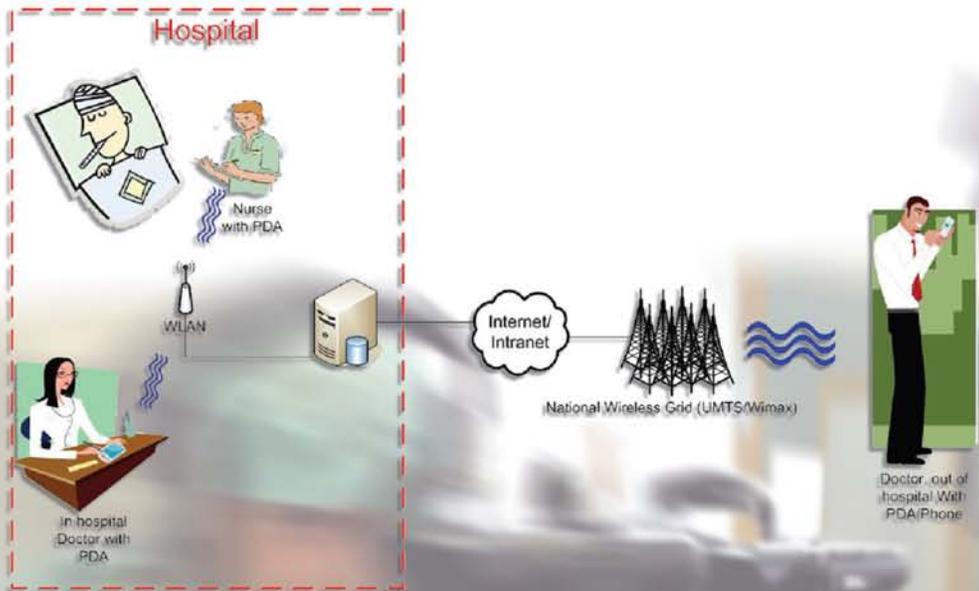


Mobile Telemedicine

A Computing and Networking Perspective



Edited by
Yang Xiao
Hui Chen

 CRC Press
Taylor & Francis Group
AN AUERBACH BOOK

Mobile Telemedicine

**A Computing
and Networking Perspective**

Mobile Telemedicine

A Computing
and Networking Perspective

Edited by
Yang Xiao
Hui Chen



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

AN AUERBACH BOOK

Auerbach Publications
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2008 by Taylor & Francis Group, LLC
Auerbach is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-6046-1 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The Authors and Publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Mobile telemedicine : a computing and networking perspective / editors, Yang Xiao and Hui Chen.

p. ; cm.

Includes bibliographical references and index.

ISBN-13: 978-1-4200-6046-1 (hardcover : alk. paper)

ISBN-10: 1-4200-6046-5 (hardcover : alk. paper)

1. Telecommunication in medicine. 2. Mobile communication systems. 3.

Wireless communication systems. I. Xiao, Yang, 1966- II. Chen, Hui, 1972-

[DNLM: 1. Monitoring, Ambulatory. 2. Telemedicine. 3. Computer

Communication Networks--organization & administration. W 83.1 M687 2008]

R119.9.M63 2008

610.285--dc22

2007050727

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the Auerbach Web site at
<http://www.auerbach-publications.com>

Contents

Prefaceix
Acknowledgmentsxi
About the Editors xiii
Contributorsxv

1 PATIENT CARE AND MONITORING

Chapter 1

Personal Supervision and Alarming Systems3
ETIENNE HIRT AND MICHAEL SCHEFFLER

Chapter 2

Integrated Alarm Monitoring System in the ICU29
A. MURAKAMI, M. AKUTAGAWA, Y. OHNISHI, Y. KURODA, AND
Y. KINOCHI

Chapter 3

Remote Wireless Patients' Data Access System49
ZIAD HUNAITI, AMMAR RAHMAN, GREGORY SVELIS,
ZAYED HUNEITI, AND WAMADEVA BALACHANDRAN

2 CARDIOLOGY

Chapter 4

Mobile, Secure Tele-Cardiology Based on Wireless and Sensor
Networks63
FEI HU, LAURA CELENTANO, AND YANG XIAO

Chapter 5

Monitoring and Management of Congestive
Heart Failure Patients85
SAJID HUSSAIN AND SAIRA MAJID DAR

Chapter 6

Issues in Personal Cardiac Health Monitoring with Sensor Networks	103
KATHY J. LISZKA, MALINDA J. SEVER, MICHAEL E. RICHTER, AND SUDHA BHATTARAI	

3 DIABETES

Chapter 7

Automated Blood Glucose Management Techniques through Micro-sensors	119
FEI HU, MICHAEL LEWIS, AND YANG XIAO	

Chapter 8

Mobile Telemedicine for Diabetes Care	143
IÑAKI MARTÍNEZ-SARRIEGUI, GEMA GARCÍA SÁEZ, M ^A . ELENA HERNANDO, MERCEDES RIGLA, EULALIA BRUGUÉS, ALBERTO DE LEIVA, AND ENRIQUE J. GÓMEZ	

Chapter 9

Telemedicine: A Way to Improve Glycemic Control among Elderly Diabetics	161
SHEILA BLACK	

4 SECURITY AND PRIVACY IN TELEMEDICINE

Chapter 10

Security and Privacy in Mobile Telemedicine	175
JUNGWOO RYOO, YOUNG B. CHOI, AND TAE HWAN OH	

Chapter 11

Security of Body Sensor Networks	195
SHU-DI BAO, CARMEN C.Y. POON, AND YUAN-TING ZHANG	

Chapter 12

A Survey of Security in Telemedicine with Wireless Sensor Networks	209
DAISUKE TAKAHASHI, YANG XIAO, AND FEI HU	

Chapter 13

Power Management and Security in IEEE 802.15.4 Clusters: How to Balance?	237
FERESHTEH AMINI, MOAZZAM KHAN, AND JELENA MIŠIĆ	

5 NETWORKING SUPPORT

Chapter 14

Fourth Generation Heterogeneous Wireless Access Networks for eHealth Services: Architecture and Radio Resource Management267
 DUSIT NIYATO, EKRAM HOSSAIN, AND JEFFREY DIAMOND

Chapter 15

3G/WLAN Cross-Layer Design for Ultrasound Video Transmission in a Robotic Tele-Ultrasonography System.....297
 MARIA G. MARTINI, ROBERT S.H. ISTEPANIAN, MATTEO MAZZOTTI, AND NADA PHILIP

Chapter 16

Enabling Mobile Adaptive Computing Environments in Teleteaching and Telemedicine Applications 319
 TUAN CAO-HUU

Chapter 17

Building a Mobile Healthcare Network within Public Networking Infrastructures339
 ZIAD HUNAITI

6 CHALLENGES AND OPPORTUNITIES

Chapter 18

Telemedicine Research: Opportunities and Challenges.....349
 PENNIE S. SEIBERT, TIFFANY A. WHITMORE, CARIN M. PATTERSON, CAITLIN C. OTTO, PATRICK D. PARKER, NICHOLE WHITENER, MICHAEL J. WARD, JEAN BASOM, AND CHRISTIAN G. ZIMMERMAN

Chapter 19

Conventional Telemedicine, Wireless Telemedicine, Sensor Networks, and Case Studies367
 LAUREN BIGGERS, YANG XIAO, AND FEI HU

Chapter 20

Telemedicine for Pervasive Healthcare.....389
 QUINTON ALEXANDER, YANG XIAO, AND FEI HU

Index405

Preface

Wireless and mobile telemedicine has drawn attention from health care providers and recipients, governments, industry, and researchers. Though various practices have been exercised, the realization of telemedicine depends on advances in computing and networking techniques. In recent decades technological development in computing and networking has largely made the delivery of health services, including medical diagnosis and patient care, possible from a distance. Many funded projects have evaluated the use of communications technology in the implementation and performance of telemedicine activities, and examined the impact of telemedicine on medical care in terms of cost, quality, and access. Telemedicine has become a growing new interdisciplinary field, which will eventually contribute to improving the quality of health care for everyone. However, successful implementation of this vision depends not only on innovative telemedicine applications but also networking and computing technical readiness. Furthermore, many ethical, social, and political problems arising in telemedicine need technical solutions.

This book studies computing and networking problems which arise from wireless and mobile telemedicine. It is a contribution of many prominent researchers working on the telemedicine field around the world. The book is divided into six parts: patient care and monitoring, cardiology, diabetes, security and privacy in telemedicine, networking support, and opportunities and challenges, including 20 chapters on a wide range of topics associated with novel telemedicine applications and pertinent networking and computing techniques. The book will shed light on future research of these areas. This book will serve as a good reference for researchers to know the state of the art and to discover uncovered territory and develop new applications, especially from a networking and computing perspective.

Acknowledgments

This book is made possible by the great efforts of our publishers and contributors. First, we are indebted to the contributors, who have sacrificed many days and nights to put write these excellent chapters for our readers. Second, we owe our special thanks to our publishers and staff members. Without their encouragement and quality work, this book would not have been possible. Finally, we would like to thank our families for their support.

About the Editors



Yang Xiao is currently with the Department of Computer Science at The University of Alabama. He worked at Micro Linear as a MAC (Medium Access Control) architect involving the IEEE 802.11 standard enhancement work before he joined the Department of Computer Science at The University of Memphis in 2002. Dr. Xiao is the director of W⁴-Net Lab, and was with the CEIA (Center for Information Assurance) at The University of Memphis.

Dr. Xiao is an IEEE senior member and a member of the American Telemedicine Association. He was a voting member of the IEEE 802.11 Working Group from 2001 to 2004. He currently serves as editor-in-chief for *International Journal of Security and Networks (IJSN)*, *International Journal of Sensor Networks (IJSNet)*, and *International Journal of Telemedicine and Applications (IJTA)* and as a (lead or associate) guest editor, an associate editor, or on editorial boards of many refereed journals. Dr. Xiao has also served as editor/co-editor for 11 books.

In addition, Dr. Xiao serves as a referee/reviewer for many funding agencies, as well as a panelist for the U.S. National Science Foundation (NSF) and as a member of Canada Foundation for Innovation (CFI)'s Telecommunications Expert Committee. He serves as TPC for more than 90 conferences such as INFOCOM, ICDCS, ICC, GLOBECOM, and WCNC. His research areas are security, telemedicine, sensor networks, and wireless networks. He has published more than 200 papers in major journals (more than 50 in various IEEE journals/magazines) and refereed conference proceedings related to these research areas. Dr. Xiao's research has been supported by the U.S. National Science Foundation (NSF).



Hui Chen is a geophysicist turned computer programmer and researcher. He received his M.S. and Ph.D. degrees in computer science, respectively, in 2003 and 2006 from The University of Memphis. Hui Chen is currently with the Department of Mathematics and Computer Science at Virginia State University. While he retains his interests in studying computational problems in various areas of earth science, he primarily works on computer system and networking research such as design and analysis of personal communications service systems, wireless LANs, wireless sensors, mobile/wireless distrib-

uted systems, and cache systems for wireless systems. He is an author of more than 30 scientific papers and articles. He serves as a guest editor for *EURASIP Journal on Wireless Communications and Networking*, special issue on “Wireless Telemedicine and Applications.” He is a member of IEEE and ACM.

Contributors

M. Akutagawa

Institute of Technology and Science
The University of Tokushima
Tokushima, Japan

Quinton Alexander

Department of Computer Science
University of Alabama
Tuscaloosa, Alabama

Fereshteh Amini

University of Manitoba
Winnipeg, Manitoba, Canada

Wamadeva Balachandran

School of Engineering and Design
Brunel University
Uxbridge, United Kingdom

Shu-Di Bao

Joint Research Centre
for Biomedical Engineering
The Chinese University of
Hong Kong
Hong Kong, China

Jean Basom

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Sudha Bhattarai

Department of Computer Science
The University of Akron
Akron, Ohio

Lauren Biggers

Department of Computer Science
University of Alabama
Tuscaloosa, Alabama

Sheila Black

Department of Psychology
University of Alabama
Tuscaloosa, Alabama

Eulalia Brugués

Endocrinology Department
Hospital Sant Pau
Barcelona, Spain

Tuan Cao-Huu

York University
and
Massachusetts General Hospital
Harvard Medical School
Boston, Massachusetts

Laura Celentano

Department of Computer Engineering
Rochester Institute of Technology
Rochester, New York

Young B. Choi

Department of Computer Information
Systems and Management Science
James Madison University
Harrisonburg, Virginia

Saira Majid Dar

Wagoner Medical Center
Kokomo, Indiana

Jeffrey Diamond

TRLabs and Department of Electrical
and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada

Enrique J. Gómez

Bioengineering and Telemedicine Group
Technical University of Madrid
Madrid, Spain

Ma. Elena Hernando

Bioengineering and Telemedicine Group
Technical University of Madrid
Madrid, Spain

Etienne Hirt

Art of Technology AG
Zürich, Switzerland

Ekram Hossain

TRLabs and Department of Electrical
and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada

Fei Hu

Department of Computer
Engineering
Rochester Institute of Technology
Rochester, New York

Ziad Hunaiti

Faculty of Science and Technology
Anglia Ruskin University
Chelmsford, United Kingdom

Zayed Huneiti

Electrical Engineering Department
University of Hail
Hail, Saudi Arabia

Sajid Hussain

Acadia University
Wolfville, Nova Scotia, Canada

Robert S.H. Istepanian

CNIT, DEIS
University of Bologna
Bologna, Italy

Moazzam Khan

University of Manitoba
Winnipeg, Manitoba, Canada

Y. Kinouchi

Institute of Technology and Science
The University of Tokushima
Tokushima, Japan

Y. Kuroda

Institute of Technology and Science
The University of Tokushima
Tokushima, Japan

Alberto de Leiva

Endocrinology Department
Hospital Sant Pau
Barcelona, Spain

Michael Lewis

Department of Computer Engineering
Rochester Institute of Technology
Rochester, New York

Kathy J. Liszka

Department of Computer Science
The University of Akron
Akron, Ohio

Iñaki Martínez-Sarriegui

Bioengineering and Telemedicine Group
Technical University of Madrid
Madrid, Spain

Maria G. Martini

CNIT, DEIS
University of Bologna
Bologna, Italy

Matteo Mazzotti

CNIT, DEIS
University of Bologna
Bologna, Italy

Jelena Mišić

University of Manitoba
Winnipeg, Manitoba,
Canada

A. Murakami

Institute of Technology and Science
The University of Tokushima
Tokushima, Japan

Dusit Niyato

TRLabs and Department of Electrical
and Computer Engineering
University of Manitoba
Winnipeg, Manitoba,
Canada

Tae Hwan Oh

Department of Electrical Engineering
Southern Methodist University
Dallas, Texas

Y. Ohnishi

Institute of Technology and
Science
The University of Tokushima
Tokushima, Japan

Caitlin C. Otto

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Patrick D. Parker

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Carin M. Patterson

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Nada Philip

CNIT, DEIS
University of Bologna
Bologna, Italy

Carmen C.Y. Poon

Joint Research Centre for
Biomedical Engineering
The Chinese University of
Hong Kong
Hong Kong, China

Ammar Rahman

School of Engineering and Design
Brunel University
Uxbridge, United Kingdom

Michael E. Richter

Department of Computer Science
The University of Akron
Akron, Ohio

Mercedes Rigla

Endocrinology Department
Hospital Sant Pau
Barcelona, Spain

Jungwoo Ryoo

Division of Business and Engineering
The Pennsylvania State University
Altoona, Pennsylvania

Gema García Sáez

Bioengineering and Telemedicine Group
Technical University of Madrid
Madrid, Spain

Gregory Savelis

School of Engineering and Design
Brunel University
Uxbridge, United Kingdom

Michael Scheffler

QIAGEN Instruments AG
Zürich, Switzerland

Pennie S. Seibert

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Malinda J. Sever

Department of Computer Science
The University of Akron
Akron, Ohio

Daisuke Takahashi

Department of Computer Science
The University of Alabama
Tuscaloosa, Alabama

Michael J. Ward

Saint Alphonsus Regional
Medical Center
Boise, Idaho

Nichole Whitener

Saint Alphonsus Regional
Medical Center
Boise, Idaho

Tiffany A. Whitmore

Saint Alphonsus Regional
Medical Center
Boise State University
Boise, Idaho

Yang Xiao

Department of Computer Science
University of Alabama
Tuscaloosa, Alabama

Yuan-Ting Zhang

Joint Research Centre for
Biomedical Engineering
The Chinese University of
Hong Kong
Hong Kong, China

Christian G. Zimmerman

Saint Alphonsus Regional
Medical Center
Boise, Idaho

**PATIENT CARE
AND
MONITORING**

1

Chapter 1

Personal Supervision and Alarming Systems

Etienne Hirt and Michael Scheffler

CONTENTS

1.1 Introduction.....	5
1.2 Target Groups and Suitable Parameters to Be Measured.....	5
1.2.1 Target Groups	6
1.2.1.1 Elderly Person Surveillance	6
1.2.1.2 Post-Trauma Care	6
1.2.1.3 Personal Health Devices	7
1.2.2 Vital Parameters to Be Measured	7
1.2.2.1 Pulse and Heart Rhythm.....	7
1.2.2.2 Movement/Fall.....	7
1.2.2.3 Temperature	7
1.2.2.4 Skin Humidity	8
1.2.2.5 Blood Pressure	8
1.2.2.6 Pulse Wave Velocity (PWV).....	8
1.2.2.7 Blood Oxygen Saturation (SpO ₂)	9
1.2.2.8 Electrocardiogram (ECG)	9
1.2.2.9 Blood Values.....	9
1.2.3 Design Guidelines and Challenges of Wrist Wearable Devices.....	9

1.3 Technologies for Wrist Wearable Devices.....	10
1.3.1 Sensors to Be Integrated into Wrist Wearable Devices.....	10
1.3.1.1 Pulse	10
1.3.1.2 Movement/Fall	12
1.3.1.3 Skin Temperature.....	12
1.3.1.4 Skin Humidity.....	13
1.3.2 Location.....	13
1.3.2.1 Indoor Location.....	13
1.3.2.2 Horizontal Position	16
1.3.3 Additional Sensors for Further Patient Surveillance	16
1.3.3.1 Blood Pressure	16
1.3.3.2 Blood Oxygen Saturation	16
1.3.3.3 ECG	17
1.3.3.4 Pulse Wave Velocity (PWV)	17
1.3.4 Networking and Communication Technologies	18
1.4 System Examples.....	18
1.4.1 Commercially Available Devices.....	18
1.4.1.1 OMRON Medical Home-Use Devices	18
1.4.1.2 BodyMedia® Lifestyle Monitoring	19
1.4.1.3 Tunstall Supervision Approach	19
1.4.2 A MON Approach	20
1.4.2.1 Wrist Device.....	20
1.4.2.2 Sensors and Clinical Results.....	21
1.4.2.3 Infrastructure	22
1.4.3 E MERGE Approach	22
1.4.3.1 Wrist Device and Its Integrated Sensors.....	23
1.4.3.2 Other Sensors and Signs	24
1.4.3.3 Infrastructure	25
1.5 Conclusions and Outlook	25
1.5.1 Tunstall Solution	25
1.5.2 A MON Approach	26
1.5.3 Fusion Approach.....	26
Acknowledgments	26
References	27

This chapter describes the motivation for personal supervision and alarming systems, the vital parameters to be measured, and the wearable and environmental sensors required. Such systems are found to be suitable for the supervision of elderly people, outpatients, high-risk populations (i.e., cured cardiac patients), and health-conscious people. The different solutions are discussed and assessed.

The fusion of these approaches is proposed for working out an optimal solution based on a sophisticated wrist wearable device measuring vital parameters for

supervision but not diagnostic. The wrist device is supported by a minimal environmental installation.

1.1 Introduction

Nowadays in the Western world, more and more elderly people live in their own households and without any relatives in their vicinity. Although their physical fitness is continuously decreasing, they want to keep their personal autonomy and want to stay in their own environment as long as possible without sacrificing their medical safety.

Moreover, post-operative intensive care in hospitals is minimized as much as possible to reduce overall health costs, thus causing an increase in hospital outpatients. Additionally, the general population of high-risk patients (e.g., those exhibiting a cardiac deficiency) is becoming larger and larger.

To support all these groups in their daily life and to improve their personal well-being at reasonable cost, personal supervision and alarming devices can fill a gap to alert relatives or activate neighborhood or emergency medical services (EMS) early enough in case of a medical incident. Such personal devices can replace part of the capabilities of a nurse or a caring family member in the same household—the automated version that can identify a medical condition and trigger help.

Apart from the obvious societal benefit, also from a commercial perspective such devices are very interesting: the target groups present a significant emerging market. However, it is not a single mass market because the required solutions are still specific to the environment and the target group. Therefore this fact also opens business opportunities for niche players and small- and medium-size enterprises.

This chapter describes sensor technologies suitable for integration into personal supervision and alarming wrist devices, and location and networking technologies required. We present examples of the medical approach of the AMON research project and the network approach of the EMERGE research project, and we propose a fusion approach of both solutions.^{1,2}

Cost-benefit calculations are not included in this chapter because they very much depend on usage and the reimbursement of insurances; also descriptions, requirements, and examples for corresponding infrastructures such as telemedicine centers can be found elsewhere.

1.2 Target Groups and Suitable Parameters to Be Measured

The user group of personal supervision and alarming devices is far from being homogeneous. In order to identify their requirements, this section defines typical target applications, their purpose, and the vital parameters to be measured. Also, we present general design guidelines for wearable devices.

1.2.1 Target Groups

For the purpose of this chapter, we divide the target groups into:

- Elderly person surveillance
- Post-trauma care
- Personal health devices

The requirements of these groups are described in the following subsections.

1.2.1.1 Elderly Person Surveillance

A lot of elderly people would like to live in a self-determined way in a familiar environment as long as possible. As long as a person is not living alone, the elderly supervise each other constantly and therefore are not ready for technical supervision. But as soon as a person is living alone at home or even in a managed care facility, an automated alarm system* at least is strongly recommended.

For elderly people not visited daily by relatives or neighbors further supervision of drinking, eating, mobility, and medication should be performed to recognize typical problems early.

In addition, any people having certain health problems should be supervised specifically, i.e., regular ECG and/or blood pressure measurements as described in the further subsections.

1.2.1.2 Post-Trauma Care

People often have to stay in hospital for supervision only. Health costs can be reduced if these people could be released to go home or even to work, provided they could wear the required supervision and alarming equipment. Such equipment has to provide medical measurements comparable to hospital standard, which is really a challenge.

High-risk populations such as cardiac patients after successful convalescence should be supervised with the same devices. An example for this target group is described in Section 1.0.

* An automated alarm system automatically alerts a caregiver, relative, or the EMS on the suspicion that something may be wrong.³ These systems can continuously monitor variables sensitive to changes in functional health status and behavior. They generate an alarm when significant changes are observed. They can be as simple as consisting of combined movement detectors only. With a n intelligent decision-support system using robust algorithms, false alarms are unlikely. These systems are most likely fully integrated within a home network.

1.2.1.3 Personal Health Devices

People interested in continuous supervising of their health are possibly a third target group. Their requirements are lower than those of the other groups, because they neither require an excellent alarming system nor medical-grade measurements. Their focus is rather on unobtrusive devices with many features. Typical examples are pulse watches such as the well-known Polar heart rate monitors.

1.2.2 Vital Parameters to Be Measured

The following vital parameters are assessed for supervision of the target groups.

1.2.2.1 Pulse and Heart Rhythm

Measuring the pulse can give very important information about the health of a person. Any deviation from normal heart rate can indicate a medical condition. Fast pulse may signal the presence of an infection or dehydration.⁴ In emergency situations, the pulse rate can help determine if the patient's heart is pumping and allows detecting a cardiac arrest immediately. Furthermore, pulse measurement during or immediately after exercise can give information about the fitness level and the health of a person. A pulse measurement shall be provided for all three target groups.

ECG-type techniques additionally allow measuring the R–R interval (peak-to-peak time in ECG) and provide better supervision because arrhythmias show cardiac problems that are not recognized with pulse measurements. However, the rhythm is usually only required for post-trauma care and often a noncontinuous measurement is sufficient.

1.2.2.2 Movement/Fall

Movement supervision is very important for an automatic alarm device. It enables the device to recognize a fall or a situation where the patient is helpless, and calls for assistance without user intervention. This supervision capability is required especially for elderly people surveillance.

1.2.2.3 Temperature

Of medical relevance is the body (core) temperature, but because fever is usually only an additional medical sign, it is not very important. Body temperature measurement is not required for any of the target groups.

It is much easier for a device to measure the skin temperature, a mixture between body temperature and environment temperature. The skin temperature

has no medical value but provides a context for the other vital sensors required, e.g., for temperature compensation.

1.2.2.4 *Skin Humidity*

Skin humidity is hardly measured by medical people but according to the Mayo Clinic,⁵ it would be a good parameter to monitor the (de-)hydration of people and would therefore be recommended for the supervision of elderly people. However, the physicians within the EMERGE² project do not support this recommendation.

1.2.2.5 *Blood Pressure*

High blood pressure increases the risk of heart failure, heart attack, stroke, and kidney failure. For people who have high blood pressure, this test is a way of monitoring the effectiveness of medications and dietary modifications.⁴

Low blood pressure may be a sign of a variety of illnesses, including heart failure, infection, gland disorders, and dehydration.

Repeated measurements are important. A single high measurement does not necessarily mean that the patient has a high blood pressure condition. On the other hand, a single normal measurement does not mean that the patient has no such condition.

Blood pressure readings taken at home can provide important information to the doctor. Such readings may be a better measure of the current blood pressure than those taken at the doctor's office where people are known to become nervous ("white-coat hypertension"), resulting in higher readings than normal.

Continuous blood pressure measurement is recommended for cardiac outpatients but noncontinuous measurements are sufficient and suitable for all target groups.

1.2.2.6 *Pulse Wave Velocity (PWV)*

The pressure pulse travels much faster than the blood itself. PWV describes how quickly a blood pressure pulse travels from one point to another in the human body. The time difference between these two locations is known as the pulse transit time (PTT). PWV is typically measured between the carotid and the femoral artery. Atherosclerosis causes the arterial walls to become thicker and harder, and narrows the arterial lumen. The increased inflexibility of the arterial walls increases PWV, because the energy of the blood pressure pulse cannot be stored in an inflexible wall. PWV can be used as an index of arterial distensibility.

In terms of medical diagnosis, PWV is a highly interesting subject because it provides an estimate of the condition of the cardiovascular system based on a large area of the human body.

Furthermore, as blood pressure is essentially sensitive on the pulse wave velocity, the velocity pulse, and the arterial diameter, it can be calculated from PWV after an individual initial calibration by means of a standard blood pressure meter.

1.2.2.7 Blood Oxygen Saturation (SpO₂)

The SpO₂ measurement is routinely used for patient supervision as well as supervision during surgery. An SpO₂ monitor is an easily installable device that can be used by rescue teams for patient survey or even triage.⁶ Although the measurement provides an easy installable pulse measurement as well as supervision of the oxygen saturation during surgery, the main problem is that the sensor itself requires supervision to ensure proper installation to avoid false readings and alarms. Therefore, continuous SpO₂ measurement is only recommended for post-trauma care.

1.2.2.8 Electrocardiogram (ECG)

An ECG is very useful in determining whether a person has heart disease. If a person has chest pain or palpitations, an ECG is helpful in determining if the heart is beating normally. If a person is on medications that may affect the heart or if the patient is on a pacemaker, an ECG can readily determine the immediate effects of changes in activity or medication levels. Some heart conditions are not detectable all the time, and others may never produce any specific ECG changes. A person who suspects heart disease or has had a heart attack may need more than one ECG.⁴

For cardiac outpatients as well as long-term measurements for heart disease investigations, ECG measurement is required continuously. For all other target groups, no ECG is required.

1.2.2.9 Blood Values

Blood values such as glucose concentration and others are only required for specific patients. These values can hardly be measured without taking blood samples or other (semi-)invasive techniques. Noninvasive measurement⁷ methods are still subject to further research and development and are therefore not further considered in this chapter.

1.2.3 Design Guidelines and Challenges of Wrist Wearable Devices

The developer of wearable medical devices (WMD),⁸ compared to stationary equipment, has to take additional user requirements into account. Also, instead of first designing a functional prototype and then making it wearable, both tasks need to be tackled concurrently, in order to avoid costly redesigns.

Small and lightweight: To suit the size of a forearm, typical dimensions would be about $60 \times 50 \times 5$ mm.⁹ Therefore, the inner dimensions are fixed and volume/weight restrictions apply. These restrictions require mechanical and electrical co-design throughout development. The WMD needs to be unobtrusive in order to be worn as a daily accessory without looking like a medical device.

Low power: Power is required for at least one working day without recharging. For an application with very low power consumption, a primary battery is suitable. Otherwise, a secondary/rechargeable accumulator is required.

Life cycle: High reliability and a minimum four-year field life are necessary in order to be eligible for possible reimbursement by health insurance plans.

Housing: The device needs to be shockproof and must be biocompatible where exposed to the user.

Input–output connection: If a plug and socket are chosen for input–output, there are mechanical issues and relatively large and expensive hardware is required. If wireless connections (see Section 1.3.4) are used, they will require much more power, contradicting the above-mentioned low power postulation.

Sensors: Novel applications depend on new sensor concepts, which cannot easily be integrated into standard electronics or housings. Also, where direct physical contact with the user is required, biocompatibility issues may influence the sensor principles and signal post-processing.

Sensor technologies and the suitable approaches for integration are analyzed in the following section.

1.3 Technologies for Wrist Wearable Devices

The user shall wear one device that incorporates the required and feasible sensors. Other sensors shall be attached to this device or to the infrastructure. The preferred embodiment of this device is a watch-like device that can be worn day and night. This device is also responsible for communication with any room installation.

Besides vital parameters, movement and location of a person shall also be supervised in order to detect emergency situations.

1.3.1 Sensors to Be Integrated into Wrist Wearable Devices

1.3.1.1 Pulse

State-of-the-art pulse measurement methods are

- ECG-type techniques
- Transmissive PPG (photoplethysmography) on ear, finger, or foot, reflective PPG on the forehead
- Pulse detection during oscillometric blood pressure measurement

Because all the above methods have their disadvantages, alternative methods such as using a reflective PPG on the wrist or capacitive or pressure sensor-based approaches should be investigated and research effort invested in order to develop a sensor technology that can be integrated into a wrist device.

1.3.1.1.1 ECG-Type Technique

Sensors for measuring heart activity continuously (such as Polar heart rate monitors) are built into separate chest-worn belts or integrated into a shirt. This shall be avoided for all applications but for cardiac outpatients. A further disadvantage is that even if the electrical signal for the heart is present the heart might no longer or only insufficiently transport the blood. This can only be detected by measuring the pulse that is not present or weak. This failure is not detected with the ECG-type technique. However, the ECG measurement technique can be used for heart rate measurement on demand by pressing a finger of the other hand onto a wrist device. This approach is not recommended for integration into a wrist device.

1.3.1.1.2 Photoplethysmography (PPG)

The commercially available devices require a non-optical measurement at the earlobe (www.sentec.ch, www.csem.ch), a finger clip (ChipOx module from www.corescience.de, www.spo-medical.com, and others), or the measurement might be integrated into a finger ring (MIT⁹). All successful devices use the transmission PPG configuration (left side of Figure 1.1).

1.3.1.1.3 Oscillometric Blood Pressure Measurement

During oscillometric blood pressure measurement by means of a cuff, provided by Omron and other manufacturers, the pulse is also measured. However, these devices are intended for periodic use and not for continuous wear. Further, sufficient battery power is required (11 mAh per measurements for pump and valve) to pump

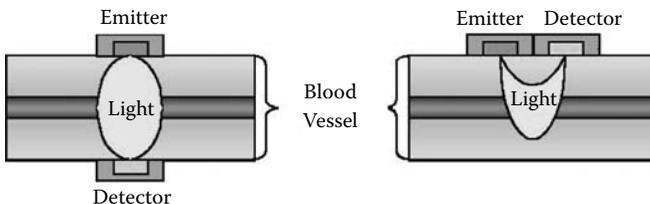


Figure 1.1 Transmissive versus reflective PPG measurement for pulse and oxygen saturation.

up the cuff for blood pressure measurement. This technique is not recommended for integration into a wrist device because an obstructive cuff is required.

1.3.1.1.4 Capacitive or Pressure Sensor-Based

In order to integrate a pulse measurement into a wrist device using minimum power, the most promising approaches are piezoelectric pressure sensors and capacitive pulse measurements quite similar to the pressure sensor approach. These approaches demonstrated promising results in prototype pre-tests but they also showed that the signal-to-noise ratio and the sensor placement are very critical.

1.3.1.2 Movement/Fall

Fall detection is provided by several commercial devices. The only known approach at the wrist is a research device of ETH¹⁰. According to their findings, fall detection cannot be 100% warranted because it causes too many false alarms (e.g., hits of the arm on a table). This sensor requires further research in order to identify positively a situation in which a person is rendered helpless. The most difficult situations to detect are faints, falls against a wall, and unconscious and seated, as illustrated in Figure 1.2.

1.3.1.3 Skin Temperature

Skin temperature measurement can easily be integrated into a wrist device by means of an analog or digital off-the-shelf temperature sensors. To do so, a good thermal path between the skin and the sensor has to be ensured through the enclosure. However, as skin temperature is different from body temperature, this information can be used only to recognize if the device is worn.

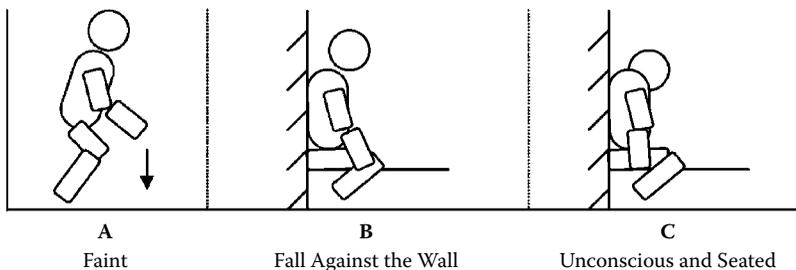


Figure 1.2 Difficult-to-detect fall situations. (Source: Noury, N., Barralon, P., Virone, G., Boissy, P., Hamel, M., and Rumeau, P., A smart sensor based on rules and its evaluation in daily routines, *Proceedings of the IEEE EMBS 2003*. With permission.)

1.3.1.4 Skin Humidity

Skin humidity can be estimated by electrical skin impedance spectroscopy.¹² The measurement frequencies range from 10 Hz to 5 MHz with a driven current of 1 μ A to 5 mA. However, most approaches use 1 to 1000 kHz with a current below 1 mA. The frequency selection depends on the target to be measured. In Martinsen, Grimness, and Haug¹³ it is found that the impedance measured mainly reflects the stratum corneum at low frequencies, in this case below 1 kHz, and that the viable skin dominates at higher frequencies. At 1 MHz the stratum corneum only accounts for 5.4% of the measured impedance.

The used electrodes are either one concentric ring electrodes¹³ or a four-point measurement with two driving electrodes and two measurement electrodes.

1.3.2 Location

In outdoor environments, we can obtain precise location information ranging from 1 to 10 m from GPS (global positioning system) but only at the price of significant power consumption.* But with GPS alone, only self-location is possible. For location by emergency personnel, this information then needs to be transmitted.

1.3.2.1 Indoor Location

For indoor location, GPS availability is limited. Besides intelligent (sensor-equipped) carpets, the following common transmission technologies can be used:

- Ultrasonic
- Infrared
- Microwave
- RF

The transmission-based technologies (RF) are all active technologies requiring the user to wear a device. A completely different technology to be mentioned is the use of cameras.

1.3.2.1.1 Passive Infrared (PIR) Motion Detection Sensors

Infrared systems rely on the user taking explicit actions to identify their presence. The sensor measures the changes in the received infrared signal radiated from every human being as well as animals. These sensors are very inexpensive but usually require movements, at least small ones, and they often detect pets, too.

* 3 mAh for one position per minute and 35 mAh for one position per second based on the ublox LEA 4A module.

A usual range is around 10 m, but the covering area cannot be well adjusted. Wireless ZigBee PIR sensors are available.

1.3.2.1.2 Microwave Motion Detection Sensor

Microwave sensors are usually based on Doppler shift radar technology in order to detect movements. The required movements can be as small as 6 mm per sec, thus recognizing a person sitting.

Unlike passive infrared detectors, the low power radar beam may be aimed to cover a specific area. Also, the range, or sensitivity, is easily adjusted over a range of approximately 1 to 10 m.

1.3.2.1.3 RF Received Signal Strength Indication (RSSI)

By measuring the received signal strength from several fixed nodes, a location accuracy of 1.6 for 80% of the measurements was achieved.¹⁴ The accuracy mainly depends on the algorithm and fingerprinting and can further be improved by utilizing sophisticated antennas that have a uniform radiation pattern and by considering that the best distance measurement accuracy is achieved within proximity of a fixed node due to exponential decrease of the received signal strength. According to theoretical analysis by TI with their CC2430, this is within 7 m from any fixed node as shown in Figure 1.3.

This technique is promising primarily if an RF sensor network is installed in an apartment providing fixed nodes for free. The disadvantage is that due to multipath reflections and other noise sources the RSSI value is quite noisy and might be smaller than expected from the distance due to unexpected damping.

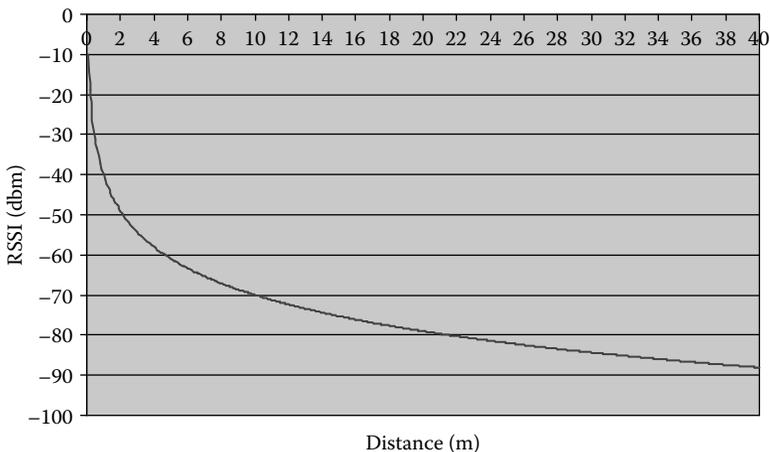


Figure 1.3 RSSI versus distance.

1.3.2.1.4 RF Angle of Arrival

Another option for determining an active nodes position is to measure the angle of arrival of an RF signal and perform a triangulation. To do so, sophisticated antennas and receiving circuits are required that are very different from the transceiver systems used in off-the-shelf sensors with RF communication. Therefore, this approach is not recommended.

1.3.2.1.5 RF Ultra-Wide Band

Based on the IEEE802.15.4a ultra-wideband (UWB) technology, first products are available for very accurately determining the location of an object or a person. With a combined approach that uses both time difference of arrival (TDOA) and angle of arrival (AOA) only two to four fixed nodes are required to reach a standard deviation of the position as low as 15 cm.¹⁵ The UWB approach is much less affected by multipath distortion than conventional RF systems and the calculation is not based on signal strength.

UWB technology can also be used for data communication, which makes it very attractive at first glance. Unfortunately, components for integration into a wrist wearable device are not yet available and the stand-alone devices are expensive. Therefore this technology can only be used as an additional device worn separately, which is not recommended.

1.3.2.1.6 Ultrasonic

Ultrasonic devices offer a low-cost solution providing a high accuracy of measured positions. Accuracy varies between 0.03 to 0.5 m depending on the specific system.

The clear disadvantage of ultrasonic waves compared to an RF-based location sensing system is that infrastructure is more complicated and needs to be installed very precisely. In fact, all narrowband ultrasonic positioning systems are faced with the problem of loss of signal due to obstruction, false signals by reflections, and interference from high-frequency sounds. However, most of these limitations can be reduced through careful planning, resulting in a highly accurate system. But as this is a completely separate infrastructure in addition to other sensor installations, it is not recommended for this application. For distance measurement, however, it is a very promising approach.

1.3.2.1.7 Camera

Location tracking can also be performed by means of a camera. It further allows a good insight in case of an emergency but will not be investigated further in this chapter.

1.3.2.2 Horizontal Position

A specific location task is a horizontal position determination, a critical item in order to distinguish between someone lying on a bed or chair (normal for a certain time) and lying on the floor.

Currently there is no specialized method to determine the height of a person relative to the ground. Even worse, when using triangulation with signal strength indication, it is a requirement that the floor, where the person is, is estimated. This allows taking into account the signal loss of an RF signal passing through the ceiling, usually built of concrete and steel.

A novel approach is currently investigating the possibility of using an altitude sensor for horizontal position tracking. A change of 4 Pa is equivalent of 27 cm and the noise of the sensor is ± 1 Pa. Thus, an accuracy of better than ± 10 cm was found. However, these are only first trials and many effects such as several rooms, several base stations, sensor orientations, etc., have to be investigated before such an approach is ready for implementation.

Besides the air pressure sensor in the wearable device, a base station with a known “height” is required to track the environmental pressure that can change rapidly, dependent on the weather condition.

1.3.3 Additional Sensors for Further Patient Surveillance

The additional sensors might be added separately into a sensor network or communicate their readings to the wearable device. It is, however, better not to integrate them into it. This is mainly recommended for medical-grade vital parameter measurement.

1.3.3.1 Blood Pressure

A blood pressure sensor shall not be integrated into the wrist device because no ready-to-use technology is available. A big research effort would be required. However, there are commercially available stand-alone devices available that might be integrated into a sensor network.

1.3.3.2 Blood Oxygen Saturation

The technology is well known¹⁷ but not easy to implement on a wrist. For reliable measurements, a transmissive measurement device (Figure 1.1.) shall be used. As SPO_2 saturation measurement is not required for continuous supervision, it is recommended to add this parameter by means of a commercial device referenced in Section 1.3.1.1.2.

1.3.3.3 ECG

A one-lead ECG can be integrated into a wrist device by providing one electrode on the device bottom and a second on the device top. The top electrode shall be in contact with a finger of the second hand in order to measure the derivation I. A similar result is achieved with the belt described in Section 1.3.1.1.1.

For reliable and medical-grade measurements including more derivations, glued electrodes are required but not appropriate for a wrist device. An alternative approach is the newly developed dry electrode technology that can be built into common items of clothing like bras, shorts, or waist belts.¹⁸ Because the clothing-integrated ECG is not yet commercially available, a stand-alone device is recommended even for target groups requiring a continuous supervision.

1.3.3.4 Pulse Wave Velocity (PWV)

In principle there are three methods to determine pulse transit time (PTT) used to calculate PWV:

1. ECG pulse to laser Doppler flow pulse on arm or leg as shown in Figure 1.4
2. ECG pulse to PPG pulse on arm or leg
3. Time between two PPG pulses or laser Doppler flow pulses measured at least 100 mm apart on arm or leg

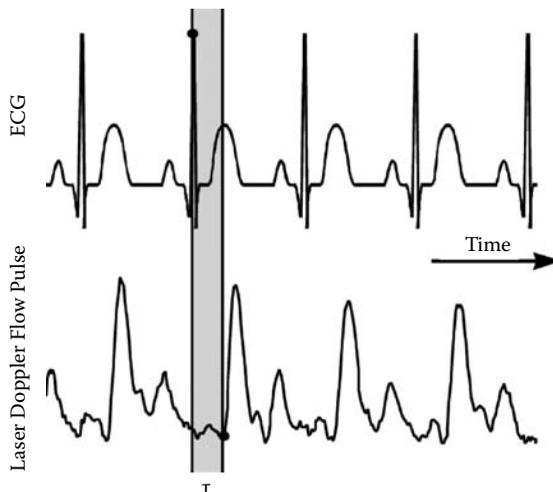


Figure 1.4 Principle of the determination of the pulse transit time using an ECG R-wave and the footpoint of a laser Doppler flow pulse. (Source: Elter, P. et al., *Noninvasive and nonocclusive determination of blood pressure using laser Doppler flowmetry*, *Proceedings of SPIE Specialty Fiber Optics for Medical Applications*, 3596: 188–196. With permission.)

The laser Doppler flow meters are very large and therefore not suited for integration. Methods 2 and 3 might be used as a combination of the sensors described in Section 1.3.3.2 and Section 1.3.3.3.

1.3.4 Networking and Communication Technologies

There are many communication technologies available. The analysis required to select the appropriate one is not the intention of this subsection—instead we only provide an overview.

For long-range communication a device has to rely on available infrastructure. Therefore, a cellular standard such as GSM/UMTS/CDMA2000 is the best choice. However, any cellular standard requires regular power in the range of 30 mA* and therefore a large battery that makes a device obstructive and heavy.

For indoor and local communication the following standards might be considered:

- 433/868/2400 MHz proprietary ISM band communication or 869 MHz European social alarm frequency
- ZigBee
- Bluetooth
- UWB

The most promising standard is ZigBee because it is a nonproprietary standard, low power, and fully integrated components are available. It can further be used for localization as described in Section 1.3.2.1.3. UWB might be considered as soon as the required components are available.

1.4 System Examples

For supervision of the target groups many research and commercial approaches exist or are ongoing. The purpose of this section is to provide a few examples. There are some further commercial approaches and many other research projects other than the two described in detail here.

1.4.1 Commercially Available Devices

1.4.1.1 OMRON Medical Home-Use Devices

OMRON does not provide a supervision system but provides medical devices for home use for measuring blood pressure, one-lead ECG, and digital temperature. These devices are a good supplement for other approaches that require supervision.

* Telit GE 864PY GSM/GPRS modem assuming 1.2 min data transmission per hour but being logged-in all the time.

1.4.1.2 BodyMedia® Lifestyle Monitoring

The SenseWear® armband from BodyMedia, worn on the back of the upper arm, provides personal metabolic, physical activity, and lifestyle monitoring. The integrated sensors and algorithms determine the following parameters:

- Total energy expenditure (calories burned)
- Active energy expenditure
- Physical activity duration
- Number of steps
- Sleep duration

The data collected on the device can be transferred to a physician by means of a USB connection to the PC. The analysis is therefore a posteriori without a new alarming function.

1.4.1.3 Tunstall Supervision Approach

Tunstall provides a broad range of (environmental) sensors, as summarized in Section 1.4.1.3.1). These self-contained sensors raise individual alarms that are collected at a central device. The interconnection of these sensors is a proprietary wireless solution on the 869 MHz European social alarm frequency.

1.4.1.3.1 Sensors and Signs

The Tunstall solution uses the sensors listed below in order to supervise a person. Their description is based on Tunstall's Website.²⁰ Tunstall further provides environmental sensors for gas alarm, smoke alarm, etc. (not listed).

Bed/chair occupancy sensor: Provides an early warning by alerting that the user has left his bed or chair and not returned within a preset time period, indicating a potential fall. The sensor can also be programmed to switch on lights.

Enuresis sensor: Placed between mattress and sheet, this sensor provides immediate warning on detection of moisture, allowing effective action to be taken.

Epilepsy sensor: This sensor monitors the user's vital signs including heart rate and breathing patterns to detect a range of epileptic seizures.

Fall detector: See Section 1.4.1.3.2 for a description.

Medication dispenser: Automatically dispenses medication and provides audible and visual alerts to the user each time medication should be taken. If the user fails to access the medication, an alert is raised to the monitoring center or designated care giver.

Movement detector: A passive infrared (PIR) movement detector that can be used for both activity and inactivity monitoring.

Pressure mat: Detects if somebody is in a specific area.

Door opening sensor: Monitors, for example, the main exit door.

Wireless alarm button: Wearable on the neck or to be installed at fixed locations.

Vibration alarm: Designed to support hearing-impaired people, the device vibrates to provide a smoke alarm to a sleeping user.

1.4.1.3.2 Wearable Devices

Tunstall provides two wearable devices. One is only a wireless alarm button. The other one is a fall detector that is part of a fall management system. The fall detector is worn on the belt or in a pocket around the waist and will automatically raise an alarm if the unit senses a fall. The detector senses both impact and angle in order to distinguish between normal impact (user is vertical) and fall. Trials showed that this sensor is able to detect the worst falls but we must assume that this approach is not able to detect the falls sketched in Figure 1.2.

A further disadvantage is that it is a single function device only and it does not enable further reasoning together with other sensors on the body or in the apartment.

1.4.1.3.3 Infrastructure

All the sensors are wirelessly attached to a main unit called a lifeline. Upon alarm of any sensors a call of the 24-h response center or a local alarm is initiated based in the main unit's configuration.

1.4.2 AMON Approach

Designed to be worn by cardiac outpatients and high-risk people, the AMON device allows remote monitoring of vital signs such as blood pressure, pulse, oxygen saturation, skin temperature, and two-channel ECG signals (heart rate, QRS duration, ST-elevation, etc.). Operating autonomously, the wearer can be continuously monitored and all the data is stored on-device. A built-in expert system can issue a warning or alarm locally or to a telemedicine center if necessary; together with a built-in GSM link, doctors at the health center can download this data for immediate or later analysis and emergency care as, sketched in Figure 1.5.

After the development of AMON, Telcomed developed and now markets the WristClinic™²¹ that integrates similar sensors but uses a wireless transmission to a separate gateway for data transmission to a center. Also, an expert system is not included in this commercial device.

1.4.2.1 Wrist Device

Due to the design decision to use off-the-shelf modules with no miniaturization potential, the first prototype was rather bulky, at $6 \times 60 \times 30 \text{ mm}^3$ in size (see Figure 1.6). It consisted of 10 submodules, folded together in order to embrace

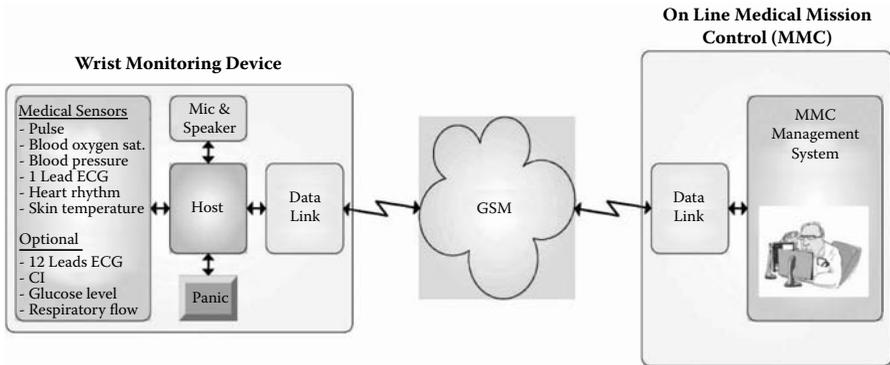


Figure 1.5 AMON system overview: Wrist worn medical device with GSM/UMTS link to the telemedicine center.

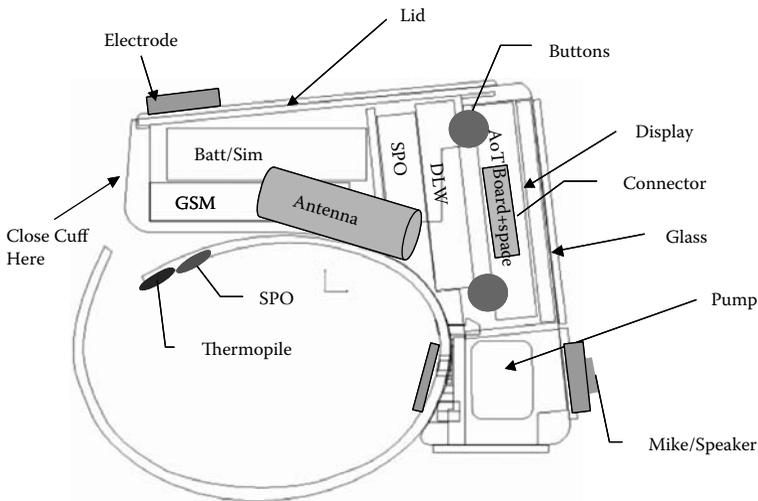


Figure 1.6 Cross-section of first (milled) AMON enclosure.

the wrist. Electronics and sensors are mounted in a plastic enclosure containing a blood pressure cuff. The enclosure was built with a rapid prototyping, laser sintering method to avoid an injection molding tool or complicated milling processes. The second enclosure version (see Figure 1.7) was an improved design.

In principle, four measuring devices plus a mobile phone can be miniaturized to meet the form factor of a conventional off-the-shelf blood pressure monitor.

1.4.2.2 Sensors and Clinical Results

The blood oxygen saturation (SpO₂) sensor is a prototype for wrist measurements, based on a reflective sensor principle (standard sensors transmit light through the



Figure 1.7 Second AMON enclosure on wrist.

fingertip). The ECG sensor and the blood pressure measurements are standard but adapted to the WMD requirements.

A thermopile was chosen for temperature measurement. This should provide better results than a simple thermistor but the relationship between the measurement, skin temperature, and body temperature is still a topic for research. Therefore, this sensor has not yet been validated.

The AMON device also included acceleration sensors but due to effort constraints no algorithm was implemented at all.

The ECG, the SpO₂ sensor, and the blood pressure meter were tested with 29 subjects, using two AMON devices. The sensors were found to be functional, but as expected the data processing algorithms will need some fine-tuning.

1.4.2.3 Infrastructure

AMON does not require any infrastructure in an apartment. All required sensors are integrated into the wrist device as described above. For alarms and data communication AMON uses the GSM network over which it sends encrypted data to a telemedicine center.

The AMON telemedicine center software provides data storage for vital signs history and display of the actual measurement values including rating, as shown in Figure 1.8.

1.4.3 EMERGE Approach

The approach in EMERGE² is to reason about situations based on information collected from ambient, unobtrusive, and noninvasive sensors in the home

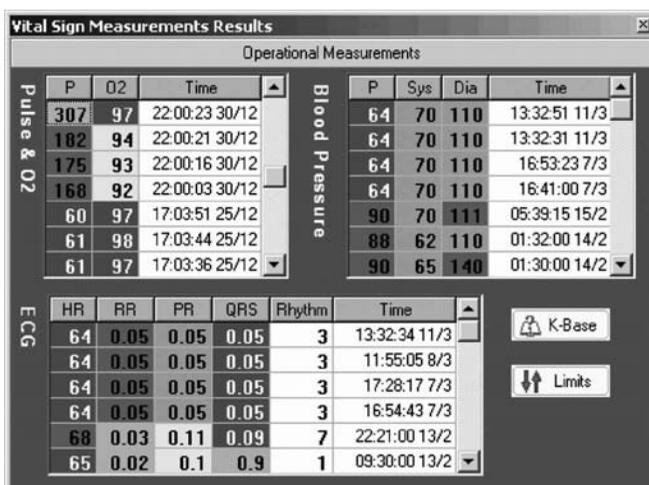


Figure 1.8 AMON telemedicine center vital sign display including rating shown in case of an alarm.

environment of elderly people. This raises the challenge to cope with inherently unreliable and imprecise data as well as the need to adapt the system to the specific conditions and demands of the assisted persons.

EMERGE tackles this problem with a n i n t e l l i g e n t s e n s o r f u s i o n a p p r o a c h in combination with a sound emergency model, describing the environment, individual diseases, parameters to monitor, potential emergency situations, and corresponding treatment options.

In case of emergency detection the system will automatically connect to an emergency dispatch center for immediate intervention if required. In case of longer-term events such as change in behavior due to depression or starting dementia the family doctor or relatives are informed.

1.4.3.1 Wrist Device and Its Integrated Sensors

The concept of EMERGE is to use environmental (installed) sensors whenever possible. The wrist device sensors are therefore reduced to the required minimum. Furthermore, no external accessories such as ECG belts shall be used but everything should, if possible, be integrated into the wrist device. The wrist device is designed to be worn day and night. The wrist device will therefore feature

- Pulse measurements by means of PPG (Section 1.3.1.1.2) or capacitive/pressure sensor (Section 1.3.1.1.4)
- Fall/movement and motionlessness detection by means of triaxial acceleration sensors (Section 1.3.1.2)

24 ■ *Mobile Telemedicine*

- Skin temperature (Section 1.3.1.3)
- Horizontal position (Section 1.3.2.2)
- Alarm button
- Rechargeable battery and watch display
- RF communication (ZigBee or similar)

and optionally:

- Skin impedance (Section 1.3.1.4)
- RSSI location (Section 1.3.2.1.3)

The wrist device controls the integrated sensors and extracts the measurement values from the raw data. In case of event detection (and also regularly), the wrist device sends measurement data to the gateway for further processing and/or alarm.

1.4.3.2 Other Sensors and Signs

In order to provide supervision of people for at least the following actions:

- Toilet usage
- Sleeping behavior
- Meal preparation
- General movement including wandering and remaining at the same place

and for supervision of their weight, the sensors described next might be installed. The preferred sensor communication (see Section 1.3.4) is the ZigBee wireless standard. For users that require regular medical-grade measurements the sensors described in Section 1.3.3 can be added.

1.4.3.2.1 Activity Sensors

- Bed occupancy sensors including vital data such as biomedical bed-clothes²²
- Chair occupancy sensors (see Section 1.4.1.3.1)
- Door exit/entry sensor (see Section 1.4.1.3.1)
- Video: Location tracking and communication with the user in case of emergency
- Pressure mat/floor sensors (see Section 1.4.1.3.1)
- Intelligent cups supervise drinking
- Intelligent walking stick raises an alarm if not vertical for a longer time in order to detect a fall
- Intelligent medicine bottle allows tracking of medicine consumed

- Location tracking (see Section 1.3.2)
- Intelligent light switches and dimmers report switching by means of RF or X-10

1.4.3.2.2 Environmental Sensors

- Temperature and humidity sensors
- Gas detector sensor/activator
- Smoke detector sensor
- Carbon monoxide detector
- Flood detector sensor/activator

1.4.3.2.3 Weight Sensors

The weight of a person is measured daily by means of a wireless scale or better by a scale integrated into the bed(-posts) or similar.

1.4.3.3 Infrastructure

As EMERGE relies mainly on environmental sensors, a major infrastructure and therefore installation is required. Most of the installations are the sensors listed in Section 1.4.3.2. In addition a central server and a gateway to it are required. The central server is further connected to communication installations for a direct connection to the local emergency dispatch center as well as telemedical center and to the family doctor and the relatives. The external centers have to provide means to integrate the extended information from the EMERGE system, as most of them are prepared for phone connections only.

1.5 Conclusions and Outlook

Supervision of elderly persons, patients in post-trauma care, and personal health devices are only partly available. There is ongoing research, also supported by the EU IST frameworks. The provided system examples are concluded below, and based on them a fusion approach is proposed.

1.5.1 Tunstall Solution

The approach provides an environmental supervision for elderly and impaired people. Their approach does not include a wrist device but is limited to two wearable devices that do not collect any vital data from the user. Besides the lack of vital data

measurement, their system is not able to provide reasoning based on long-term data and/or based on the combination of several sensors.

1.5.2 AMON Approach

The AMON approach is a serious effort to provide medical-grade vital parameter measurement. Due to required sensors the device's size is not minimal and therefore only recommended for target groups requiring this tight supervision. A major advantage of the device is the provided freedom due to the widely available GSM network. But in case of emergency, positioning based on the GSM network only is not very accurate.

1.5.3 Fusion Approach

Based on the experience in both the AMON and the EMERGE projects, we propose an approach combining the advantages without the drawbacks. On one hand the installation requirements shall be minimized and mainly rely on a wrist device. On the other hand the wrist device shall not provide medical measurements but only the required detection of alarming signs. In order to minimize size and weight a short-range communication shall be used.

The proposed wrist device shall feature the same functions as the EMERGE wrist device (Section 1.4.3.1). The RSSI location technique is not optional for the fusion approach as no other location technique shall be mandatory. However, the number of fixed nodes will hardly be sufficient for exact location derivation, but the technology shall be used to detect a wearer staying longer than usual at the same location.

The minimum installation will be a base station for the height reference (see Section 1.3.2.2) and acting as gateway to a fixed phone line, long-range communication, or any other network for dispatching an alarm or a warning. For the RSSI location technique three or more fixed nodes are required. One is the base station; further nodes can be without function or be combined with assisting sensors or with nodes used for home automation.

The system might be extended by means of a bed occupancy sensor and some PIR motion detectors to support the activity tracking in combination with the wrist device and to continue supervision at reduced quality during the night if the user does not wear the wrist device.

Acknowledgments

The AMON project (IST-2000-25239) has been funded by the EU IST FP5 program. The EMERGE project (IST-2005-045056) has been funded by the EU IST FP6 program.

References

1. Anliker, A. et al., AMON: A wearable multiparameter medical monitoring and alert system, *IEEE Transactions on Information Technology in Biomedicine*, 8(4): 415–427, 2004.
2. EMERGE, <http://www.emerge-project.eu/> Emerge Consortium, Emerge Project Fact Sheet, 2007; http://cordis.europa.eu/search/index.cfm?fuseaction=proj.simple.documentlucene&HD_ID=9074973&CFID=359732&CFTOKEN=49382868.
3. Branko, C., Lovell, N., and Chan, D., The potential impact of home telecare of clinical practice, *The Medical Journal of Australia*, 518–521, 1999.
4. Medical Online Encyclopedia, <http://www.nlm.nih.gov/medlineplus/encyclopedia.html>
5. www.mayoclinic.com/health/dehydration/DS00561/DSECTION=2
6. Wendelken, S. et al., The Feasibility of Using a Forehead Reflectance Pulse Oximeter for Automated Remote Triage, New England Engineering in Medicine and Biology Conference, Springfield, MA, 2004.
7. Schrepfer, T., Caduff, A., Hirt, E., and Süssstrunk, H., Method and device for determining the concentration of a substance in body liquid, Patent WO02069791, March 6, 2001.
8. Scheffer, M. and Hirt, E., Wearable devices for telemedicine applications, *Journal of Telemedicine and Telecare*, 11(Supplement 1), 11–14(4), July 2005.
9. Asada, H.H. and Hutchinson, R.C., Sensor design for improved motion artefact reduction without circulatory interference, *MIT Home Automation and Healthcare Consortium Progress Report 3-3*, October 2001–March 2002.
10. Wyss, S. and Rufer, M., Wearable Lifesaver 2, Diploma thesis, IFE, ETH Zurich, 2003.
11. Noury, N., Barralon, P., Virone, G., Boissy, P., Hamel, M., and Rumeau, P., A smart sensor based on rules and its evaluation in daily routines, *Proceedings of the IEEE EMBS 2003*.
12. Guricci, S., Hartriyanti, Y., Hautvast, J., and Deurenberg, P., Prediction of extracellular water and total body water by multifrequency bio-electrical impedance in a Southeast Asian population, *Asia Pacific Journal of Clinical Nutrition*, 8(2): 155–159, 1999.
13. Martinsen, O.G., Grimnes, S., and Haug, E., Measuring depth depends on frequency in electrical skin impedance measurements, *Skin Research and Technology*, 5: 179–181, 1999.
14. Lorincz, K. and Welsh, M., MotTrack: A robust, decentralized approach to RF-based location tracking, *Personal and Ubiquitous Computing*, 11(6): 489–503, 2006.
15. Tüchler, M., Schwarz, V., and Huber, A., Location Accuracy of an UWB Localization System in a Multi-Path Environment, IEEE International Conference on Ultra-Wideband, 2005 Sept. 5–8, 2005, pp. 414–419.
16. Hazas, M. and Ward, A., A novel broadband ultrasonic location system, *Proceedings of UbiComp*, Göteborg, Sweden, September 2002, 264–280.
17. Rusch, T. et al., Signal processing methods for pulse oximetry, *Computers in Biology and Medicine*, 26(2): 143–159, 1996.

18. Pacelli, M. et al., Sensing fabrics for monitoring physiological and biomechanical variables: E-textile solutions, *Proceedings of the Third IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors*, 2006.
19. Elter, P. et al., Noninvasive and nonocclusive determination of blood pressure using laser Doppler flowmetry, *Proceedings of SPIE Specialty Fiber Optics for Medical Applications*, 3596: 188–196.
20. Tunstall Website, www.tunstall.co.uk
21. WristClinic™, <http://www.telcomed.ie/allinone.html>
22. Philips Research, Novel solutions to improve detection of sleep disturbance and early indicators of heart failure, *Password Magazine*, 29, February 2007.

Chapter 2

Integrated Alarm Monitoring System in the ICU

A. Murakami, M. Akutagawa, Y. Ohnishi,
Y. Kuroda, and Y. Kinouchi

CONTENTS

2.1 Introduction	30
2.2 Transparent Error Reporting in the ICU	32
2.3 Specification of Integrated Alarm Monitoring System	34
2.3.1 Data Collection Interface	34
2.3.2 Data Format and Bidirectional Communication.....	36
2.3.3 Structure of System for Indicating Information on Medical Devices	37
2.4 Communication Result with Medical Devices.....	39
2.4.1 Data Collection of Medical Devices.....	39
2.5 Integrated Alarm Monitoring System for Medical Device.....	39
2.5.1 Appearance of System Screen	39
2.5.2 Indication of Alarm Information in System	40
2.5.3 Indication of Alarm Information.....	44

2.5.4 Indication of History of Communication and Alarm
 Information on Medical Devices 44
 2.6 Conclusions and Outlook 46
 References 47

The prevention of human error relevant to medical devices is an important and serious issue in hospitals. In order to guarantee patients’ safety, it is necessary to analyze the cause of errors in detail using error reporting with the electronic Error Reporting System (ERS). In the intensive care unit (ICU) and the high care unit (HCU), a network-based alarm monitoring system that can monitor the state of patients and medical devices from a distance is convenient and essential. This chapter describes the integrated alarm monitoring system for a life-support medical device with alarm functions in the ICU. An advantage of the integrated alarm monitoring system is to provide a fail-safe structure to protect patient life. A change in patient status may not be notified when power source trouble occurs in a system that monitors only a single device. Integrated monitoring enables notification for abnormalities in complementary form. The system communicates with the medical device using a dedicated interface, i.e., a Data Collection Interface (DCI), which converts the communication protocol from RS-232C to TCP/IP. The development of integrated monitoring may be useful to realize a new concept of error reporting that integrates the monitoring system and the ERS.

2.1 Introduction

The prevention of errors and accidents caused by staff in hospitals is an important and serious issue. In order to guarantee patients’ safety, it is necessary to analyze the cause of errors (see Table 2.1) in detail using error reporting with the electronic Error Reporting System (ERS).¹⁻⁶ Moreover, a network-based monitoring system that can monitor the status of patients and medical devices is required.⁷⁻¹⁰

Table 2.1 Errors Relevant to Medical Care

Medical error	An incorrect action or plan that may or may not cause harm to a patient.
Adverse event	An injury to a patient because of medical management, not necessarily because of error.
Near miss	An error that does not reach a patient.
Latent error	Defects in the hospital environment and operations that can lead to medical errors and adverse events.

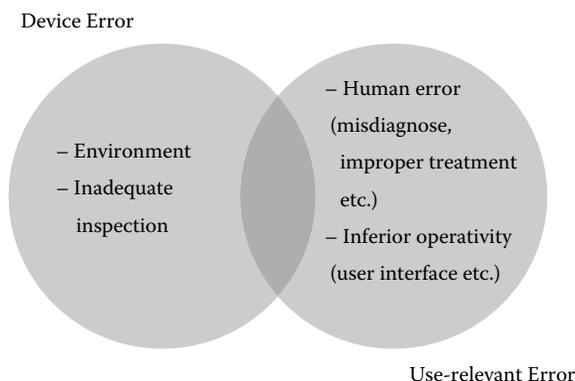


Figure 2.1 Classification of errors relevant to the medical device.

In the intensive care unit (ICU) and the high care unit (HCU), various kinds of medical devices support patients' safety. However, medical devices require high-level expert knowledge and attention in the delivery of care. The concern about accidents relevant to the medical device has increased yearly.^{11,12}

Errors relevant to the medical device are mainly classified into device errors and use-relevant errors as shown in Figure 2.1.

The use-relevant errors are mainly caused by human error during device operation. In order to reduce and prevent these errors, it is effective to analyze their causes using error reporting.

Human errors have been reported by various medical institutions. The style of reporting is divided into mandatory and voluntary.¹ Mandatory reporting may involve accountability in cases of serious harm or death; on the other hand, minor errors that have less urgency are difficult to report. In contrast, minor errors or latent errors that may lead to fatal accidents are mainly reported in a voluntary style. The Institute of Medicine (IOM) strongly endorsed voluntary error reporting in medical care to prevent minor errors.¹

If the reporting procedures are simplified and automated, medical institutions that report error cases voluntarily will increase. For the automation of reporting, it is necessary to collect information of medical care in hospitals. The medical device installed in the hospital is adequately computerized to collect information of medical care.

In the ICU, most of the medical devices have alarm functions for notifying staff to the change of patient state and malfunction of the device. However, it is difficult for staff positioned far from the bedside (e.g., the nurses' station) to notice the occurrence of alarms. In addition, with limited number of staff in the ICU, constant attention must be paid to abnormalities which alarms show. In practice, only a few staff members respond to alarm functions effectively, because there are unsolved problems about the existence of false alarms and issues about urgency ranking and the like.¹³⁻²⁰

In order to improve circumstances for staff in the ICU, installation of an alarm monitoring system using the network is effective. With network communication, the system installed in the nurses' station can monitor various medical devices in integrated form. An advantage of utilizing integrated monitoring is that staff can estimate whether the alarms are valid or not, using various device information.

Another advantage of integrated alarm monitoring is to provide a fail-safe structure to protect patient life. When the system only monitors a single medical device where power source trouble occurs, the change of patient state may not be notified. Integrated monitoring enables notification of the abnormality in complementary form. Especially information on vital monitors such as a pulse oximeter and a capnometer are preferable and important in order to notify the system of patient abnormality. Information on medical devices such as ventilators, infusion pumps, artificial dialyzers, and intra-aortic balloon pumping (IABP) devices, which are installed in the ICU, is also important because it relates directly to life support.

This chapter describes the integrated alarm monitoring system for these life-support medical devices with alarm functions in the ICU. The system collects information on medical devices using a dedicated interface, i.e., the Data Collection Interface (DCI), which converts the communication protocol from RS-232C to TCP/IP.

Moreover, we propose a new style of error reporting that is relevant to the integrated alarm monitoring. The details of this new style of error reporting is described in the following section.

2.2 Transparent Error Reporting in the ICU

Various errors have been reported by error reporting in hospitals, but there is no guarantee that all information is correct. Reports include documenting errors because compilation of erroneous documents is also included in errors.⁵ The documenting errors include under-reporting, underestimation of reporting, misunderstanding, and staff's lapse of memory.

Collecting accurate information about the occurrences of error is dependent not only on staff's memory but also on utilizing information on medical devices installed in hospitals. The ICU and the HCU are replete with computerized medical devices dedicated to collecting information. By collecting data from the medical device, the state of the device and the occurrence of alarms are recorded to an external device. Information that a medical device outputs is important to analyze device operation done by staff.

To prevent documenting errors relevant to error reporting by staff, report contents should be compiled automatically using medical device information. Demands for automatic error reporting may be increased with advancement of hospital facility and device, and it is necessary to improve the style of reporting.

Automation of the error reporting is related to the introduction of transparency. The transparency in this chapter means the concept of user friendliness in human-computer interaction, which relieves the user of the need to worry about technical details (like installation, updating, downloading, or device drivers). For instance, a program that automatically detects the monitor resolution is more transparent compared to one that asks the user to enter it manually. A concept of transparency is related to the realization of a ubiquitous network,²¹ which will be introduced into the field of medical care at the time when network research advances.

The outline of Transparent Error Reporting (TER) utilizing the integrated alarm monitoring system in the ICU is shown in Figure 2.2.

The integrated alarm monitoring system reports the device error and the use-relevant error using information on device operations done by staff. The device operations by staff are recorded with data collection from the medical device. The system reports errors automatically, excluding the human operation of document compilation. Regardless of the importance of the error, all information should be reported.

Using transparent error reporting, information collected by voluntary and mandatory error reporting systems is included in the reported data or documents, because the monitoring system reports all device information constantly regardless of the characteristic and the importance of the error. As for an advantage of utilizing transparent error reporting, it is possible to detect the error type for the medical device. For instance, errors are classified as the device error and the use-relevant error done by staff by examining the report at the time when error occurs.

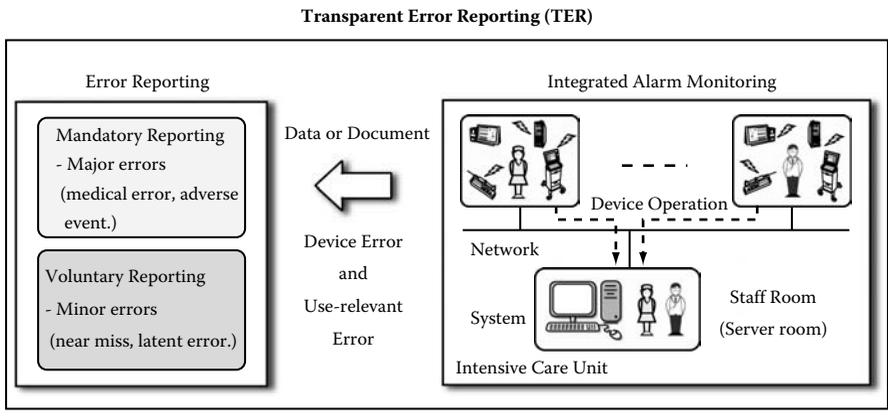


Figure 2.2 Outline of Transparent Error Reporting (TER) in the ICU. The integrated alarm monitoring system reports the device error and the use-relevant error using information on device operations done by staff. The system reports errors automatically, excluding the human operation of document compilation. Regardless of the importance of the error, it should be reported.

2.3 Specification of Integrated Alarm Monitoring System

This section describes the specification of the Web-based integrated alarm monitoring system developed by the authors. The system collects information on medical devices used for care of patients in the ICU. The system communicates with medical devices using a dedicated interface of the DCI on the network. The specification of the interface is also described.

2.3.1 Data Collection Interface

Most of the medical devices installed in the ICU have an output interface like RS-232C. On the other hand, utilizing the network with TCP/IP in place of RS-232C is desired, because installation of network components for medical devices in hospitals is advancing to integrate the information on the devices from a distance. In consideration of cost, performance, and convenience due to operation with existing medical devices, it is difficult to replace a device that can connect to the network.

This study developed the Data Collection Interface (DCI), which converts the communication protocol from RS-232C to TCP/IP. It becomes possible to collect

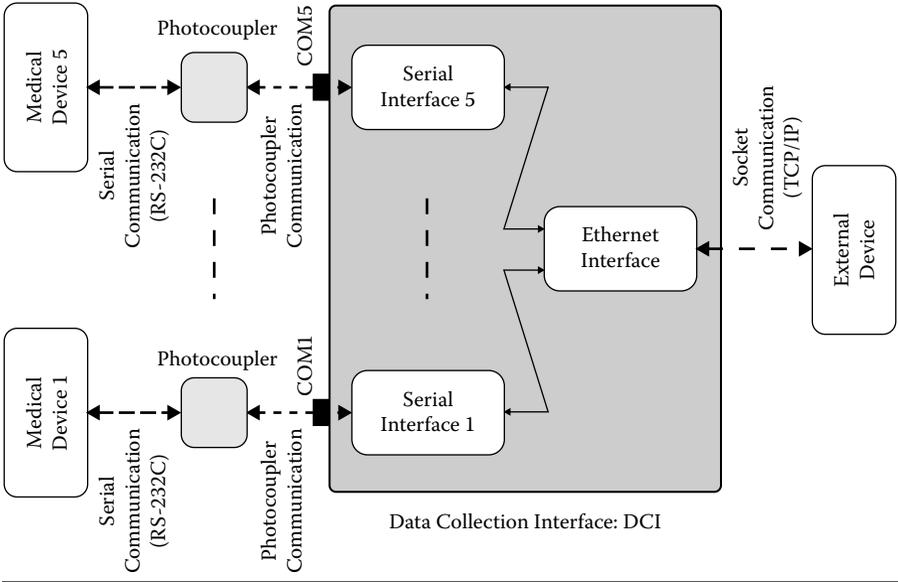


Figure 2.3 A Data Collection Interface (DCI) that communicates with a medical device by RS-232C and transfers data to the network by TCP/IP. The DCI is insulated by photocoupler to prevent an influence of macroshock and the commercial power source.

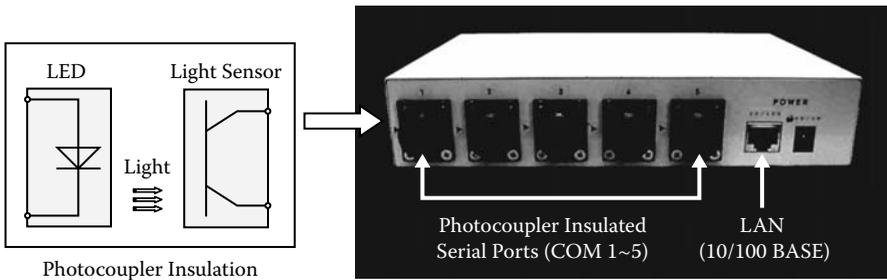


Figure 2.4 Structure of the DCI for communication with a medical device using RS-232C. The photocoupler unit converts an electrical signal to an optical signal. An external device can communicate with a desired medical device connected to the DCI by appointing the serial port from COM1 to COM5.

Table 2.2 Specification of Serial Interface of the DCI^a

Baud rate (bps)	1200/2400/4800/9800/14400/19200/38400
Data length	7 bit/8 bit
Parity	None/even/odd
Stop bits	1 bit/2 bit
Flow control	Xon/Xoff, RTS/CTS

^a The DCI can specify the communication parameters required in serial communication to respective serial ports.

information of existing medical devices installed in the ICU by utilizing the DCI (Figure 2.3), making use of the network.

The DCI is insulated by photocouplers between the serial port and the internal circuit, as shown in Figure 2.3 and Figure 2.4, in order to prevent an influence of the macroshock and the commercial power source. The photocoupler is a device that uses a short optical transmission path to transfer a signal between elements of a circuit. Because the signal goes from an electrical signal to an optical signal back to an electrical signal, electrical contact along the path is isolated.

The serial interface of the DCI communicates with a medical device using serial communication by RS-232C. The Ethernet interface of the DCI communicates with an external device using the socket communication by TCP/IP. Table 2.2 shows the specification of serial interface. The DCI has to specify parameters in serial communication (e.g., baud rate of transmission, number of data bits encoding a character, number of stop bits) to the respective device (see Table 2.3). The DCI can specify communication parameters by request from the external device connected to the network. The DCI can communicate simultaneously with a maximum of five devices using serial ports from COM1 to COM5. The Ethernet interface specifies the IP address required to communicate with the network, and communicates with an external device connected to the network by TCP/IP.

Table 2.3 Specification of Serial Communication of Medical Devices Used in this Study

Type	Infusion Pump	Ventilator	Vital Monitor
Product name	TE-161/TE-172, TE-311/TE-312	Servo 300/300A/i	DS-5400
Manufacturer	TERUMO Co.	SIEMENS Co.	FUKUDA DENSHI Co.
Baud rate (bps)	9600	9600	2400
Parity	None	Even	None
Data length	8 bit	7 bit/8 bit	8 bit
Stop bits	2 bit	1 bit/2 bit	1 bit

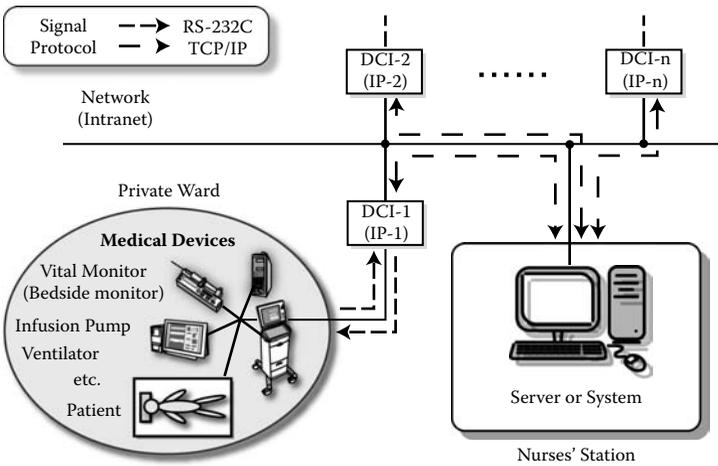


Figure 2.5 It is possible for the server or the system to collect information on several medical devices used for care of patients in the private ward by communication with the DCI using the network.

It is possible for the server or the system installed in the nurses' station to communicate with several medical devices in the private ward by communication with the DCI, as shown in Figure 2.5.

2.3.2 Data Format and Bidirectional Communication

Data format of the DCI includes header information in the head and a cyclic redundancy check (CRC) to verify communication data in the end, as shown in Figure 2.6. The header information includes a desired serial port number and the type of device. The DCI recognizes information in the header, and performs communication with the desired device. We adopt the CRC ITU-T (generation

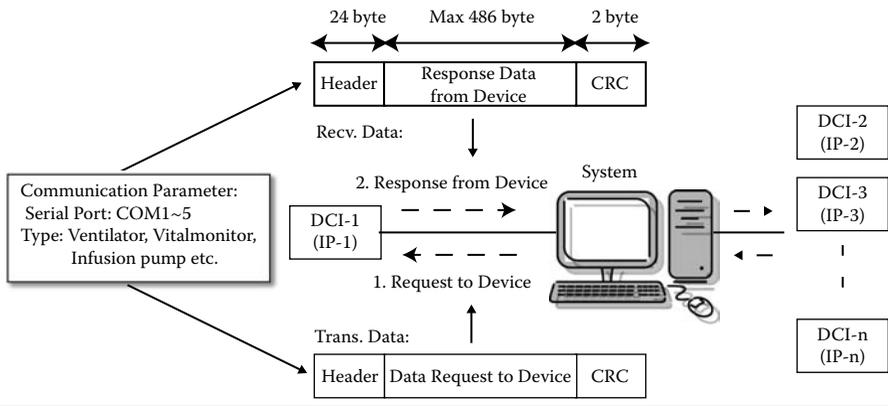


Figure 2.6 Data format of communication using the DCI. The data format includes the header and the footer of the CRC. The system transfers and receives data from medical devices making use of the DCI in bidirectional communication.

polynomial: $X^{16} + X^{12} + X^5 + 1$), and calculation range is taken as whole data length (header + data + CRC).^{22,23} Compared with a usual parity check, error detection capability of the CRC is high, and it can improve the accuracy of communication.

It is possible to monitor contents of alarm notified by medical device accurately, because the system transfers and receives data from medical devices in bidirectional communication, as shown in Figure 2.6. Moreover, the system can notify the occurrence of abnormality to staff instantaneously, because contents of communication are observed continuously. Communication data of the DCI includes transmission data from medical devices and the system. If the system sends an incorrect request or header data, the medical device does not send a response or an error response. The system can detect the medical devices connected to the DCI by analyzing responses from the DCI, and confirms abnormalities of the network connection and the device by presence of response.

There is another method of communication with medical devices that utilizes a wireless device like a medical telemeter. However, the system cannot send information to the medical device at the bedside, because a telemetry system communicates in a unidirectional way, which means there is no method to verify information from a medical telemeter. Moreover, it is not possible to select output of a medical device. Therefore, bidirectional communication is preferable in data collection of medical devices which include an alarm device.

2.3.3 Structure of System for Indicating Information on Medical Devices

The system generates a Web document that includes information about the medical device and the patient, and publishes the information on the Web in the local area.

A block diagram indicating structure of the Web document in the system is shown in Figure 2.7.

[MD monitor] communicates with the DCI and collects data from medical devices and records information to [data storage]. [TCP/IP monitor] communicates with the DCI by using the Internet Control Message Protocol (ICMP), and monitors the state of network connection to the DCI.

[Alarm checker] detects the abnormality of the medical device and communication using information of [TCP/IP monitor] and [MD monitor], and records information to [alarm database]. [Alarm checker] also detects the abnormality of communication on serial communication and network, and power source trouble of the DCI and medical device, then notifies the staff by using the Simple Mail Transfer Protocol (SMTP).

When there is no response from the DCI, the system detects that the device or the network has a problem. These problems are classified as “problem in communication.”

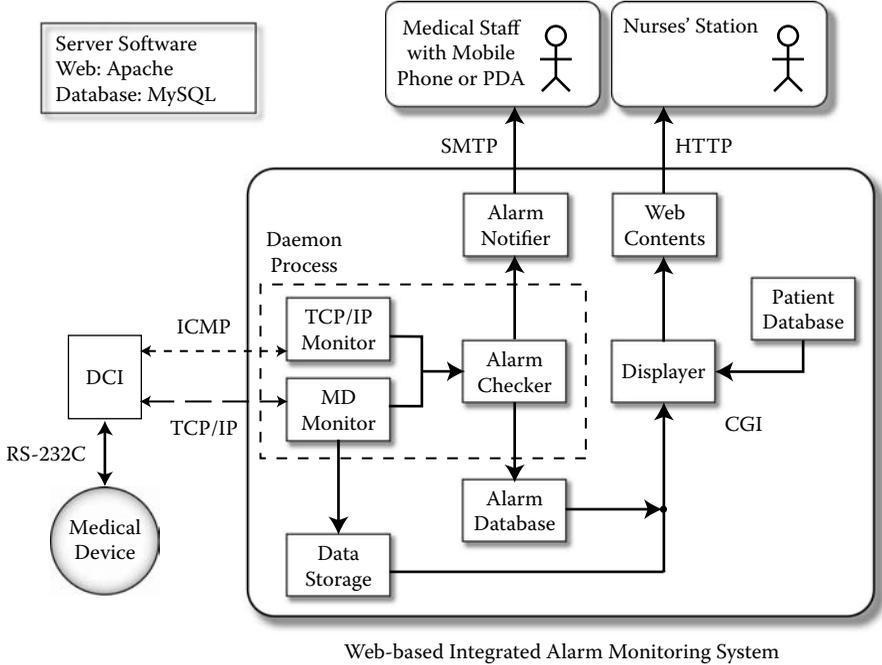


Figure 2.7 The system generates a Web document and publishes it on the Web in the local area. Figure shows a block diagram indicating structure of the Web document in the system. TCP/IP monitor and MD monitor run in the background as a daemon process. Displayer runs on http server process as a CGI.

When the medical device causes an alarm, the system can detect the occurrence of a problem related to the contents of the alarm. These problems are classified as “problem in device” or “problem in patient.”

The [displayer] generates a Web document when the system receives a request from the Web browser. Information on [alarm database], [data storage], and [Patient database] are included in the Web document displayed in the Web browser. The system displays information following the priority of care such as generation of alarm and the change of patient state.

2.4 Communication Result with Medical Devices

The system collects information on medical devices with network communication using the DCI. The system can specify the interval of communication with a medical device manually, because response from the device is obtained by demand from the system. In this study, communication interval by the system is specified from several seconds to several dozen seconds in order to collect alarm information and the device operation done by staff.

2.4.1 Data Collection of Medical Devices

Examples of data collection using the DCI are shown in Figure 2.8, Table 2.4, and Table 2.5.

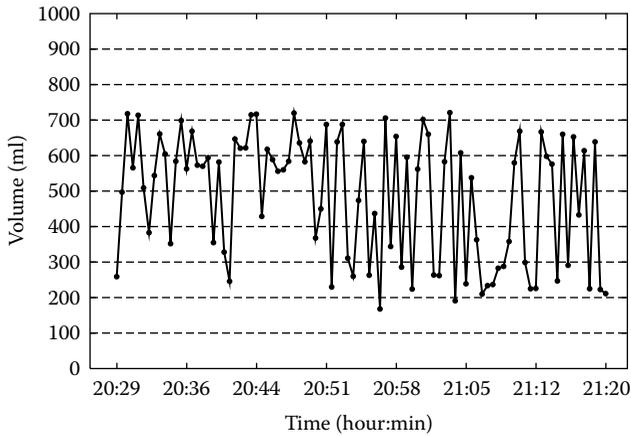
Figure 2.8 shows the values which ventilator (Servo 300/300A) outputs. The system has recorded values from the ventilator every 30 sec for 1 h.

Table 2.4 and Table 2.5 show the values specified to the infusion pump (TE-172) and the ventilator (Servo 300/300A) by staff. These setting values are useful to analyze the device operation done by staff, when an accident occurs.

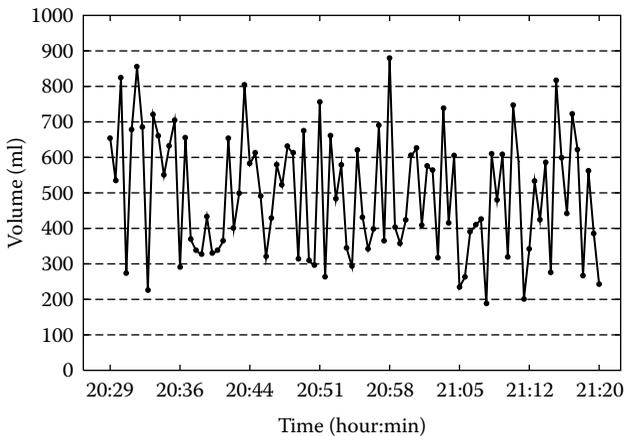
2.5 Integrated Alarm Monitoring System for Medical Device

2.5.1 Appearance of System Screen

Figure 2.9 shows a system screen. Staff can verify the state of patient and medical device, because a Web document that the system generates is easily browsed with any Internet-connected PC in the hospital, using the Web browser. In the system screen, a number of private wards and information of patients in the ICU are displayed in one screen. The number and the types of medical devices connected to a patient are also displayed in the patient screen. The Web document is automatically updated with the change of patient state or medical device.



(a) Inspiratory Tidal Volume



(b) Expiratory Tidal Volume

Figure 2.8 Results of data collection with ventilator (Servo 300/300A). The values are recorded every 30 seconds for 1 hour.

2.5.2 Indication of Alarm Information in System

Figure 2.10 shows an example of the indication of alarm information displayed in the system screen. Patient information like the gender and name of the patient are displayed in the patient screen, and the number and the types of medical devices that are used for the patient are also displayed by respective marks. If staff clicks a desired patient screen, the system displays a menu that shows detailed patient information in pop-up style.

Table 2.4 Examples of Setting Values Recorded by Data Collection with Infusion Pump (TE-172)

<i>Time</i>	<i>Weight</i>	<i>Medication Volume</i>	<i>Solution Volume</i>	<i>Flow Volume</i>	<i>Scheduled Volume</i>	<i>Integrating Volume</i>
14:10:30	90	100	110	8	100	1
14:10:40	90	100	110	8	100	1
14:10:50	90	100	110	8	100	1
14:11:00	90	100	110	8	100	2
14:11:10	90	100	110	8	100	2
14:11:20	90	100	110	11	100	2
14:11:30	90	100	110	11	100	2

Table 2.5 Examples of Setting Values Recorded by Data Collection with Ventilator (Servo 300/300A)

Pause time percent set	10
SIMV frequency set	15
Inspiratory rise time percent set	5
Volume set	377.5
Pressure control level above PEEP set	9.8
Pressure support level above PEEP set	8.8
PEEP set	0
Ventilation mode set	VCVPS
Expiratory minute volume upper alarm limit set	18.4
Expiratory minute volume lower alarm limit set	2.9
Upper pressure limit set	48
O2 concentration upper alarm limit	27
O2 concentration lower alarm limit	18
O2 concentration set	20.9
Trigger sensitivity level below PEEP set	0

The system automatically detects the medical devices used for a patient and displays the information in the patient screen. We classify four types of medical devices used for patients in the ICU: vital monitor, ventilator, syringe pump, and transfusion pump. The system displays patient screens with green when the state of communication, patient, and medical device has no problem. Moreover, the system displays patient screens with colors of orange, yellow, and red, which mean “problem in device,” “problem in communication,” and “problem in patient,” respectively. “Problem in patient” means the state of the vital monitor is in “problem in device,” which means the state of

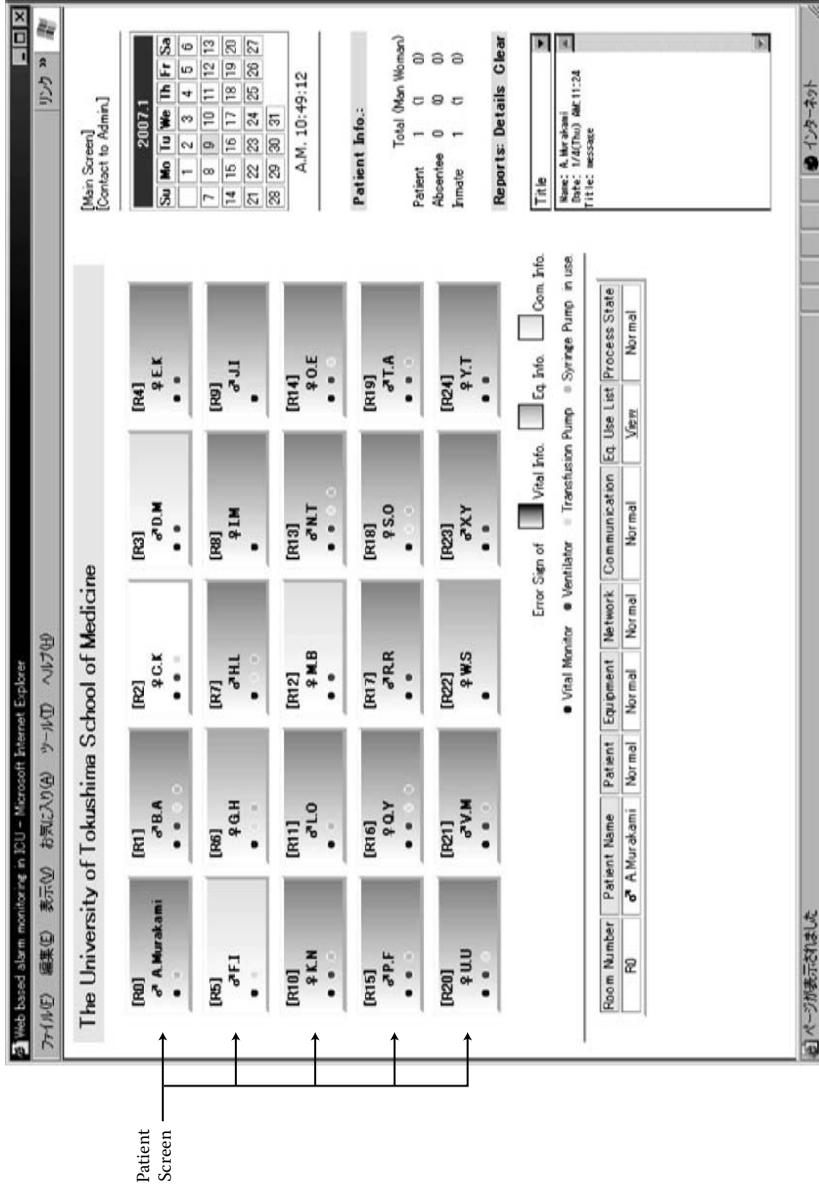


Figure 2.9 Appearance of the system screen that is displayed on the Web browser. The system displays a number of patient screens which include information of the medical device used for the patient.

1. Staff Inputs Patient Information

2. The System Indicates Patient Information and Medical Devices

Room color (Notification of abnormality):
 Red: Patient's condition (Vital monitor)
 Orange: State of device (Other device)
 Yellow: State of communication
 Green: No alarm and no abnormality

Room No. → [R0] Patient Name
 Gender → ♂ A. Murakami
 (Male/Female)

Device Type
 Connected to a Patient

Mark color (Device type):
 Red: Vital monitor
 Blue: Ventilator
 Orange: Syringe pump
 Yellow: Transfusion pump

Pop-up Menu

- Equipment List
- Alarm History
- Confirmation of Connection
- Call off Alarm
- Save Ed. Details

Room Number: R0
 Name: A. Murakami
 Age: 027 [Year]
 Gender: ♂ Blood Type: AB
 Height: 170 [cm] Weight: 60 [kg]
 Infectious Disease: None
 Remarks:

Figure 2.10 Example of the indication of alarm information displayed in the system screen. After the input of patient information, the patient screen in the system screen is updated. Information of the medical devices used for the patient is also displayed in the patient screen. Abnormalities of patient and the medical devices are displayed as color of the patient screen and the mark of the device.

the medical device is notifying alarm. “Problem in communication” means the state of communication with TCP/IP or RS-232C has abnormalities in the communication path between the system, the DCI, and the medical device.

Patient screen blinks when the system detects the abnormality in the ICU, then the indication of medical device that is notifying the alarm blinks.

2.5.3 Indication of Alarm Information

Figure 2.11 shows an example of the indication in the system screen when vital monitor (DS-5400) is notifying alarms.

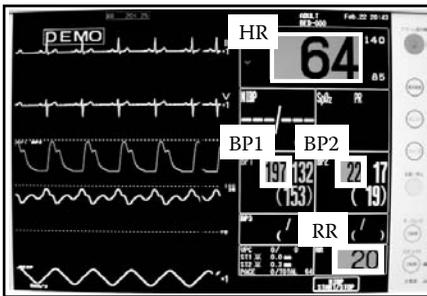
The vital monitor in Figure 2.11 (1) is notifying alarms of heart rate (HR), blood pressure (BP1 and BP2), and RR interval (RR) by showing emphasis indication on its monitor with alarm sounds. When a medical device notifies alarms, the system detects the private ward where the device is located and updates the indication of patient screen.

In the system screen shown in Figure 2.11(b), the indicators of patient screen and the vital monitor are blinking. If staff clicks the patient screen, the system displays a menu that shows detailed information for alarms.

2.5.4 Indication of History of Communication and Alarm Information on Medical Devices

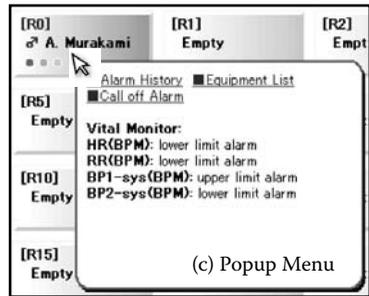
It is necessary to record the state of staff and medical devices in detail, in order to analyze circumstances of medical care and to automate error reporting in the ICU. The system records information on communication and medical device as histories.

(a) Vital Monitor (DS-5400) (Bedside monitor)



1. Generation of Alarms

(b) System Screen



(c) Popup Menu

2. Modification of Indication

Figure 2.11 Example of the indication in the system screen when vital monitor (DS-5400) is notifying alarms. When alarm is notified by vital monitor, the color of patient screen is changed to red and the mark of the vital monitor (red) blinks. Then staff can confirm information by the pop-up menu displayed on the patient screen.

Date	Time	Room	Contents
04/17	17:19	R0	Communication error [Soclet]
04/17	17:16	R0	Communication error [Syringe pump: no response]
04/17	16:41	R0	Communication starts with equipments
04/17	16:41	R0	COM4: Transfusion pump [TE-172]
04/17	16:41	R0	COM6: Syringe pump [TE-1221C]
04/17	16:41	R0	COM7: Ventilator [SV-300]
04/17	16:41	R0	COM1: vital monitor [DS-5400]
04/17	16:41	R0	Confirming equipment connected with DCI
04/17	16:10	R0	Communication ends
04/17	16:15	R0	Communication starts with equipments
04/17	16:15	R0	COM1: Syringe pump [TE-3121C]
04/17	16:15	R0	Confirming equipment connected with DCI
04/17	16:14	R0	Communication ends
04/17	16:14	R0	Communication starts with equipments
04/17	16:14	R0	COM1: Syringe pump [TE-3121C]
04/17	16:14	R0	Confirming equipment connected with DCI
04/17	16:05	R0	Communication ends
04/17	16:05	R0	Communication starts with equipments
04/17	16:05	R0	COM1: Syringe pump [TE-3121C]
04/17	16:05	R0	Confirming equipment connected with DCI

(a) History of Communication

Date	Time	Room	Contents
04/17	17:12	R0	Vital monitor [DS-5400]: asystole alarm
04/17	17:12	R0	Vital monitor [DS-5400]: SpO ₂ lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: APNEA alarm
04/17	17:12	R0	Vital monitor [DS-5400]: RR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: HR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: asystole alarm
04/17	17:12	R0	Vital monitor [DS-5400]: SpO ₂ lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: APNEA alarm
04/17	17:12	R0	Vital monitor [DS-5400]: RR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: HR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: asystole alarm
04/17	17:12	R0	Vital monitor [DS-5400]: SpO ₂ lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: APNEA alarm
04/17	17:12	R0	Vital monitor [DS-5400]: RR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: HR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: asystole alarm
04/17	17:12	R0	Vital monitor [DS-5400]: SpO ₂ lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: APNEA alarm
04/17	17:12	R0	Vital monitor [DS-5400]: RR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: HR(BPM) lower limit alarm
04/17	17:12	R0	Vital monitor [DS-5400]: asystole alarm

(b) History of Alarm Information of Medical Device

Figure 2.12 History of communication and alarm information notified by the medical device.

Figure 2.12(a) shows the history of information on communication. The history shows the times of the beginning and the end of communication between the system and the medical devices. Moreover, the system records errors of communication at the time of disconnection on serial communication and the network.

Figure 2.12(b) shows the history of alarm information notified by the medical device.

2.6 Conclusions and Outlook

In this chapter, a summary of the integrated alarm monitoring system developed by the authors is described. In order to collect information in the ICU, the system has to collect information of medical devices in integrated form, making use of an interface like the DCI.

The DCI converts the output of medical devices from RS-232C to TCP/IP. Staff becomes able to monitor the state of medical devices far from the nurses' station using a single monitor which shows a Web document as shown in Figure 2.9 generated by the system. Moreover, the system records the alarm information that a medical device outputs as a history, as shown in Figure 2.12. Using the system, the history information can be verified.

The history of information on communication is as important as the information of the alarm. There is a possibility of communication breaking off with various primary factors, because various people (staff, patient, and patient's family, etc.) and medical devices exist together in the ICU. The occurrence of abnormality on the medical device must correspond instantaneously because there is direct impact to the patient. Likewise, the abnormality of communication that monitors the state of patient and medical device must be corresponding instantaneously.

Table 2.4 and Table 2.5 are the setting values that staff specifies to respective medical devices. The device operation done by staff is recorded by analyzing the change of these values. The history function to record information regarding the device operation is installed to several medical devices (TE-261, TE-361: TERUMO Co.). However, these items of information are preferable to analyze in detail with the state of the patient by a vital monitor. In order to determine the occasion where an accident occurs, the history function is useful to provide information regarding the device operation as an electronic record. The history function will become even more important with development of error reporting. It is necessary to record the device operation continuously because of the concept of transparent error reporting.

Moreover, it is possible to provide safer medical care by providing the interface which gives attention to the erroneous operation done by staff. Construction of the database and the interface in order to improve the relationship between the medical device, the staff, and the patient becomes necessary.

The necessity to argue the concept of inspection and safety for the medical device becomes high because the opportunity to use the device in the intensive care

area has been increased. As one of our goals, it is necessary to provide the system or the environment to develop the transparent error reporting described in this chapter. Therefore, communication with various medical devices and the cooperation of medical institutions become necessary.

References

1. Institute of Medicine Committee on Quality of Health Care in America, *To Err Is Human: Building a Safer Health System*, Washington, D.C.: National Academies Press, 1999.
2. Institute of Medicine Committee on Quality of Health Care in America, *Crossing the Quality Chasm: A New Health System for the 21st Century*, Washington, D.C.: National Academies Press, 2001.
3. P.Y. Boelle, P. Garnerin, F. Clergue, J.F. Sicard, and F. Bonnet, Voluntary reporting system in anaesthesia: Is there a link between undesirable and critical events?, *Qual. Health Care*, 9, 203–209, 2000.
4. Report of the QuIC to the President, Doing What Counts for Patient Safety: Federal Actions to Reduce Medical Errors and Their Impact, The Quality Interagency Coordination Task Force (QuIC), 2000, <http://www.quic.gov/report/>
5. J.A. Taylor, D. Brownstein, D.A. Christakis, S. Blackburn, T.P. Strandjord, E.J. Klein, and J. Shafii, Use of incident reports by physicians and nurses to document medical errors in pediatric patients, *Pediatrics*, 114(3): 729, 2004.
6. S. Nishigaki, J. Vavrin, N. Kano, T. Haga, J.C. Kunz, and K. Law, Humanware, human error, and hiyari-hat: A template of unsafe symptoms, *ASCE, J. Construction Eng. Manage.*, 120 (2): 421–442, 1994.
7. H. Takeda, Y. Matsumura, S. Kuwata, H. Nakano, N. Sakamoto, and R. Yamamoto, Architecture for networked electronic patient record systems, *Int. J. Med. Inform.*, 60, 161–167, 2000.
8. A.W. Kushniruk, C. Patel, V.L. Patel, and J.J. Cimino, “Televaluation” of clinical information systems: An integrative approach to assessing Web-based systems, *Int. J. Med. Inform.*, 61, 45–70, 2001.
9. D.I. Shin, S.J. Huh, T.S. Lee, and I.Y. Kim, Web-based remote monitoring of infant incubators in the ICU, *Int. J. Med. Inform.*, 71, 151–156, 2003.
10. M.C. Chambrin, P. Ravoux, D.C. Aros, A. Jaborska, C. Chopin, and B. Boniface, Multicentric study of monitoring alarms in the adult intensive care unit (ICU): A descriptive analysis, *Intensive Care Med.*, 25, 1360–1366, 1994.
11. Guidance for Industry and FDA Premarket and Design Control Reviewers, *Medical Device Use-Safety: Incorporating Human Factors Engineering into Risk Management*, FDA, 2000, <http://www.fda.gov/cdrh/humfac/1497.html>
12. N.J. Bahr, *System Safety Engineering and Risk Assessment: A Practical Approach*, Washington, D.C.: Taylor and Francis, 1997.
13. S. Lawless, Crying wolf: False alarms in a pediatric intensive care unit, *Crit. Care Med.*, 22(6): 981–985, 1984.
14. J. Beneken and J. Van der Aa, Alarms and their limits in monitoring, *J. Clin. Monitoring*, 5 (3): 205–210, 1989.

15. G.A. Finley and A.J. Cohen, Perceived urgency and the anaesthetist: responses to common operating room monitor alarms, *Can. J. Anesth.*, 38: 958–964, 1991.
16. M.C. Chambrin, Alarms in the intensive care unit: How can the number of false alarms be reduced?, *Crit. Care*, 5(4): 2001.
17. T.A. Mondor and G.A. Finley, The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate, *Can. J. Anesth.*, 50(3): 221–228, 2003.
18. J. Edworthy and R. Hards, Learning auditory warnings: The effects of sound type, verbal labelling and imagery on the identification of alarm sounds, *Int. J. Ind. Ergo.*, 24, 603–618, 2004.
19. CAN/CSA-ISO9703.1-97, Anaesthesia and Respiratory Care Alarm Signals—Part 1: Visual Alarm Signals, Canadian Standards Assoc., 1997.
20. CAN/CSA-ISO 9703.2-97, Anaesthesia and Respiratory Care Alarm Signals—Part 2: Auditory Alarm Signals, Canadian Standards Assoc., 1997.
21. H. Morikawa and T. Aoyama, Realizing the ubiquitous network: The Internet and beyond, *J.Telecommun. Syst.*, 25 (3-4): 449–468, 2004.
22. B. Sklar, *Digital Communications: Fundamentals and Applications*, Prentice Hall, 1988.
23. S.B. Wicker, *Error Control Systems for Digital Communication and Storage*, Prentice Hall, 1995.

Chapter 3

Remote Wireless Patients' Data Access System

Ziad Hunaiti, Ammar Rahman, Gregory Savelis, Zayed Huneiti, and Wamadeva Balachandran

CONTENTS

- 3.1 Introduction.....50
- 3.2 Vital Signs Monitoring50
- 3.3 System Architecture..... 51
 - 3.3.1 Database Server.....52
 - 3.3.2 Handheld Device and Interface.....53
- 3.4 System Evaluation.....56
- 3.5 Conclusion.....59
- References60

This chapter presents a mobile medical data access system, specifically designed for the vital patients' information remote monitoring. The chapter describes the design of the complete system that enables medical staff global access for monitoring their patients and providing relevant feedback. The system is designed to

operate over different types of networks depending on the availability of coverage. The chapter then goes into a detailed presentation of the prototype designed and tested by Brunel University, which provides in-hospital wireless access through a WLAN. A pre-evaluation is conducted with doctors and nurses to assess system usability and functionality.

3.1 Introduction

In the last two decades technology in general and information and communications technologies (ICT) in particular have been advancing at a fast pace. Medical services are one of the major beneficiaries of this development. Most of the medical systems have moved from traditional paper form to digital forms. Nowadays, the majority of medical records are stored in digital format. In turn, digital transmission and exchange of medical information has become possible. This was facilitated by rapid advancements in digital telecommunications networks. Thanks to the Internet, authorized medical personnel can have access to patients' information globally through any device connected to the Internet.¹

3.2 Vital Signs Monitoring

Vital signs refer to a set of readings that can express the basic health status of patients. These include temperature, pulse, blood pressure, respiratory rate, and blood gases. For some patients this list can be expanded to include other parameters such as fluid intake and output. Vital signs monitoring is one of the main routines that all inpatients have to go through. Even though most medical equipment has substantially developed into digital form with many of them including computer interfaces, vital signs monitoring has not taken full advantage of this development.² The protocol currently in use requires the nurse/care assistant to record a patient's clinical data at the point of care several times a day. The clinical data is then plotted on a sheet of paper which is clamped to the patient's bed for doctors to review during their daily ward rounds. This protocol does possess at least a handful of disadvantages, including but not limited to the following:

- Inefficient utilization of doctors' time due to the fact that they have to physically pass by every patient's bed to look at their records on a daily basis even if such a visit is not warranted (e.g., patient's status is clearly stable).
- The need for the doctor to be available on-site for the monitoring of his patients. This in turn implies the need for extra staff for contingency plans in case of doctor's absence.
- Inefficient utilization of nurses' time because of the need to record the digital clinical data in paper form.

- Possible inaccuracies in graphical representations of patients' clinical data graphs due to the manual presentation in paper form.
- The need for extra staff time for transferring record data from paper form to electronic form for storing in computer systems.
- Confidentiality of patients' clinical data cannot be secure and protected because it is attached to their beds. This can be seen by any person passing by, including visitors and staff who are not directly involved in patients' treatment.

This chapter investigates and proposes an extension to the availability of medical data records for clinical and health care applications. A solution for providing support for efficient remote patients' monitoring is researched. It makes use of the new technological developments in computing and wireless communications systems to provide mobile access to patients' data for authorized medical professionals. This system can result in a protocol for monitoring inpatients that is time efficient, more accurate, more secure, and could directly reduce the total cost of treatment.

3.3 System Architecture

The system consists of three major parts: a mobile unit, a medical information server, and a suitable communications link. The communications link consists of two parts: one for local access and the second for wide area communication. The local access network is for communication within the hospital and the wide area network is to enable doctors to access patients' clinical data remotely.^{3,4} Figure 3.1

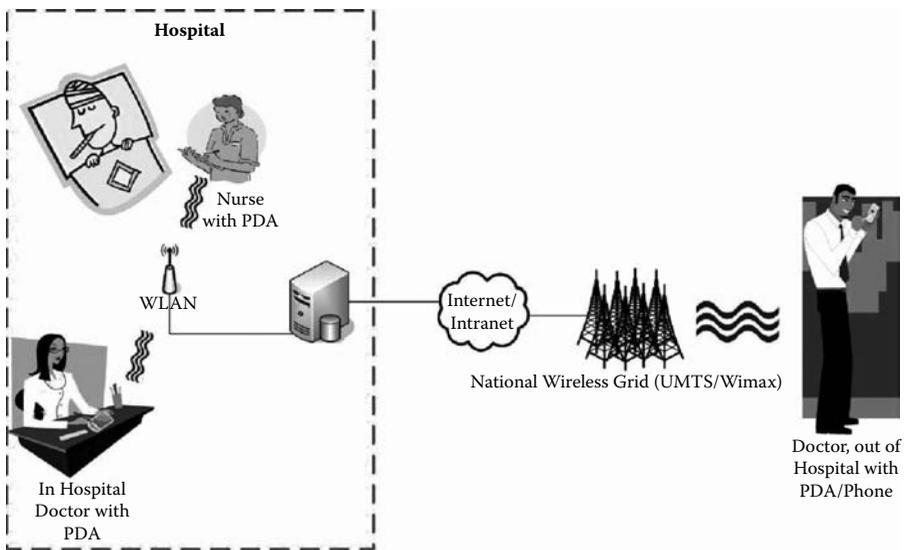


Figure 3.1 General structure of the wireless inpatients monitoring system.

shows a functional diagram of the wireless inpatients monitoring system. The system connects to the hospital's database using a WLAN within the hospital and using a Universal Mobile Telecommunications System (UMTS) from outside the hospital. However, the system is platform independent. This enables the migration to newer platform such as WiMAX.

3.3.1 Database Server

In accordance with the recommendation of the Connecting For Health (CFH) initiative,⁵ vital information monitoring is considered part of a “detailed care record,” therefore the information collected should be stored in a local database within the hospital itself. For the design of the database, a Web client/server approach was used⁶ (Figure 3.2). This approach provides a Web interface that is accessible from any Web-enabled device. This provides several advantages³:

- Providing a unified interface for in-hospital and off-hospital access from any device equipped with a Web browser.
- Simplifying the process of software installation and maintenance. For example, once an administrator installs new a software tool or upgrades an existing application, where the installation is done on the server side, all the clients can use the upgraded software version, whereas in the stand-alone application the installation has to be done on every PC.

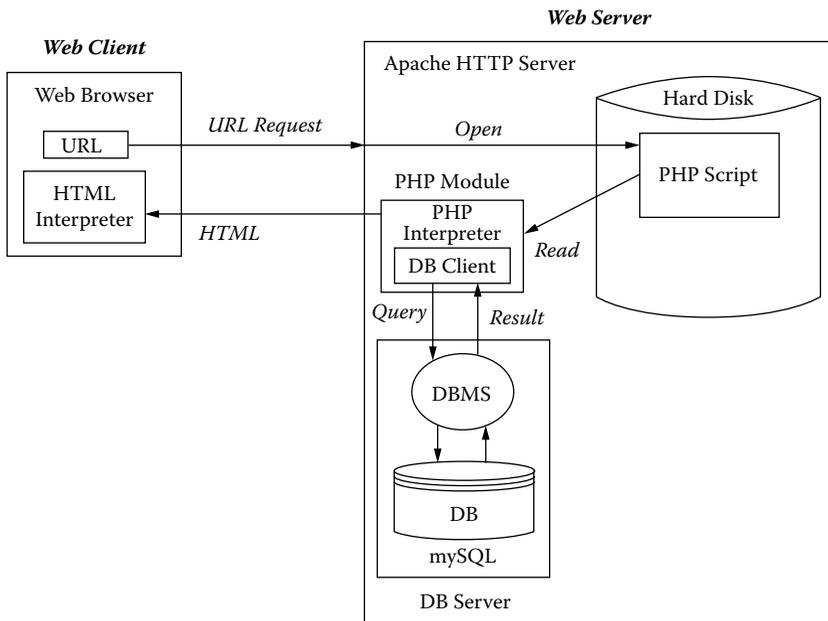


Figure 3.2 Database server's software interaction.

- Upgrading the hardware also becomes a far simpler operation. Because all processing work is being done on the server side, major upgrades are only needed on the server side.
- Enhancing the security. Due to the fact that no data is stored or processed within the mobile device, a security breach would not occur in case of the loss or theft of those devices.

Designing an interface for database access had to be done in a dynamic language to allow different access rights to be implemented. PHP (Hypertext Preprocessor) was the language of choice to accompany the basic HTML in designing this interface. The choice of language was highly influenced by the fact that it is an open source language with no royalties involved for using it. Other reasons included the immense amount of support that is available on the Internet for developers.

A sample database needed to be generated to reflect the United Kingdom NHS database server. For this prototype, MySQL was chosen for the implementation. The choice of this database can be justified for use on this system because of its free availability and compatibility with PHP. MySQL is a thoroughly tested application known to be a very reliable database server.

A simple database was designed for the implementation of this prototype. Figure 3.3 shows the database which was made into five different tables that can easily be incorporated into a real-life system. In real-life implementation, these tables are designed to be parts of different servers. Doctor and Nurse tables can be implemented in an authentication database to provide access rights to the relevant information. The Patient table reflects the recommendation by CFH for a “summary care record” which contains the patient’s basic health information. This is to be stored in a centralized NHS database that is accessible anywhere in the United Kingdom, although the PatientData table and the RPatientData are to be stored in the local hospital database. The PatientData table contains the collected vital readings and the RPatientData table contains the diagnoses and recommendations provided by the monitoring doctor. The reason for these two tables being separated is because they could fall into different ethical categories when sharing the information with other involved medical professionals (e.g., GP or nurses).

3.3.2 Handheld Device and Interface

The prototype designed for testing the system was based on the use of off-the-shelf devices. A suitable device was chosen to meet the following criteria:

- Portability
 - i. Size and weight should be suitable for the medical staff to carry around in their pocket

Database
record_system
Relationships

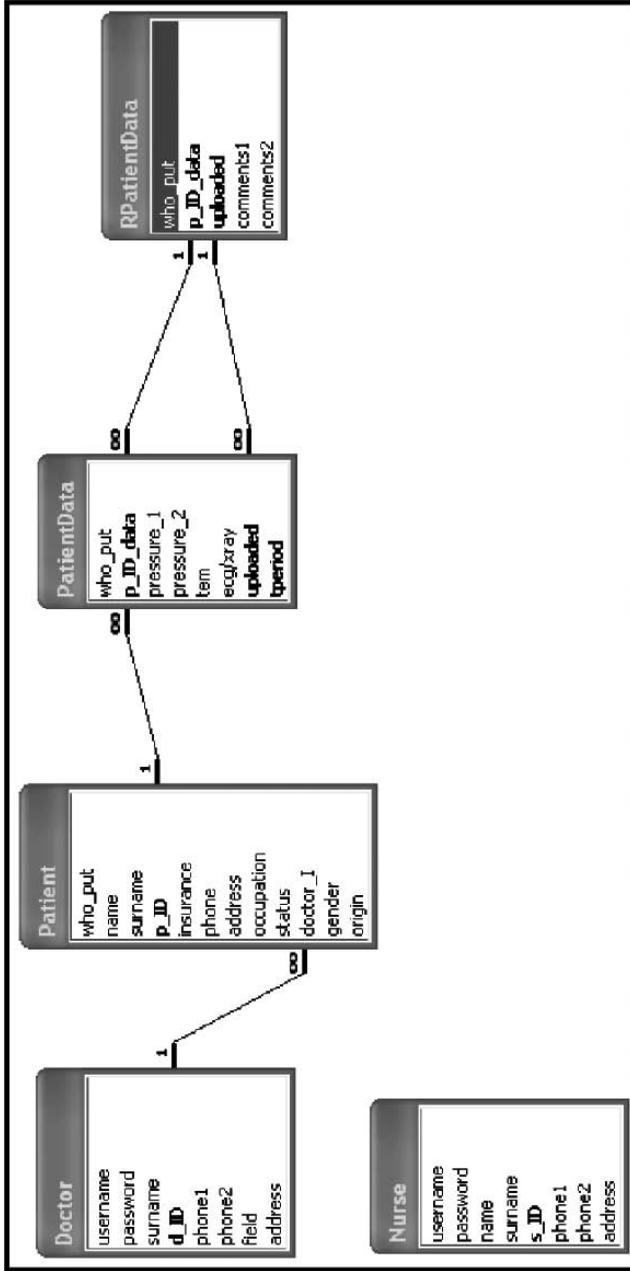


Figure 3.3 Relational database structure.

- ii. Battery life is long enough for the staff to use during their ward rounds without the need to recharge often
- The screen size and resolution should be sufficient to display and edit patients' data
- Provides a suitable wireless interface for the transmission of data; in this case the choice was WLAN 802.11b or 802.11g
- Provides a Web interface with support for PHP for accessing the database

Based on the aforementioned criteria, PDAs were chosen as the most suitable devices for building and testing the prototype. An HP iPAQ h4150 was chosen for the prototype implementation. However, for building the final device, a few customizations may be necessary to ensure flawless medical service. These include implementation of a card and pin authentication standard as recommended by the CFH initiative and the addition of RFID sensors to the devices for the flawless recognition of patients' beds. For the benefit of the prototype, a standard user name and password combination was used (Figure 3.4). For patient identification and access, the system was limited for accessing the patients by name only. Figure 3.5 shows the patient selection screen and display.

As Figure 3.5 shows, once medical personnel are authenticated, they are presented with a drop-down menu of the patients they are authorized to view. This conforms to the guidelines of CFH for the security of patients' records. To make the selection process simpler, medical personnel are required to select



Figure 3.4 Authentication page.

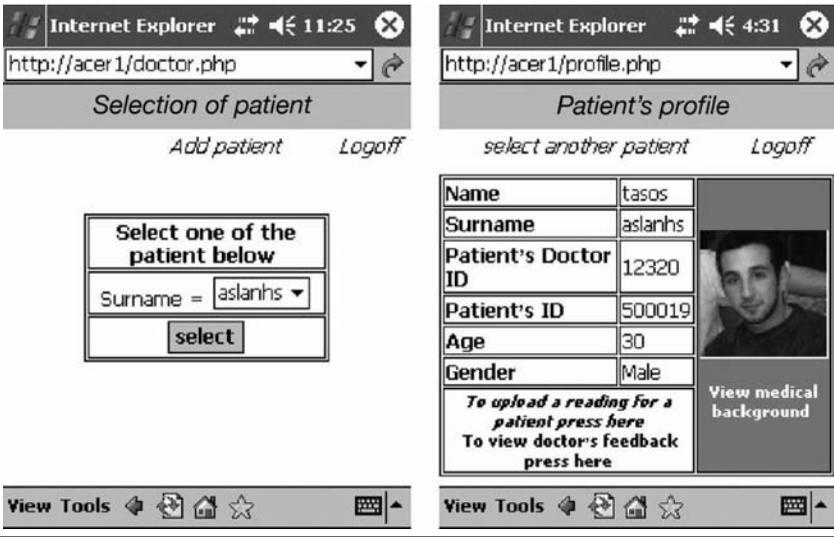


Figure 3.5 Patient selection and view.

the surname, from which they are presented with a following screen to select the specific patient from. These screens are common to all medical personnel involved in the treatment of the patient. However, the following screens distinguish between nurses and doctors. Nurses are presented with an interface where they can record the readings they take, and doctors are able to view the results and provide feedback to the nurses. Several pages were designed to reflect the needs and recommendations of medical personnel. Figure 3.6 shows samples of the data entry pages used by the nurses and the result review pages for the doctors where color coding is one of the features that doctors recommended for ease of reading.

Other pages designed include access to patients' examination results such as electrocardiogram (ECG), X-ray images, and laboratory tests.

Administration of the database and setting of access rights were assigned to administrative staff. These are faced with administrative interface as they log into the system. In larger system implementations, this interface can be substituted by a larger management system.

3.4 System Evaluation

An evaluation of the prototype was carried out with doctors and nurses from two medical centers in the United Kingdom and Greece. A set of questionnaires was compiled to measure the usability of the system. The questionnaires were composed of ten different questions using a five-point Likert scale. It was designed to measure

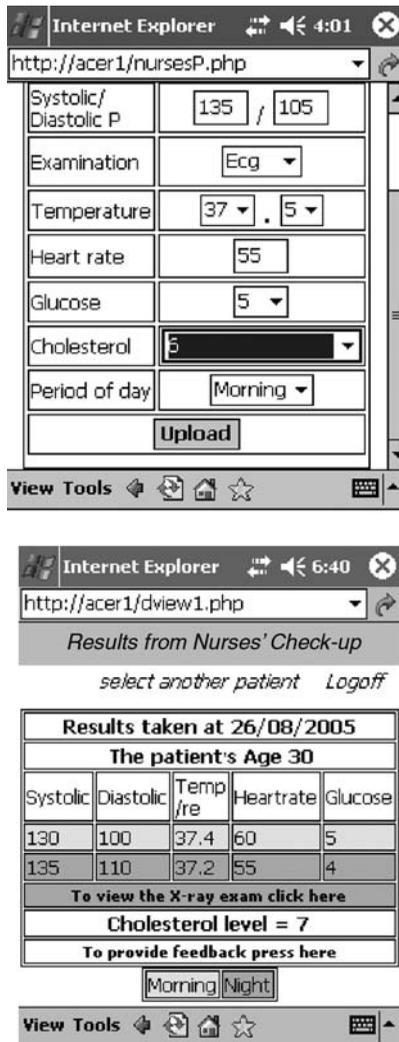


Figure 3.6 Sample pages of the interface.

the general satisfaction with carrying and using the system. The questionnaires were handed out, along with basic instructions for the use of the system. The participants were asked to answer the questions after trying the system with the author. Table 3.1 shows the questions included in the questionnaire.

The questionnaire was handed to two different groups: doctors and nurses. Results from participants of each group were then analyzed by taking the mean of the answers from participants within the group. With the used scale, higher numbers represent more satisfaction, with the exception of question 8. For the consistency of

Table 3.1 Evaluation Questions

Number	Question
Q1	The wireless device is easy to carry around with me
Q2	The PDA is easy to use
Q3	The PDA is suitable for this application
Q4	The application is easy to learn
Q5	This application can be useful for healthcare staff
Q6	I would have preferred instructions that were easier to understand
Q7	Sufficient information for diagnostic
Q8	I found it difficult to use the application
Q9	Screen directions are consistent and easy to follow
Q10	Text and images on each screen are clear enough

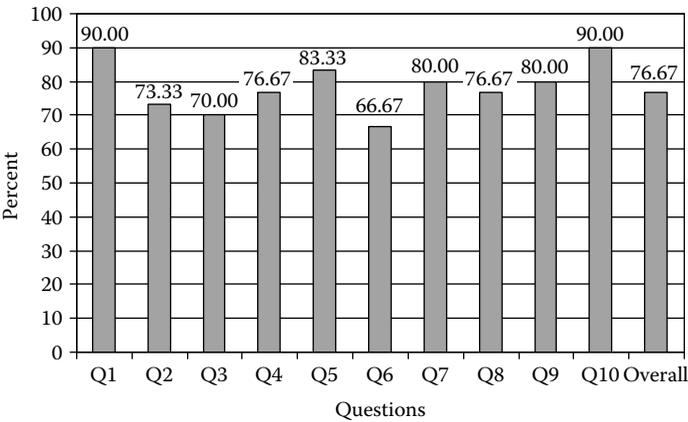


Figure 3.7 Doctors' satisfaction.

presentation, the results of question 8 were inverted to give 5 for totally disagree and 1 for totally agree. Figure 3.7 shows the satisfaction of doctors and Figure 3.8 shows it for nurses. The graph is presented in percentages, where 100% denotes complete satisfaction with the demonstrated prototype.

Examination of the figures shows a general satisfaction with handling and usage of the PDA for this application. The use of the screen for medical data was very satisfactory. In addition, participants agreed on the importance of this application for medical staff. However, the application interface was not intuitive to all participants. The results showed the need for a more detailed instruction manual to ease the use of this application.

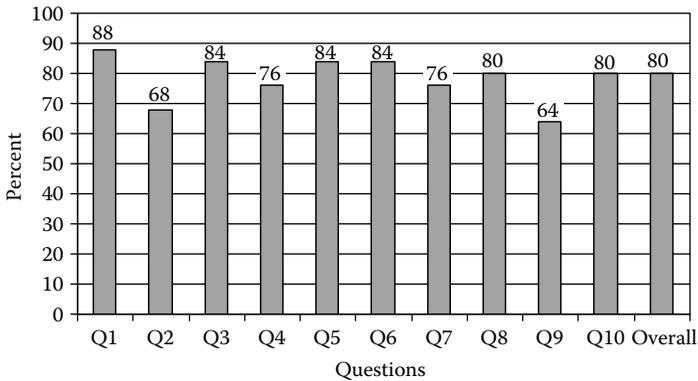


Figure 3.8 Nurses' satisfaction.

3.5 Conclusion

This research resulted in designing a new protocol that conforms to the modern approach that the Department of Health is trying to implement through the Connecting For Health program. A working prototype was developed and tested using off-the-shelf software and hardware. The designed system was based on the use of PDAs that connect to the hospital server using a WLAN. The interfaces were designed as part of the complete system. The first is a doctors' interface that enables them to view the vital signs of the patients and provide relevant consultation. The second is the nurses' interface enabling them to log the patients' information directly into the hospital database. The last is the administrative interface where administrators can specify and control access rights for medical staff. The database and the interfaces were designed to be accessed through Web browsers. This enables the system to be used remotely over any local or wide area network technology available from any Web browser-equipped device. Initial evaluation of the prototype showed general satisfaction with the system. As a result of the feedback, minor modifications to the user interface are to be incorporated before wider trials are conducted.

The designed system was platform independent because it uses standard Web technologies. This simplifies the upgrade process for incorporating newer technologies for increased performance or security.

A pre-evaluation of the system was conducted to pinpoint the deficiencies and the strengths of the system. Medical staff showed general satisfaction with the system. However, the need for clearer instruction and documentation for the use of the PDA in addition to the system have been recognized. Further and more comprehensive evaluations are to be conducted when the system is in the final prototype stage.

References

1. R. Spanjers, M. Smiths, and W. Hasselbring, Exploring ICT-enabled networking in hospital organizations, in R.A. Stegwee and T.A. Spil (Eds.), *Strategies for Healthcare Information System*, Hershey, PA: Idea Group Publisher, 2001, p. 155.
2. A.E. Carroll, S. Saluja, and P. Tarczy-Hornoch, The implementation of a personal digital assistant (PDA) based patient record and charting system: Lessons learned, *Proceedings of the AMIA Symposium*, 2002, pp. 111–115.
3. Z. Hunaiti, A. Rahman, Z. Huneiti, and W. Balachandran, 3G Mobile Health System, Second International Conference on Cybernetics and Information Technologies, Systems and Applications: CITSA 2005, Orlando, Florida, July 14–17, 2005.
4. Z. Hunaiti, A. Rahman, Z. Huneiti, and W. Balachandran, Mobile Medical Data Access System, Fifteenth International Conference on Electronics, Communications, and Computers: CONIELECOMP 2005, Puebla, Mexico, February/March 2005, pp. 2–6.
5. UK National Health Service (NHS), vol. 2006, 04/07/2006 (2006). Available at <http://www.connectingforhealth.nhs.uk/>
6. E. Turban, D. Leidner, E. McLean, and J. Wetherbe, *Information Technology for Management: Transforming Organizations in the Digital Economy*, John Wiley & Sons Inc, 2005, p. 784.

CARDIOLOGY

2

Chapter 4

Mobile, Secure Tele-Cardiology Based on Wireless and Sensor Networks

Fei Hu, Laura Celentano, and Yang Xiao

CONTENTS

- 4.1 Introduction..... 64
- 4.2 Significance of Next-Generation Wireless Networks for
Tele-Cardiology65
- 4.3 Results on Tele-Cardiology Based on Integrated
Wireless Networks.....67
 - 4.3.1 Routing in Simplified Heterogeneous Wireless
Tele-Cardiology Networks.....67
 - 4.3.2 Performance Analysis68
- 4.4 Cardiac Monitoring Based on Wireless Sensor Networks69
- 4.5 MSN-Based Tele-Cardiology Design73
 - 4.5.1 Low-Power, Small-Size ECG Micro-Sensors.....73
 - 4.5.2 Secure ECG Transmission.....75

4.5.2.1 Single-Patient Case75
4.5.2.2 Multi-Patient Case.....78
4.6 Conclusions 80
Acknowledgment81
References81

In this chapter, we first present a mobile tele-cardiology architecture that is based on the next-generation wireless networking platforms, which are able to switch among different wireless networks (including cellular networks, wireless LAN, WiMAX, ad hoc networks, and satellite networks) seamlessly and automatically when cardiac patients move to different locations (at home, large buildings, suburbs, or highways). Then, we discuss the importance of wireless sensor networks in cardiac monitoring. Finally, we provide our tele-cardiology results on secure ECG signal transmission based on a Skipjack cryptography algorithm.

4.1 Introduction

Health care costs, coverage problems, and demographic pressures mean system overload; its formal institutions can't cope with the future. What will ease the pain? A major shift, enabled by technology, to self-care, mobile care, home care.

— Forrester Research¹

Cardiovascular diseases are among the most widespread health problems and the single largest cause of morbidity and mortality in the United States and the Western world.² Based on the World Health Report 2000,³ each year coronary artery disease kills an estimated 7 million people representing 13% of all male deaths and 12% of all female deaths. No country spends more per capita on health care delivery than the United States. The entire nation has doubled its health care expenditure over the last two decades. Thus, low-cost and high-quality cardiac health care delivery is a critical challenge.

Tele-health monitoring can be defined as “mobile computing, medical sensor, and communications technologies for health care.” This represents the evolution of e-health systems from traditional desktop “telemedicine” platforms to wireless/mobile configurations. Tele-health for cardiac monitoring would largely benefit our society (1) by enhancing accessibility to care for underserved populations (such as in rural/remote areas), (2) by containing cost inflation as a result of providing appropriate care to cardiac patients in their homes/communities, and (3) by improving quality as a result of providing coordinated and continuous care for cardiac patients and highly effective tools for decision support.

Among commercial telemetry systems, CardioNet is the first provider of mobile cardiac outpatient telemetry (MCOT) service in the United States for continuous monitoring of a patient's ECG and heartbeat.^{4,5} It automatically detects and transmits (using cellular networks) abnormal heart rhythms to the CardioNet monitoring center, where certified cardiac technicians analyze the transmissions and respond appropriately 24 hours a day, 7 days a week, 365 days a year. Philips provides a telemetry system, which consists of a portable TeleMon companion monitor and the IntelliVue information center to offer an integrated surveillance and monitoring solution for ambulating patients who require vigilant oversight of ECG and SpO₂.⁶ The GMP Wireless Medicine Corp. has developed the LifeSync wireless ECG system for bedside monitoring.^{7,8} The systems described in Geier⁹ and ten Duis and van der Werken¹⁰ show promising research in using wireless LAN (WLAN) radios in buildings to transfer cardiac patient information to the hospital in real-time, and analyze the signal interference with other available WLAN hotspots. A cardiac patient monitoring solution using ad hoc networks, which can be formed dynamically among mobile and wearable devices, is presented in Varshney.¹¹ It is limited within a small transmission distance. A third-generation universal mobile telecommunications system (UMTS) solution for the delivery of cardiac information from an ambulance to a hospital is presented in Gállego et al.¹² A combined hardware and software platform known as CodeBlue¹³ provides protocols for device discovery and publishes/subscribes multi-hop routing as well as a simple query interface for medical monitoring. Other examples of sensor-based cardiac monitoring systems are SMART¹⁴ and WiSARD.¹⁵ A real-time patient monitoring system that uses vital signs and location sensors, ad-hoc networking, electronic patient records, and Web portal technology to allow remote monitoring of patient status was designed and developed in Gao et al.¹⁶ The Advanced Health and Disaster Aid Network (AID-N),¹⁷ being developed at the Johns Hopkins University, explores and showcases how these advances in technology can be employed to assist victims and responders in times of emergency. The comprehensive overview of some of these existing wireless telemedicine applications and research can be found in recent publications.^{18–21}

4.2 Significance of Next-Generation Wireless Networks for Tele-Cardiology

The evaluation of the above prototypes points to a number of critical areas for future tele-cardiology developments. One of the biggest shortcomings of most existing cardiac monitoring systems is that they are based on a single type of wireless network (most of them use cellular networks and some others use WLANs in buildings). However, a reliable cellular connection may not be available in many areas such as rural areas/roads, mountains, and recreational forests (due to poor

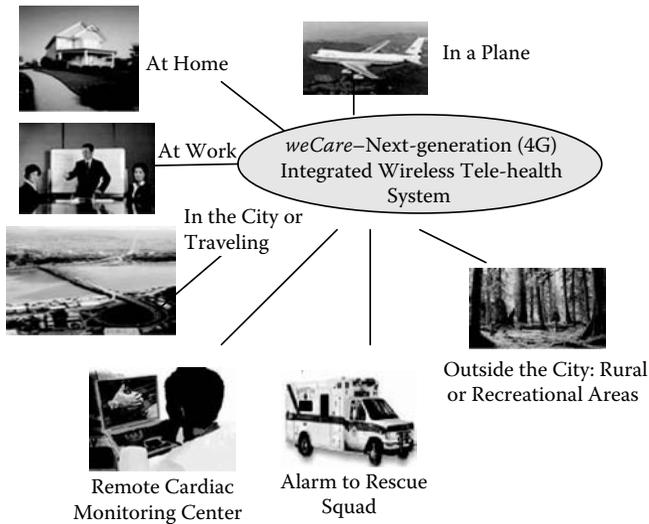


Figure 4.1 Achieve “anywhere anytime” cardiac care through integrated wireless networks.

coverage); indoor and underground areas (due to dead spots); and higher call blocking in dense areas (due to hotspots).

Similarly, WLANs are available only inside buildings and their signals experience attenuation/loss due to fading, multipath interference, and multipath distortion. Thus the cellular or WLAN networks alone cannot achieve “anywhere anytime” cardiac monitoring (shown in Figure 4.1).

Moreover, the health provider will need to install many telephone lines for receiving tele-health data when many patients use cellular networks in the future. On the other hand, these health providers already have high-speed Internet connections.

Another serious issue is the lack of comprehensive wireless quality-of-service (QoS) support (including not only delay, bandwidth, and jitter but also packet loss rate, cellular call dropping rate, and other metrics) in a single wireless network.

Another area worth exploring is the impact of wireless bandwidth limitations and effective techniques for sharing bandwidth across patient sensors. For example, certain types of queries and patient data could be assigned a higher priority to give it better resources than others in the presence of radio congestion.

The tele-health monitoring systems can be made more reliable and economic, if the patient data and hospital command/control/query data could be transmitted through different wireless communication technologies (instead of a single type) depending on the patient location, network availability, and other QoS requirements. Note that each type of wireless network has different data rates, end-to-end delay, radio coverage range, deployment cost, and allowable user mobility level (see Table 4.1).

Table 4.1 The Features of Different Wireless Networks That Could Be Used for Cardiac Monitoring

	<i>Radio Coverage</i>	<i>End-to-End Delay</i>	<i>Data Transmission Rates</i>	<i>Allowable Patient Mobility</i>	<i>Deployment Cost</i>
Cellular networks	Approx. 35 km	Medium/high	144 kbps ~ 1 Mbps	High	High/very high
WiMAX	Approx. 20 km	Low	Approx. 10 Mbps	Very high	Medium/high
WLAN	50 m ~ 300 m	Very low	11 ~ 54 Mbps	Medium	Medium
Satellite	World	Very high	< 144 kbps	High	Very high
Ad hoc networks	Typically < 1 km	Low/medium	300 kbps ~ 2 Mbps	Medium	Low

We can fully utilize the features of different wireless networks to design an “anywhere, anytime, real-time” cardiac tele-health system. For example:

- When a cardiac patient stays at home, some home Internet access technologies (such as DSL, cable modem, etc.) can be used to send the patient’s data to the health provider.
- When a patient is driving/shopping in the city, the cellular network or WiMAX may be a better choice because it has long-distance, high-speed radio communication.
- When the patient is at work or stays in a nursing home or hospital, typically WLAN (high-speed, covers building range) is available for local wireless Internet transmission.
- Cellular networks can also be used for cardiac data transmission when out of the WiMAX radio range, say, in suburban areas.
- Satellite networks could be the only choice when traveling in a plane or a desolate place.
- Ad hoc networks could be used to organize a temporary small area hop-to-hop network when many patients are close to each other and other networks are not available.

4.3 Results on Tele-Cardiology Based on Integrated Wireless Networks

4.3.1 Routing in Simplified Heterogeneous Wireless Tele-Cardiology Networks

We have investigated the routing scheme in a mini weCare scenario, which integrates the cellular networks and ad hoc networks together, shown in Figure 4.2. A

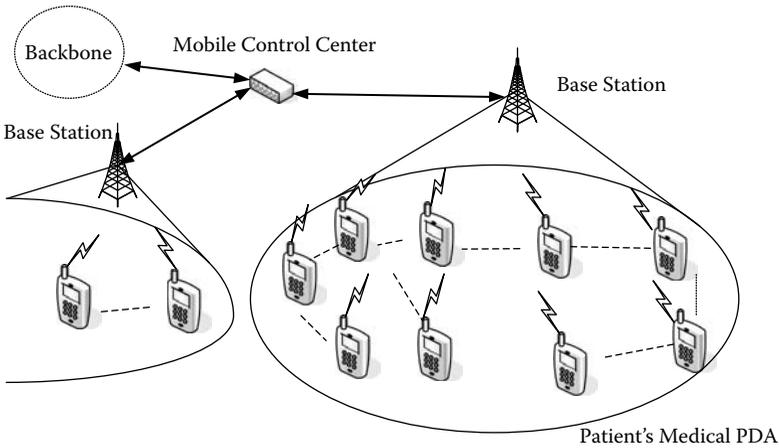


Figure 4.2 The integration of cellular networks and ad hoc networks.

cardiac patient’s monitoring device utilizes other patients’ monitoring devices to relay ECG data in a hop-by-hop topology until it finally reaches a cellular base station (BS). The BS can then transmit the ECG data to the health provider. Because the ad hoc network can forward data to its neighboring devices at a short distance at up to 1 Mbps, the data rate is much higher than the long-distance direct device-to-BS data forwarding (the cellular network data rates < 300 kbps).

To find a shortest multi-hop path from a patient’s device to the BS quickly, we have designed a dynamic routing scheme that adopts the controlled flooding approach for route discovery. First, the source cardiac monitoring device broadcasts a route discovery packet. The intermediate devices that receive this packet will rebroadcast it until it reaches the cellular BS. The BS sends a route reply packet to the source device and thus a route is formed, which is recorded in the routing table at the source device. To satisfy individual patient’s delay and data rate requirements or maintain the system efficiency, our routing protocol can allocate different paths for adaptive adjustment. For example, if a patient requires short delay and low data rate, our routing protocol can choose the cellular way (i.e., direct PDA-to-BS communication). On the other hand, if a patient requires transmission of high data rates, our routing protocol can choose the hop-to-hop relay for the source PDA.

4.3.2 Performance Analysis

We have conducted network simulation using OPNET²² to analyze the data rate and delay. We assume that the ECG data traffic follows the Poisson distribution model and each cellular “cell” covers 0.5 miles. We assume the cellular network communication rate as 1 Mbps, which is achievable under 4G wideband CDMA

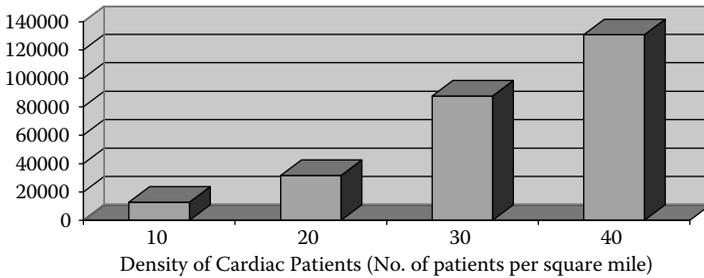


Figure 4.3 ECG throughput (bits per second) vs. patient density.

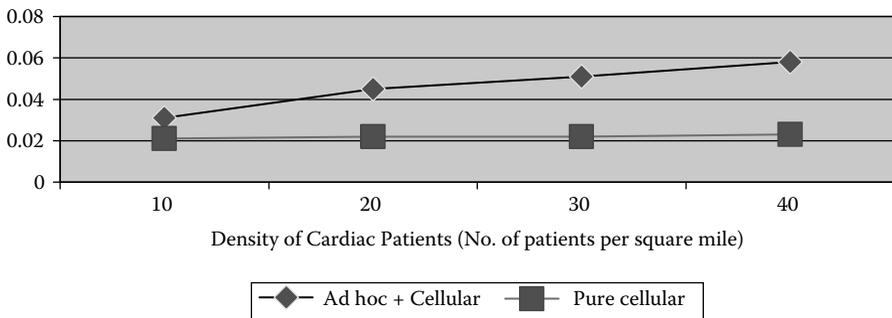


Figure 4.4 ECG transmission delay (seconds) versus patient density.

environments. The ad hoc network throughput is assumed to be 2 Mbps, which is achievable when the average communication distance between two neighbors is less than 200 feet and the radio frequency is 2.4 GHz. Figure 4.3 shows the relationship between ECG data throughput (bits per second) and the density of cardiac patients (number of patients per square mile). Figure 4.4 shows that the end-to-end ECG communication delay could be longer when using hop-to-hop relay than in pure cellular network.

4.4 Cardiac Monitoring Based on Wireless Sensor Networks

A cardiac patient with “multiple” health conditions needs a special telemedicine platform that is able to collect “multiple” health parameters from the patient’s body automatically and then send a timely alert to a remote health care office if those parameters are beyond normal ranges. Those health parameters include heart rate (HR), blood oxygenation level (SpO₂), blood pressure (BP), and so on. Table 4.2 shows a partial list of physiological conditions that may cause medical alerts.^{23–26}

Table 4.2 Multiple Health Care Parameters That Could Cause Alerts

<i>Alert Type for Patients with Multiple Health Conditions</i>	<i>Detection Parameter That Goes beyond Normal Range</i>
Low SpO2	SpO2 < 90% (default values, adjustable)
Bradycardia	HR > 40 bpm (default values, adjustable)
Tachycardia	HR > 150 bpm (default values, adjustable)
HR change	$ \Delta\text{HR per 5 minutes} > 19\%$
HR stability	Max HR variability from past 4 readings > 10%
BP change	Systolic or diastolic change > $\pm 11\%$

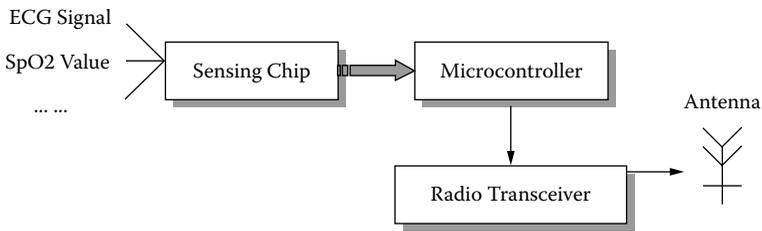


Figure 4.5 MSN sensor hardware components.

Recently, a promising wireless telemedicine technology called a medical sensor network (MSN)^{27–33} has been proposed to monitor changes in patients’ vital signs closely and provide feedback to help maintain an optimal health status.

As shown in Figure 4.5, an MSN sensor typically includes a sensing chip to sense health care parameters, a microcontroller to perform local data processing (such as data compression) and networking operations (such as communicating with a neighbor sensor), and a radio transceiver to send/receive health care sensed data wirelessly. The entire MSN sensor is powered by batteries with a lifetime of several months. Because the power storage is limited, it is very important to use “low-energy” MSN networking operations.

The MSN sensors can improve the health care quality greatly because the automatic, wireless health care data transmission can avoid patients’ frequent doctor visits and labor-intensive manual health care parameter collections. Such sensors are also important to capture medical emergency events. For instance, many serious heart problems affecting older people are transient and infrequent and can go unnoticed even by the patients. A sudden slowing of the heart rate that leads to

a fainting spell may last less than a minute and occur only once or twice a week. That is often enough to make driving a car dangerous but not frequent enough for a doctor to spot during a checkup or even by using a portable 24-hour ECG recorder called a Holter monitor. Another problem, the uncoordinated quivering of the small upper chambers of the heart, a leading cause of stroke in people over 70, can be both infrequent and without obvious symptoms. Therefore patient-triggered ECG recorders could easily miss it. However, our MSN ECG sensor can automatically collect ECG data and trigger an alert to the doctor if the ECG data mining software detects an anomaly.

Note that an MSN sensor is different from traditional wearable medical devices that are also marketed as “portable”—but this does not always indicate that they are small and have wireless communications capability. Most such appliances are much heavier and larger than an MSN sensor that can be conveniently attached to a patient’s body.

We have designed a practical MSN that has the following characteristics:

- Our MSN is able to collect multiple health care parameters continuously from a patient with multiple health conditions. If a patient has multiple chronic diseases, it is important to monitor multiple body parameters, shown in Table 4.2, instead of just one type of data. Our MSN can perform multi-sensor data collection as shown in Figure 4.6. Each sensor node can sense, sample, and process one or more physiological signals. For example, an ECG sensor can be used for monitoring heart activity, an electromyogram (EMG) sensor for monitoring muscle activities, a neuroelectroencephalogram (EEG)

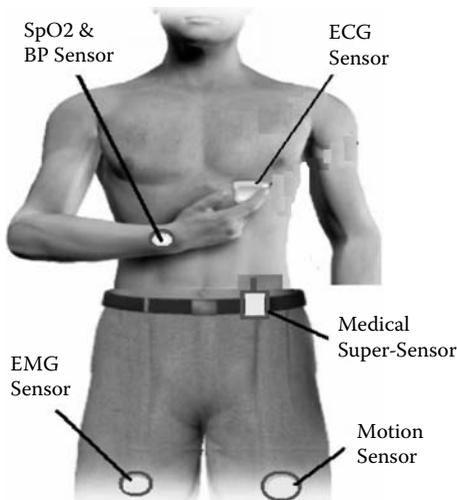


Figure 4.6 Multiple sensors.

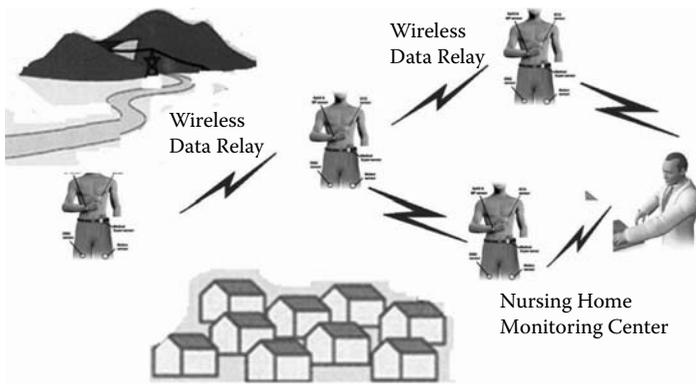


Figure 4.7 Medical sensor networks for nursing homes.

sensor for monitoring brain electrical activity, a blood pressure sensor for monitoring blood pressure, and a breathing sensor for monitoring respiration. In the next section, we will further explain the hardware architecture of those sensors.

- Our MSN uses a wireless body area network (WBAN) in each patient's body to perform multisensor data integration. A medical super-sensor (MSS), shown in Figure 4.6, is a WBAN integration center that can use a radio frequency to communicate with all body sensors. It can also use another radio frequency to send data to a center.
- Our MSN can be applied in large U.S. nursing homes through a self-managed, relay-based wireless communication architecture. We built an MSN hardware/software system that is suitable to large nursing homes with a radius of 300 to 1000 feet. Because each patient's MSS has limited wireless communication range (typically less than 100 feet) due to the low-power transceiver and tiny antenna in each sensor, this project will design a patient-to-patient (i.e., hop-to-hop) wireless transmission relay scheme. That relay scheme can automatically search neighbor patients' MSS and use them to relay the medical sensed data to a remote medical monitoring center, shown in Figure 4.7. If the distance is less than 100 feet, an MSS can directly communicate with the center.
- Our MSN design also considers patients' mobility behaviors. If the patient moves around in a nursing home, our dynamic MSN routing protocol can automatically search a new patient-to-patient path to send the remote patient's data to the monitoring center.

In summary, our nursing home MSN has automatic wireless network management (such as neighbor discovery, mobility, adaptivity, etc.) and multisensor data transmission functions.

4.5 MSN-Based Tele-Cardiology Design

4.5.1 Low-Power, Small-Size ECG Micro-Sensors

Through the collaboration with Sensorcon Inc.,³⁴ Crossbow Inc.,³⁵ and the CodeBlue research group at Harvard,³⁶ the authors have led the researchers in the Wireless Networking Lab in the Computer Engineering Department at RIT to develop a prototype of ECG sensor networks. Our ECG micro-sensor shown at the top of Figure 4.8 has three leads that attach to the patient's upper and lower chest; one lead serves to bias the patient's skin properly and the other two are used to measure cardiac activity. To send out the measured ECG data to a remote ECG server that is around 300 to 600 feet away, we have made a low-cost RF board, shown at the bottom of Figure 4.8, that has the following features:

- The heart of the RF board is the microcontroller/radio transceiver unit. There are two options to incorporate the two devices together: either using a separate microcontroller/transceiver chip or using a system-on-chip (SoC). We chose the SoC option as it is cheaper to implement, decreases programming complexity, and creates an easier printed circuit board (PCB) layout as there will be fewer parts. The Ember EM250³⁷ SoC was selected for use.
- The design of the RF board costs less than \$20 in production of 5000+ units. It is much cheaper than the Crossbow programmable RF motes³⁵ that cost more than \$120 each. It is also significantly smaller than existing portable hospital ECG monitors. The entire RF board plus the ECG sensor can fit in a package measuring $6 \times 4 \times 4$ cm.

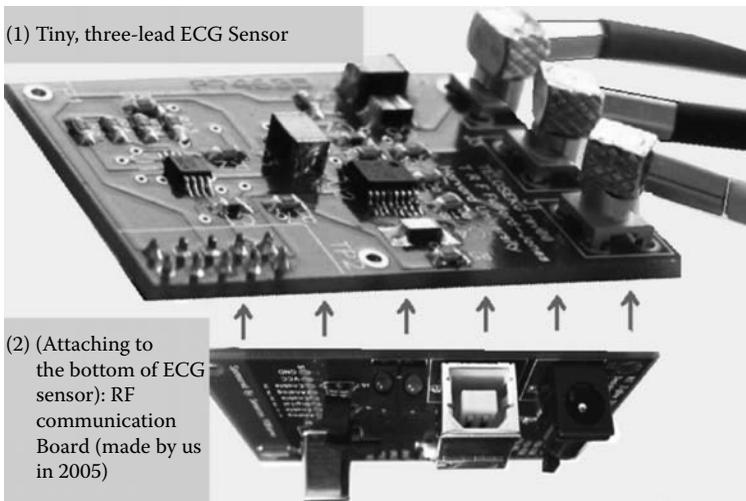


Figure 4.8 (Top) Three-lead ECG sensor. (Bottom) Wireless communication board.

- The RF board is attached to the bottom of the ECG sensor, shown in Figure 4.8. It can run sensor networking protocols such as automatic searching for a next-neighbor RF board (located in another patient's body) to help relay its ECG data if the ECG server is more than 600 feet away.

The three-lead ECG micro-sensor (shown at the top of Figure 4.8) operates from a supply of +3 V because it shares the two AA batteries with the RF board. Its circuit diagram is shown in Figure 4.8. The circuit can provide a dequate amplification of signals that are in the range of -5.0 to $+5.0$ mV because the cardiac tissue is not capable of generating signals greater than 5.0 mV in magnitude. At the heart of the ECG sensor design is Texas Instruments' INA321 integrated circuit (IC), a micropower single-supply CMOS instrumentation amplifier with a very favorable CMRR of 94 dB. The schematic consists of a biasing divider (R3, R7) which splits the supply in half and connects to the lower-left chest lead (red) as a means of setting the patient to the correct potential. The other two leads (colored white and black, respectively) receive signals from the patient's upper chest and transmit the resulting differential signal to the INA321 device.

We have also built a set of ECG networking software modules including wireless routing (in each RF board), ECG security (in both the RF board and the ECG server), ECG database management (in the ECG server), and others. Figure 4.9 shows a screen shot of our GUI (graphical user interface) in the ECG server. Our sensor network protocols can keep track of the location of each patient based on the MoteTrack algorithm.³⁶ In Figure 4.9, we can see that our software can monitor the

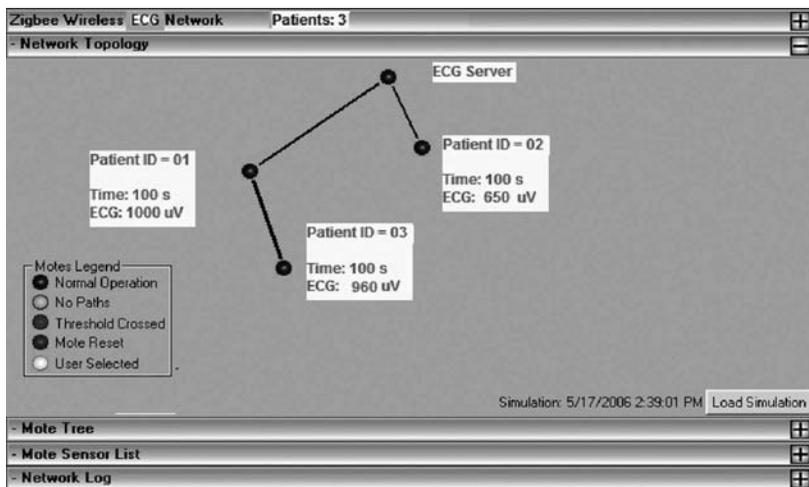


Figure 4.9 Cardiac monitoring software for a simple nursing home with three cardiac patients.

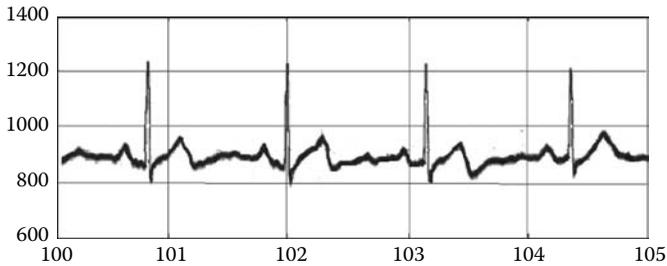


Figure 4.10 ECG signal waveform (ECG amplitude (nV) versus seconds).

sensor network topology. The topology has a tree-like architecture with the ECG server as the tree “root.”

Our software can sample the ECG signal using the built-in RF board ADC (analog-to-digital) that is capable of performing 10-bit digitization, which is comparable in accuracy to the specifications of commercial ECG devices. It is essential that an ECG sensor captures all information because medics may be interested in intermittent abnormalities that can occur at any time. Transmission of each ECG data packet occurs regularly and no more frequently than once every 50 msec. Because other patients’ ECG sensors may be in the vicinity during the operation of ECG sensor network, it would not be appropriate to send ECG data with a higher frequency than once per 50 msec, which will risk corrupting the information being sent to and from other patients’ RF devices.

A sample ECG trace captured from our ECG sensor is shown in Figure 4.10. Integrating an ECG micro-sensor on an RF board with the sensor network software yielded promising preliminary results and confirmed that a 120-Hz ECG sample rate could indeed be achieved using the differential encoding scheme. However, the resulting received waveform was nonoptimal owing to the fact that the RF board’s radio broadcasts caused interference with the analog ECG capture circuitry. This unexpected shortcoming will be remedied by the use of appropriate radio shielding techniques in the future.

4.5.2 Secure ECG Transmission

4.5.2.1 Single-Patient Case

To protect the two important aspects of cardiac patient “privacy” in tele-cardiology systems, i.e., (1) confidentiality (only the source/destination can understand the medical data through crypto-keys), and (2) integrity (no data falsifying during transmission), we need to apply strong end-to-end security mechanisms to the cardiac data packets that are transmitted between any two network entities, such as between a patient’s sensor and a physician’s PC. On the other hand, in a practical community/hospital tele-cardiology system that is based on MSN architecture

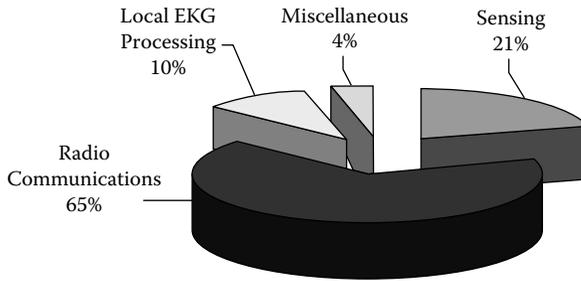


Figure 4.11 Power consumption sources.

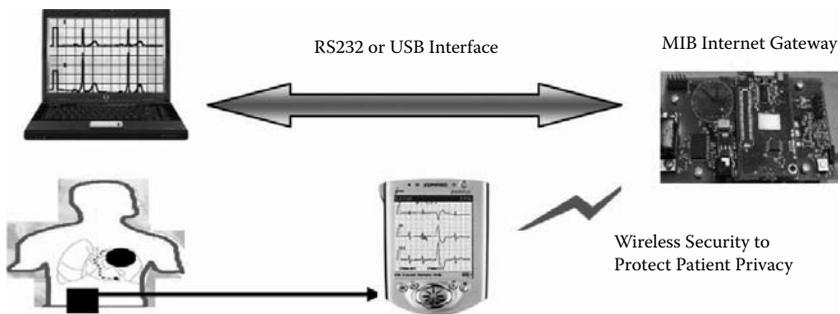


Figure 4.12 Tele-cardiology MANET security: Single-patient case.

as stated in Section 4.5.1, we should consider the following two constraints when designing privacy-preservation mechanisms:

1. *Low-energy/low-overhead security schemes.* A major concern in medical security schemes design is energy efficiency. Our experiments have shown that most sensor battery power is consumed in radio communications instead of in ECG signal processing or sensing, shown in Figure 4.11. Therefore, the security schemes should not use too many message exchanges between patients' sensors and the network. Moreover the security schemes should be of low complexity. Therefore symmetric-crypto can be a better choice than traditional asymmetric-crypto based on public/private keys having high computational overhead.
2. *Multi-hop versus single-hop security.* We should use multi-hop wireless relay among patients instead of single-hop communications (i.e., direct patient-doctor wireless forwarding) due to the following reasons. First, by deploying a multi-hop data forwarding network, packets can be routed around radio obstructions in a community. Single-hop, i.e., long-distance (>100 m), line-of-sight radio communications may not be possible. Second, packet forwarding via multiple small radii transmissions requires less energy than a single large-radius transmission for radio communications.^{38,39} The energy savings afforded by multi-hop forwarding would help conserve PDA batteries.

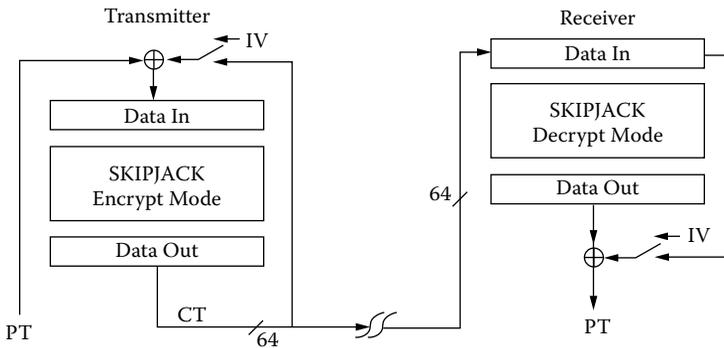


Figure 4.13 Skipjack crypto. PT = plaintext; CT = ciphertext; IV = initialization vector, which works with 64 bits.

Security in each individual hop is the prerequisite of multi-hop MANET security. As the starting point of our security research, we have implemented a low-energy, low-overhead security scheme for one-hop (e.g., patient-to-doctor) wireless communications. As shown in Figure 4.12, the security software is built in both the PDA and the mote interface board (MIB) that serves as the transition gateway between MANET and the Internet.

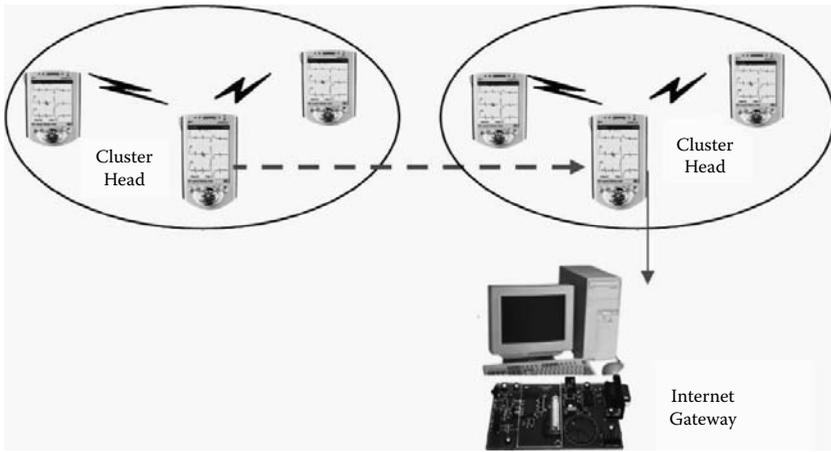
Our one-hop security mechanism uses the following two security primitives shown in Figure 4.13:

1. Initialization vectors (IVs): One implication of semantic security is that encrypting the same plaintext two times should give two different ciphertexts. The main purpose of IVs is to add variation to the encryption process when there is little variation in the set of messages.
2. Block cipher choice: Triple-DES⁴⁰ is too slow for software implementation in embedded medical PDAs or sensors. We found RC5⁵⁰⁴⁰ and SkipJack⁴¹ to be most appropriate for embedded microcontrollers. Although RC5 is slightly faster, it is patented; also, for good performance, RC5 requires the key schedule to be pre-computed, which uses 104 extra bytes of RAM per key. Because of these drawbacks, we selected SkipJack.

It is difficult to measure energy consumption of security mechanisms directly from PDAs. We have thus resorted to an accurate simulator called Power Tossim,⁴² where hardware peripherals (such as the radio, EEPROM, LEDs, and so forth) are instrumented to obtain a trace of each device's activity during the simulation run. Through the obtained real-time traces of the current drawn in our SkipJack-based symmetric crypto and RSA-based symmetric crypto,⁴⁰ we have computed the energy consumption of major components (such as CPU idle, CPU active, radio, etc.) in PDAs, as shown in Table 4.3.

Table 4.3 Security Energy Consumption Comparisons

	<i>SkipJack (mJ)</i>	<i>RSA (mJ)</i>
CPU active	26	51
Radio	1002	2542
Memory access	11	25
Total	1680	3360

**Figure 4.14 Multi-patient case: Cluster-based security.**

From Table 4.3, we can see that for the two most important components, i.e., CPU active and radio transmission, our proposed security scheme shows significant power-saving improvements over the RSA security scheme and the energy efficiency is improved by 92 and 154%, respectively.

4.5.2.2 Multi-Patient Case

To get closer to the real tele-cardiology MANET scenario, we have extended the above single-patient transmission security to a multi-patient case. Although it is currently a fixed, small MSN (with only a few sensors), it would serve as the basis of our future research work on a large-scale MSN.

It is challenging to deliver data securely from a remote sensor to an Internet gateway through multi-hop transmission as it requires integration of the security scheme with energy-efficient MSN routing schemes. As shown in Figure 4.14, we partition patients' sensors into a number of "clusters." In each cluster, exactly one sensor is chosen as the cluster head (CH). Thus each sensor only needs one-hop communication to send the ECG signals to its CH, which searches for a neighboring CH for data relay to the gateway. This cluster-based concept has also been used in many

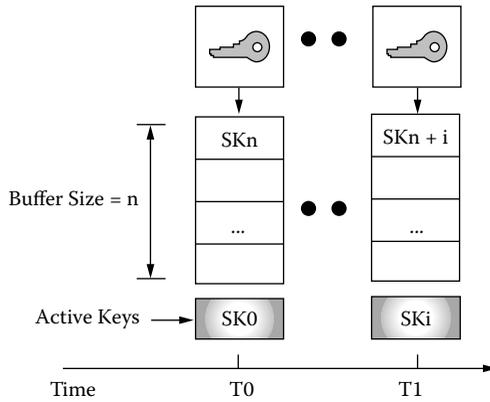


Figure 4.15 Keychain among CHs.

hierarchical routing MSN schemes to save energy. To avoid the battery overusing in a CH, the selection of CH could be rotated periodically among the sensors belonging to the same cluster.

We have used the aforementioned SkipJack to achieve intra-cluster security (i.e., inside each cluster). For secure data transmission between clusters, an inter-cluster session-key (SK) is used, shown in Figure 4.15. A new SK is periodically distributed to all CHs by the gateway. All new SKs are derived from a one-way hash function $H(\cdot)$. The gateway first pre-computes a long one-way sequence of keys: $\{SK_M, SK_{M-1}, \dots, SK_n, SK_{n-1}, \dots, SK_0\}$ (size $M \gg n$), where $SK_i = H(SK_{i+1})$. Initially, only SK_n (instead of the whole M -size key sequence) is distributed to each CH. Then a CH can utilize $H(\cdot)$ to figure out SK_{n-1}, \dots, SK_0 . The n keys $\{SK_n, SK_{n-1}, \dots, SK_1\}$ are stored in a local key buffer. However, SK_0 is not in the buffer because it is used for the current data packet encryption/decryption. After the initial SK_n delivery, the gateway periodically sends $SK_{n+1}, SK_{n+2}, \dots, SK_M$ (one key distribution in each period) to all CHs.

After receiving a new SK, the CH keeps applying $H(\cdot)$ to it for some time, in order to find a key match in its key buffer. For instance, assume that a CH receives a new key SK_j and its key buffer already holds n SKs as follows: $\{SK_j, SK_{j-1}, \dots, SK_{j-n+1}\}$. If

$$H(H(H \dots (H(SK_j))) \notin \{SK_j, SK_{j-1}, \dots, SK_{j-n+1}\} \quad (4) \quad .1)$$

the authentication fails and the SK_j will be discarded. Otherwise, if the authentication is successful, the key buffer is shifted one position and the SK shifted out of the buffer is pushed into the “active key slot” to be used as the current SK (Figure 4.15). The empty position is filled with a new key SK' , derived from the received SK_j through H , which meets the following two conditions:

$$SK' = H(H(H \dots H(SK_j))), \text{ and } H(SK') = SK_j \quad (4) \quad .2)$$

4.5.2.2.1 Security Analysis

There are several attacks listed as follows:

- *Gateway attacks*: Because the distribution of new SKs is managed by the gateway, it is possible for an attacker to compromise the gateway and thus attack any future SK disclosures. Thanks to the SK buffer, there is a delay between receiving the new SK and actually using it. If the distribution interval is Δ' (i.e., the re-keying period) and n is the buffer length, the new SK will not be used until $n \times \Delta'$ later. As long as we can detect the gateway compromise within $n \times \Delta'$ time interval and renew SKs, the cardiac data packets will maintain security performance.
- *SK attacks among CHs*: The attacker may modify the transmitting SK, inject a phony SK, or use wireless channel interference to damage security packets. Our scheme can easily defeat these attacks. Thanks to the one-way characteristics of the hash function keys, any false SKs cannot pass the authentication test, that is, after L times ($L \leq n$) of using hash function, if we still cannot satisfy the following formula (in which SK_{FAKE} is a false SK and SK_{NOW} is the currently used SK), we will regard that it is a false SK:

$$\underbrace{H(H(\dots(H(SK_{FAKE})\dots)))}_L = SK_{NOW} \quad (4.3)$$

- *Cardiac packet attacks* (such as faking the ECG data): Our scheme defeats it through SK re-keying every Δ' , and inclusion of *Sensor_ID* and per-packet *IV* (which will also be updated from packet to packet) in the generation of key-streams to counter the key-stream reuse problem.
- *Man-in-the-middle attacks*: Our scheme can also defeat man-in-the-middle attacks (where an attacker fools the CHs as if he or she were a legal CH). Our strategy is to perform a transmission of *MAC* in the re-keying procedure as follows: $Gateway \rightarrow CH : E(\Delta' | n | SK_0 | MAC(\Delta' | n | SK_0))$

4.6 Conclusions

Mobile telemedicine is an active research and development field. This chapter summarized the tele-cardiology systems based on advanced wireless networks. This includes two aspects: (1) using integrated wireless networks (such as wireless LAN, cellular networks, WiMAX, and so on) to transmit ECG and other cardiac data to any place; and (2) using low-power sensor networks to collect ECG data remotely. We have also reported our research results on secure ECG transmission in sensor networks.

Acknowledgment

This work was partially supported by the U.S. National Science Foundation (NSF) under grants CNS-0716211 and CNS-0716455.

References

1. On the Significance of Tele-healthcare, Forrester Research, available at <http://www.forrester.com/ER/Research/Brief/Excerpt/0,1317,15452,00.html>
2. M.G. Hunink, L. Goldman, A.N. Tosteson, M.A. Mittleman, P.A. Goldman, L.W. Williams, J. Tsevat, and M.C. Weinstein, The recent decline in mortality from coronary heart disease, 1980–1990. The effect of secular trends in risk factors and treatment, *JAMA*, 277, 535–542, 1997.
3. WHO World Health Report 2000, see <http://www.who.int/whr/2000/en/>
4. V. Shnayder, B. Chen, K. Lorincz, T.R.F.F. Jones, and M. Welsh, Sensor Networks for Medical Care, Technical Report TR-08-05, Division of Engineering and Applied Sciences, Harvard University, 2005.
5. SMART Project Overview web page: <http://smart.csail.mit.edu/>
6. L. Hauenstein, T. Gao, and D. White, Service-oriented architecture for disaster response: Integration of AID-N, MICHAELS, WISER, and ESSENCE, *Proc. American Medical Informatics Association Annual Conference (AMIA 2006)*, Washington, D.C., November 2006.
7. T.C. Chan, J. Killen, W. Griswold, and L. Lerner, Information technology and emergency medical care during disasters, *Acad. Emergency Med.*, 11, 1229, 2004.
8. E. Shih, V. Bychkovsky, D. Curtis, and J. Guttag, Demo abstract: Continuous, remote medical monitoring, *Proc. Second Annu. Int. Conf. on Embedded Networked Sensor Systems*, November 2004.
9. J. Geier, Saving lives with roving LANs, *Network World*, Feb. 5, 2001; <http://wireless.itworld.com/4246/NW0205bgside/pfindex.html>
10. H.J. ten Duis and C. van der Werken, Trauma care systems in The Netherlands, *Proc. Injury—Int. J. Care of the Injured*, 34(9): 722–727, 2003.
11. U. Varshney, Using wireless networks for enhanced monitoring of patients, *Proc. 10th Americas Conf on Information Systems*, August 2004, New York, 1–7.
12. J. R. Gállego, Á.H. Solana, M. Canales, J. Lafuente, A. Valdovinos, and J.F. Navajas, Performance analysis of multiplexed medical data transmission for mobile emergency care over the UMTS channel, *IEEE Trans. Info. Technol. Biomed.*, 9(1), March 2005.
13. Crossbow Inc. (producing wireless sensors), see <http://www.xbow.com>
14. L. Ohno-Machado et al., SMART: Scalable Medical Alert Response Technology, <http://smart.csail.mit.edu/>
15. L. Lerner et al., WiSARD: Wireless Internet Information System for Medical Response in Disasters, <https://wiisard.org>
16. T. Gao, D. Greenspan, M. Welsh, R.R. Juang, and A. Alm, Vital signs monitoring and patient tracking over a wireless network, *Proc. 27th Annu. Int. Conf. of IEEE EMBS*, September 2005, Shanghai, China.

17. D. White et al., AID-N: Advanced Health and Disaster Aid Network, <https://secwww.jhuapl.edu/aidn/>
18. C.S. Pattichis, E. Kyriacou, S. Voskarides, M.S. Pattichis, R.S.H. Istepanian, and C.N. S chizas, Wireless telemedicine systems: An overview, *Proc. IEEE Antennas Propagat. Mag.*, vol. (2): 143–153, 2002.
19. T.F. Budinger, Bio monitoring with wireless communications, *Proc. Ann u. Rev. Biomed. Eng.*, 5, 383–412, 2003.
20. R.S.H. Istepanian and H. Wang, Telemedicine in UK, European Telemedicine Glossary of Concepts, Standards Technologies and Users, 5th ed, L. Beolchi, Ed. Brussels, Belgium: European Commission—Information Society Directorate-General, 2003, 1159–1165.
21. R.S.H. Istepanian and J. Lecal, M-Health systems: Future directions, *Proc. 25th Annu. Int. Conf. IEEE Engineering Medicine and Biology*, Cancun, Mexico, September 2003, pp. 17–21.
22. Network Simulation Tool—OPNET, see <http://www.opnet.com>.
23. P.R. Ross, Managing care through the air, *IEEE Spectrum*, 26–31, December 2004.
24. D. Shanit and R.A. Greenbaum, Toward a comprehensive telecardiology monitoring centre for community-based services, *J. Telemed. Telecare*, 3, 60–62, 1997.
25. E. Jovanov, J. Pr ice, D. Raskovic, K. Kavi, T. Martin, and R. Adhami, Wireless personal area networks in telemedical environment, *Proc. 3rd Int. Conf. Information Technology in Biomedicine, ITAB-ITIS 2000*, 22–27.
26. S. Park and S. Jayaraman, Enhancing the quality of life through wearable technology, *IEEE Eng. Med. Biol.*, 2003, 22(3): 41–48.
27. T. Martin, E. Jovanov, and D. Raskovic, Issues in wearable computing for medical monitoring applications: A case study of a wearable ECG monitoring device, *Proc. of The International Symposium on Wearable Computers (ISWC)*, Atlanta, Georgia, 2000, 43–50.
28. Fei Hu, Yu Wang, and Hongyi Wu, Mobile telemedicine sensor networks with low-energy data query and network lifetime considerations, *IEEE Trans. Mobile Comput.*, 5(4): 404–417, 2006.
29. Fei Hu, Sunil Kumar, and Neeraj K. Sharma, Adaptive QoS Provisioning in Large-scale Mobile Sensor Networks for Next-generation Telemedicine applications (MSNT), World Wireless Congress 2004, San Francisco, May 2004.
30. R. Le Blanc, Quantitative analysis of cardiac arrhythmias, *CRC Crit. Rev. Biomed. Eng.*, 14(1): 1–43, 1986.
31. T.M. Mitchell, *Machine Learning*, McGraw Hill, 1997.
32. D. Novak, D. Cuesta-Frau, V. Eck, J.C. Perez-Cortes, and G. Andreu-Garcia, Denoising electrocardiographic signals using adaptive wavelets, *Biosignal*, 18–20, 2000.
33. The Research Resource for Complex Physiologic Signals (Sept. 8, 2004), PhysioNet [online], available at <http://www.physionet.org>
34. Sensorcon Inc. (a sensor network company): see <http://www.sensorcon.com>.
35. Crossbow Inc. (a sensor network company): see <http://www.xbow.com>.
36. D. Malan, T.R.F. Fulford-Jones, M. Welsh, and S. Moulton, CodeBlue: An ad hoc sensor network infrastructure for emergency medical care. *Proc. of the MobiSys 2004 Workshop on Applications of Mobile Embedded Systems (WAMES 2004)* 2004, 12–14.
37. Ember Inc. (a Micro-chip company), see <http://www.ember.com>.

38. L. Kleinrock and J. Silvester, Spatial reuse in multihop packet radio networks, *Proc. IEEE*, 75(1): 156-167, 1987.
- 3 9. M.B. Srivastava, Energy efficiency in mobile computing and networking, *Mobicom Tutorial*, 7, August 2000.
- 4 0. B. Schneier, *Applied Cryptography*, Second Edition, John Wiley & Sons, 1996.
- 4 1. SkipJack specifications: <http://jya.com/skipjack-spec.htm>
42. V. Shnayder et al., Simulating the power consumption of large-scale sensor network applications. In *Proc. SenSys*, 2004.

Chapter 5

Monitoring and Management of Congestive Heart Failure Patients

Sajid Hussain and Saira Majid Dar

CONTENTS

5.1 Sensor Networks for Remote Monitoring.....	86
5.2 Introduction to Congestive Heart Failure.....	87
5.3 Stages of Heart Failure.....	88
5.3.1 Minimal Monitoring (Level 1 Monitoring).....	89
5.3.2 Closer Monitoring (Level 2 Monitoring).....	89
5.3.3 Closer Monitoring and Evaluation for Intervention (Level 3 Monitoring)	91
5.3.4 Aggressive Monitoring and Preparation for Intervention (Level 4 Monitoring)	92
5.4 Smart Monitoring System	92
5.4.1 Monitoring Human Behavior in a Bedroom.....	94
5.4.2 Monitoring Example	99

5.5 Conclusion..... 101
 References101

This chapter deals with patient management strategies for congestive heart failure (CHF) patients. The computing and networking technologies are investigated to provide adaptive management strategies that can prevent further decline in the quality of health as well as, in some cases, recover the health condition. First the enabling technologies and sensor networks are briefly discussed. Second, the CHF is introduced. Third, the management and monitoring techniques for CHF are provided. Fourth, an intelligent architecture for remote monitoring is proposed. Finally, some of the challenges and techniques for implementation are identified.

5.1 Sensor Networks for Remote Monitoring

Due to recent advancement in the electronics industry, wireless sensor networks (WSNs) are used for various applications such as security, military, building automation (HVAC), and environmental and habitat monitoring^{1,2} as well as health care.³ WSNs consist of several battery-constrained sensor nodes equipped with sensor detection capabilities such as room temperature, moisture, light, and mobility as well as some computing, storage, and communication resources; the sensor nodes are commonly known as motes. The health care sensors include oxygen saturation, pulse, EKG, and weight monitoring. Due to the low cost of sensors and convenience of collecting large data samples, WSN-based applications use redundant sensors and frequent data sampling to increase reliability, precision, and robustness. Data mining and machine learning techniques are applied to extract nontrivial information from the sensor data streams. Sensor motes use license-free radio frequencies (916 MHz and 2.4 GHz) for communication; however, the implanted medical devices could use the WMTS band (wireless medical telemetry services, at 608 MHz) to avoid interference with other wireless devices.

As the Web is widely used for ubiquitous, pervasive, and seamless applications, a Web-based application can assist in remote health care applications such as in-home assistance, smart nursing homes, and clinical trials. As data will be collected and reported automatically in real-time, the number of hospitalizations and clinic visits will be reduced, which will assist in minimizing health care costs. For instance, using a database approach,⁴ the following query can be issued from a health care provider’s laptop or a personal digital assistant (PDA):

```
SELECT patient
FROM WagonerMedicalCenter
WHERE (oxygen-sat < 89) AND
```

```
(pulse < 60 OR pulse > 110)  
INTERVAL 30 minutes  
DURATION 4 months
```

In the above query, all the patients of Wagoner Medical Center who satisfy the given criteria of oxygen-saturation (SpO_2) and pulse will be retrieved. Further, the query will be active for the next 4 months, and data will be obtained after every 30 minutes. The above example can be initiated by a Web-based form where the care provider or any other authorized person can retrieve similar results using a standard interactive online form.

In the above example, patients are identified by their corresponding medical center; however, patients can be located or identified by several approaches:

1. Virtual location such as clinic, hospital, room, and lab
2. Geographical data such as GPS coordinates of a patient or geographic region, GPS coordinates of a center and radius (circular), or GPS coordinates of top-right and bottom-left corners (rectangular)
3. Referential, e.g., Dr. Taylor's patients

Although, in the above example, the exact numbers (60 and 110) are used for pulse range, the numbers can be replaced with predefined fuzzy qualitative representation such as very low, low, satisfactory, high, and very high. The results could incorporate imprecision in the accuracy of results, for instance “high probability of low saturation” or “very high probability of high pulse rate.”

5.2 Introduction to Congestive Heart Failure

Congestive heart failure (CHF) is among the top causes of morbidity and mortality in North America. Every year a large number of patients die due to CHF. However, the life expectancy and the quality of living for these patients could be improved by providing smart home care facilities (wireless devices).⁵

The heart is a muscle that contracts and relaxes in order to pump blood through the body. The blood is pumped into the lungs to absorb oxygen. Then this oxygen-rich blood returns to the heart and is pumped to the other organs. Arteries and capillaries carry oxygen-rich blood to the cells and veins and bring the blood back to the heart to get fresh oxygen from the lungs. On average, the heart beats 60 to 90 times per minute to pump 5 liters of blood every minute.

The healthy heart adjusts to the body needs in two ways: speed (low or high) and force (low or high). The heart's performance is affected by age, heart damage, and demand. For instance, when a person is ill, the heart has to work harder to assist in healing. Similarly, during heart attack, some of the muscles die and form scar tissue that cannot pump. Further, the heart workload is also increased due to stress-related activities or feelings. Systolic heart failure is a condition when the heart has

difficulty in pumping blood through the body. Diastolic heart failure, however, is the condition when the heart has trouble in relaxing, which builds pressure inside the heart as well as in the lungs. Edema is a swelling or congestion caused by collection of blood or fluids in some parts of the body such as the abdomen, ankles, feet, or lungs—shortness of breath results because of congestion in the lungs. The symptoms could be temporary and can appear or disappear occasionally.

Although congestive heart failure cannot be cured, it can be managed.^{6,7} A patient's cooperation and proper care can improve the quality of life and can reduce hospitalization. It is extremely important and beneficial if small changes are identified at an early stage. The heart is affected by many factors:

- High fat and cholesterol diet can damage the inside of blood vessels, which narrows the blood vessels so the heart has to work harder to pump the blood. Similarly, high sodium intake can cause more fluid retention, which also causes the heart to work harder.
- Like any other muscle, the heart can be strengthened by proper exercise or strained by excessive or unplanned physical activities.
- Excessive, sudden, or persistent stress can also limit the heart's performance.
- During rest and sleep, the heart slows down and relaxes, which assists in managing the body's demands.

A daily diary can be maintained to monitor the following: changes in breathing, fatigue, weight gain, swelling, pulse, and side effects of any medication. The proper recording and reporting time is very important in managing heart failure. Further, medication should be taken at the right time, even when a person is feeling well. Common medications belong to five groups:

1. *Aldosterone antagonists*: Reduce the stress to the heart and also have a weak diuretic effect.
2. *Angiotensin-converting enzyme (ACE) inhibitors*: Reduce the stress on the heart and may prevent symptoms from becoming worse.
3. *Beta blockers*: Affect the response to some nerve impulses in certain parts of the body, which decreases the need for blood and oxygen; assist the heart to beat more regularly.
4. *Diuretics*: Reduce the amount of fluid in the body, prevent or reduce swelling, shortness of breath, and bloating.
5. *Digoxin*: Increases the strength of the pumping action of the heart.

5.3 Stages of Heart Failure

The first step is to classify the stage of a heart failure. The most accurate classification can be based upon a simple treadmill test with the calculation of metabolic

equivalent expenditure at the nonischemic level and taking into account entities such as respiratory status, BMI (body mass index), baseline activity tolerance, and arthritis.

Using the Framingham criteria,⁸ heart failure can be classified based on subjective symptoms such as shortness of breath, swelling of feet associated with heart failure, nocturnal symptoms, or increasing need for a head-raising pillow for respiratory comfort in the absence of isolated respiratory illness. The stages are as follows:

- *Stage 1:* No symptoms at any activity level but objective evidence of at least some reduction in ejection fraction.
- *Stage 2:* Some or all of the above mentioned symptoms with moderate exertion (more than one flight of stairs).
- *Stage 3:* The appearance of above mentioned symptoms with mild exertion (less than one flight of stairs).
- *Stage 4:* The persistence of symptoms at rest.

5.3.1 Minimal Monitoring (Level 1 Monitoring)

The stage 1 heart failure will require monthly monitoring of dyspnea episodes, blood pressure (BP), and weight. BP should be checked sitting up and relaxed after more than one hour of any exertion or caffeine intake, where exertion is defined as any activity raising cardiac output $\leq 15\%$ of baseline. Weight should be measured at the same time of the day (preferably at waking up) without clothes. Episodes of dyspnea is a subjective description. In other words, no objective evidence is required.

5.3.2 Closer Monitoring (Level 2 Monitoring)

The stage 2 monitoring will require daily salt and fluid intake, BP monitor, episodes of dyspnea on exertion with standard activity, changes in weight, and daily expenditure of metabolic equivalents.

Figure 5.1 shows different levels of monitoring and the conditions that change the monitoring level. Initially the monitoring is at Level 1. However, the monitoring will move to Level 2 when any one of the following conditions C1, C2, or C3 is true; the conditions are defined as follows:

Definition 5.1. C1: An increase in weight is greater than 2 lbs., where the increment cannot be explained by obvious factors.

$$C1 \equiv \Delta w > w_{threshold} \quad (5.1)$$

where Δw is change in weight and $w_{threshold} = 2$ lbs.

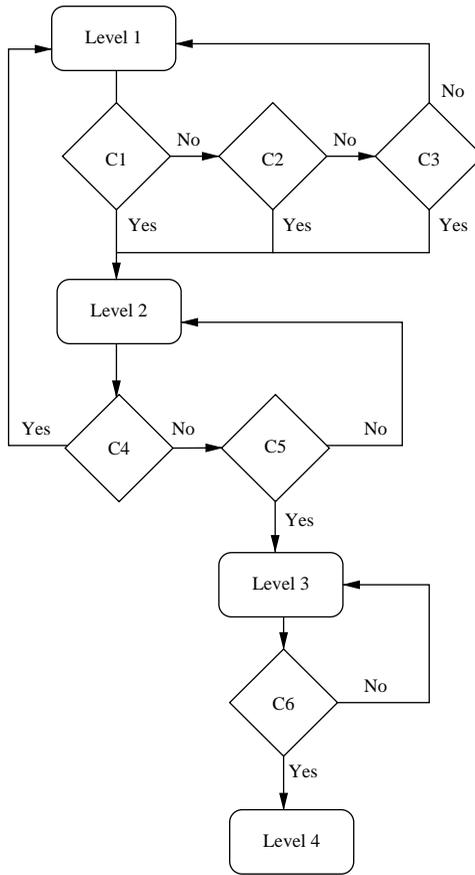


Figure 5.1 Stages of monitoring.

Definition 5.2. C2: A persistent fluctuation of more than 10 degrees systolic or 5 degrees diastolic blood pressure.

$$C2 \equiv (\Delta BP_{systolic} > n_{systolic}) \cup (\Delta BP_{diastolic} > n_{diastolic}) \quad (5.2)$$

where $n_{systolic} = 10$ and $n_{diastolic} = 5$.

Definition 5.3. C3: Episodes of dyspnea (shortness of breath).

Episodes should be a subjective description of shortness of breath, at rest, at night, or increasing need for a head-raising pillow for respiratory comfort in the absence of isolated respiratory status deterioration. Any episodes should move patient to Level 2 monitoring. The determination between Level 1 and Level 2 monitoring is based upon presence of any one of the three above-mentioned conditions. In other words, presence

of any one of the above three conditions shall move the patient from Level 1 to Level 2 monitoring. Further, in Level 1 monitoring the conditions are tested on a monthly basis whereas in Level 2 monitoring the conditions are tested on a daily basis.

After moving to Level 2 monitoring, patients can be moved back to Level 1, if weight, BP, and dyspnea stay stable for 3 months and no further clinical deterioration of cardiac failure (defined by the Framingham staging system) occurs, as shown in condition C4 of Figure 5.1.

Definition 5.4. C4: Stable weight, normal BP, and absence of dyspnea, for at least 3 months.

$$C4 \equiv \neg C1 \cap \neg C2 \cap \neg C3, \Delta t > t_{min} \quad (5.3)$$

where Δt is the duration of monitoring, and t_{min} is 3 months. In other words, if all three conditions (C1, C2, and C3) are false for more than 3 months, the monitoring will move to Level 1.

5.3.3 Closer Monitoring and Evaluation for Intervention (Level 3 Monitoring)

The stage 3 monitoring requires all of the above monitoring as well as monthly BNP (B-type natriuretic peptide) and cardiac impedance. BNP is a peptide released by cardiac muscle in direct response to strain.⁹ Higher values indicate worsening cardiac failure. Normal values are between 50 and 150. Values vary remarkably among individuals. An increasing number for one patient is a very good indicator of worsening heart failure and a good indicator to determine that worsening dyspnea is being caused by worsening cardiac status as opposed to respiratory status. On one hand, a particular patient at a BNP of 100 will clearly feel better symptomatically than at a BNP of 500. On the other hand, two different patients with a BNP of 300 may have very different symptoms from each other. It is because a rising BNP indicates worsening cardiac status with respect to a particular patient's baseline value as opposed to any absolute number. Over time, serial BNP values for a particular patient can be plotted to measure prognosis.

Definition 5.5. C5: Increase in BNP.

$$C5 \equiv \Delta BNP > n_{BNP} \quad (5.4)$$

For instance, for a patient A, if $BNP_1 = 400$ and $BNP_2 = 410$, $\Delta BNP = 10$, which does not indicate worsening cardiac status. However, for a patient B, if $BNP_1 = 100$

and $BNP_2 = 350$, $\Delta BNP = 250$, which is a significant change and indicates worsening cardiac status.

Cardiac impedance is a noninvasive method of calculating the following values: systolic pressure, diastolic pressure, mean arterial pressure, thoracic fluid content, vascular resistance, cardiac output, and ejection fraction. All the values directly indicate any worsening heart failure and vary from individual to individual. Any change from baseline to one particular patient can be used as a measure of worsening or improving cardiac status.

5.3.4 Aggressive Monitoring and Preparation for Intervention (Level 4 Monitoring)

Stage 4 monitoring requires all of the above with a monthly transthoracic echocardiogram, which gives a measure of the following:

1. Ejection fraction
2. Cardiac output
3. Valvular pressures
4. Transchamber flow gradients

Any change in values is a direct indicator of worsening or improving cardiac status.

Definition 5.6. C6: Worsening transthoracic echo cardiogram.

All of the above patients (stages 1 through 4) should have the ability to report any ischemic symptoms characterized by chest pain, sudden worsening dyspnea, etc. With such reporting, automatic cardiac enzyme levels (troponins) should be drawn with serial EKG and telemetry monitoring until ischemia or arrhythmias are ruled out.

Noncardiac causes that contribute to the appearance of decline in cardiac status should also be ruled out. These are acute insults and usually result in reversible decline in cardiac status, and if treated properly, should not require the patient to stay at a higher level of monitoring. These etiologies include but are not limited to pneumonia, septicemia, viral illnesses, anemia, worsening restrictive or obstructive lung disease, acute renal insufficiency, endocarditis, pericarditis, myocarditis, acute rheumatic fever, acute aortic dissection, and hypertensive urgency or emergency.

5.4 Smart Monitoring System

An architecture for a smart monitoring system is proposed;¹⁰ the architecture includes a sensor network, a server, a symptoms diary, users, and management. Figure 5.2 shows a sensor network that consists of a base station and the patient-related sensors such as oxygen, pulse, EKG, and weight. Using license-free radio frequencies, these sensors transmit sensor data streams to their corresponding base station. The base station collects the data from the sensors and transmits the aggregated data stream

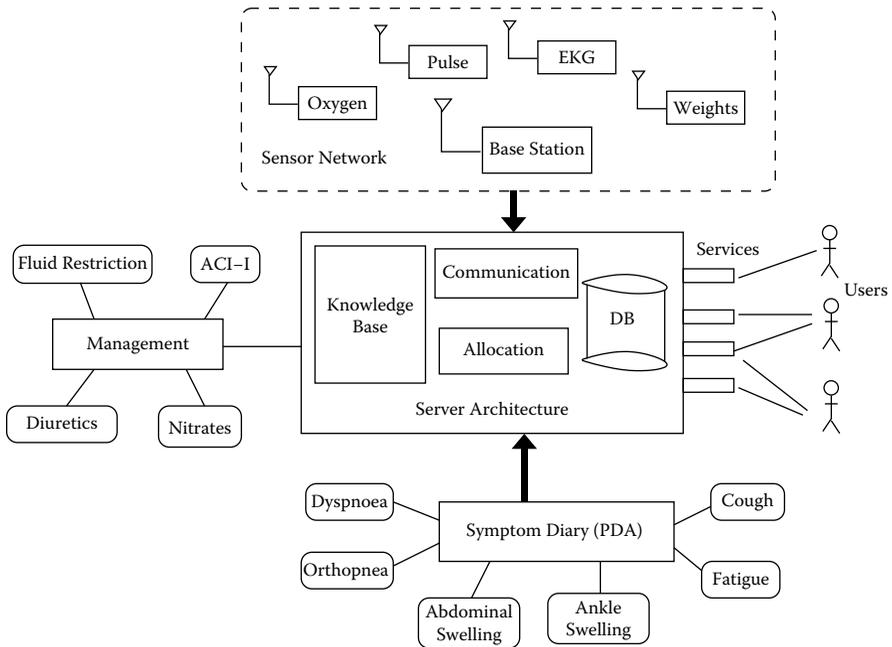


Figure 5.2 Architecture for a patient monitoring application.

to the server through the Internet. The base station is the bridge between the sensor network and the Internet.

The symptoms diary is managed through a personal digital assistant (PDA), which is also connected to the Internet. The patient fills in the Web-based form to record the entries for dyspnea, orthopnea, cough, fatigue, abdominal swelling, and ankle swelling. The entries are automatically uploaded to the server with the corresponding time stamp. If the patient forgets to record the entry, the server would generate a suitable alert for the missing or delayed entries.

The management is prescribed by the physician and the corresponding management entries are used to control the monitoring application. For CHF patients, the management entries include fluid restriction, ACI-I, diuretics, and nitrates. The management is stored as a database table and each management modification is identified by the corresponding time stamp.

The server consists of the following components:

- A communication component receives the data streams from the base station and PDA. The communication component hides the underlying communication technology; the sensor data could be received by telephone modem, local area network connection, or a serial cable.
- An allocation component stores the allocation schema, metadata, the types of sensors, symptoms, and management strategies.

- A database (DB) component stores the data collected for the patient. It stores the data received from the base station as well as the data received from the symptoms diary. The allocation component specifies “what” is being monitored. On the other hand, the actual monitored data is recorded in the database. The DB component stores the data in a relational database management system such as IBM DB2, PostgreSQL, or Apache Derby. The relational database assists in data processing for customized data visualization.
- A knowledge component is used to extract the nontrivial information from the data streams, metadata, and the usage of services. The data mining principles such as frequent item sets can be used to extract the useful information. The knowledge base will recognize the alerts in the progress of the CHF patients on regular intervals, and it will prompt the services so that patients can have timely access to health care. The expert system will be developed using patients’ histories and related case studies. The intelligent techniques such as neural networks and genetic algorithms can be used for filtering the patients with immediate needs.
- Finally, services are provided to facilitate the management and data visualization of the sensor data. The services can be used to create customized applications for the users. For instance, the services could use Simple Object Access Protocol (SOAP) to communicate using XML messages. Users interact with one or more services to achieve their objectives. Users can be software agents, programs, or patient care staff.

5.4.1 Monitoring Human Behavior in a Bedroom

An experiment is conducted to monitor human mobility and activity pattern. In this experiment, sensors are deployed in a bedroom to monitor the behavior, mobility, or lifestyle of a person. For instance, the sensors can be used to determine the time spent at the study table or on a bed.

Figure 5.3 shows the room plan of a bedroom. Moteiv’s Tmote Sky sensors¹¹ are used to monitor the temperature, humidity, and light of the room. The sensors are

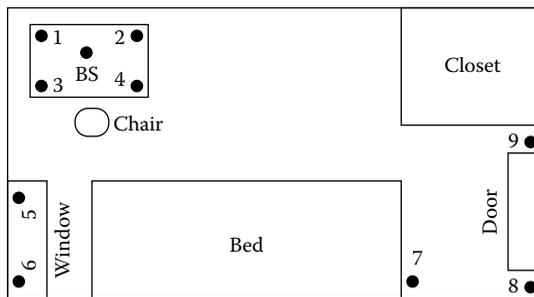


Figure 5.3 Room plan of a bedroom.

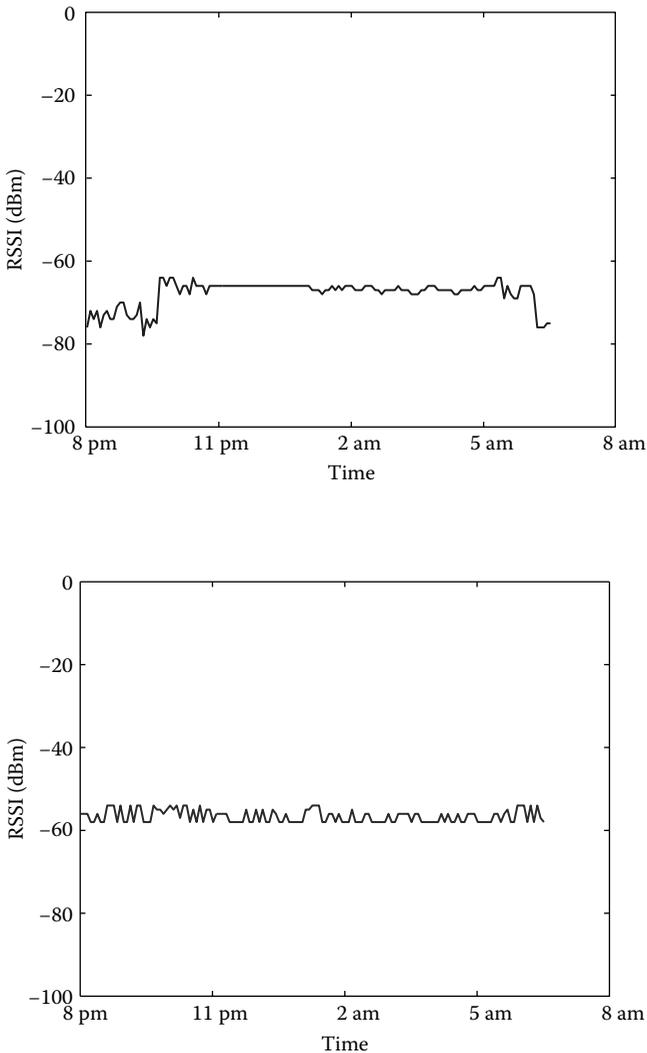


Figure 5.4 RSSI variation for sensors in a bedroom.

deployed as follows: four sensors on a study table (Mote 1, Mote 2, Mote 3, and Mote 4), two sensors near the window (Mote 5 and Mote 6), one sensor on the corner of the bed (Mote 7), and two sensors near the door (Mote 8 and Mote 9). The base station is attached to a laptop on the study table. The sensor data is collected for several days.

Figure 5.4 shows received signal strength indicator (RSSI) values for table, window, and door sensors. RSSI variation can be used to identify the activity of a person. A student was working at the study table for the initial 3 hours, which is confirmed by RSSI variation for Table motes 1 and 3, as shown in Figure 5.4(a)

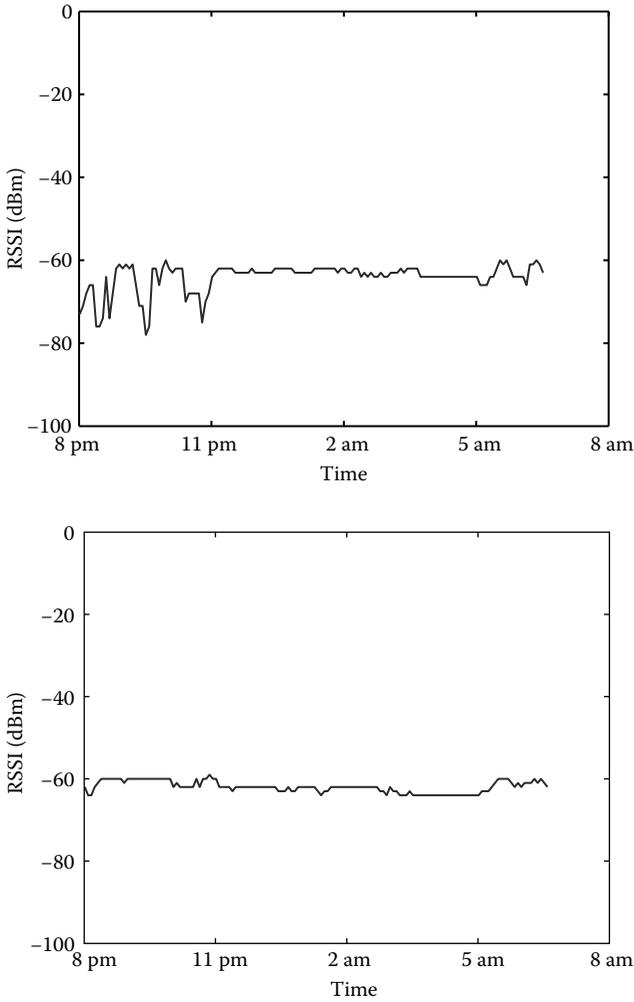


Figure 5.4 (Continued)

and Figure 5.4(c), respectively. Further, the small variation for Table motes 2 and 4 [Figure 5.4(b) and Figure 5.4(d)] indicates that the person’s sitting position did not affect these values. By comparing the RSSI values of Table motes 1, 2, 3, and 4, the sitting posture or inclination on a table can also be determined.

After 3 hours of working, the person slept for 7 to 8 hours. The sleeping time is also evident from RSSI values of window and door motes. Further, the RSSI variations for door motes are not consistent. For instance, Mote 8 [Figure 5.5(c)] near the bed confirms the sleep behavior; however, the mote near the closet (Mote 9) shows

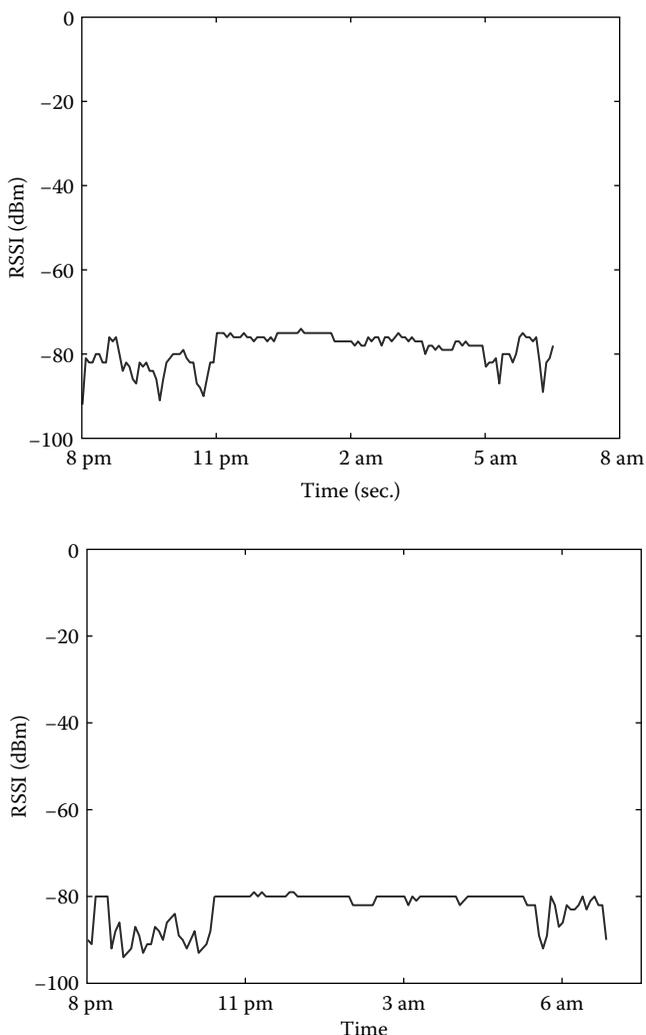


Figure 5.5 RSSI variation for sensors in a bedroom.

some unexpected RSSI variation. Finally, the RSSI variation in the last couple of hours shows the morning activity, which is confirmed by most of the notes.

The above results can be used in maintaining a daily activity diary for a person. The results obtained at the base station can be logged in a database and can be retrieved through a Web-based application. Consequently, a lifestyle of a person, active or sedentary, can be estimated by these results. It should be noted that sensors will automatically record the activity and duration of the activity. Further, sensors can be used to determine mobility during sleep, which could indicate stress level.

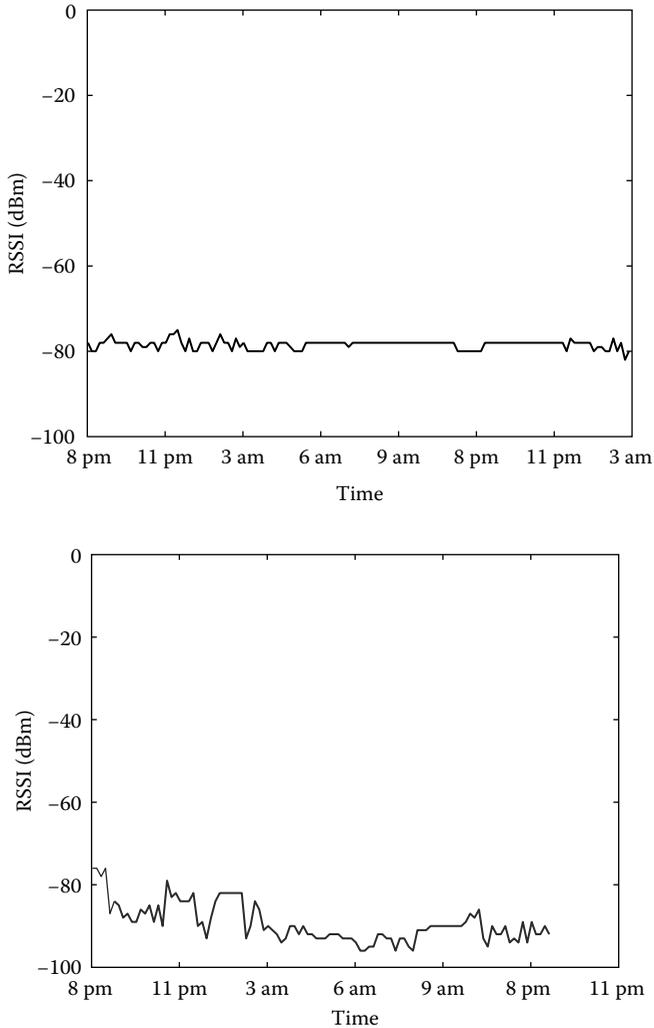


Figure 5.5 (Continued)

Figure 5.6 shows the humidity variation for the same experiment. The humidity variation confirms the presence of a person. Initially, the humidity level was low; however, due to presence of a person in a small room, the humidity level reaches a higher value. Further, in the morning when the door was opened, there is a sharp decrease in the humidity level, as shown around 5 a.m. in the graphs of Figure 5.6. These rapid changes in humidity level can be used to determine the opening and closing of the door. Further, the humidity profile can be used to determine the comfort level of the room environment.

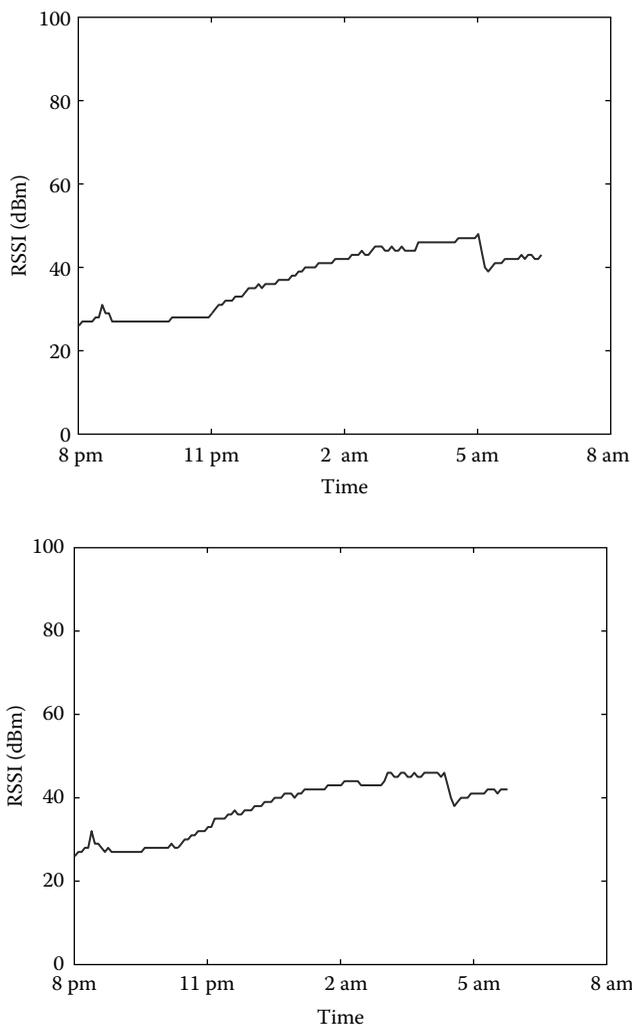


Figure 5.6 Humidity variation.

5.4.2 Monitoring Example

Several symptoms and signs can be used to give a clue on the health status of a typical patient with congestive heart failure. These include patient's weight, heart rate, oxygen saturation (SpO_2), EKG, and a symptom scale. Following is an example of these symptoms:

Ms. Jane Smith is a 65-year-old patient who was diagnosed with congestive heart failure three years ago. Her symptoms are being managed

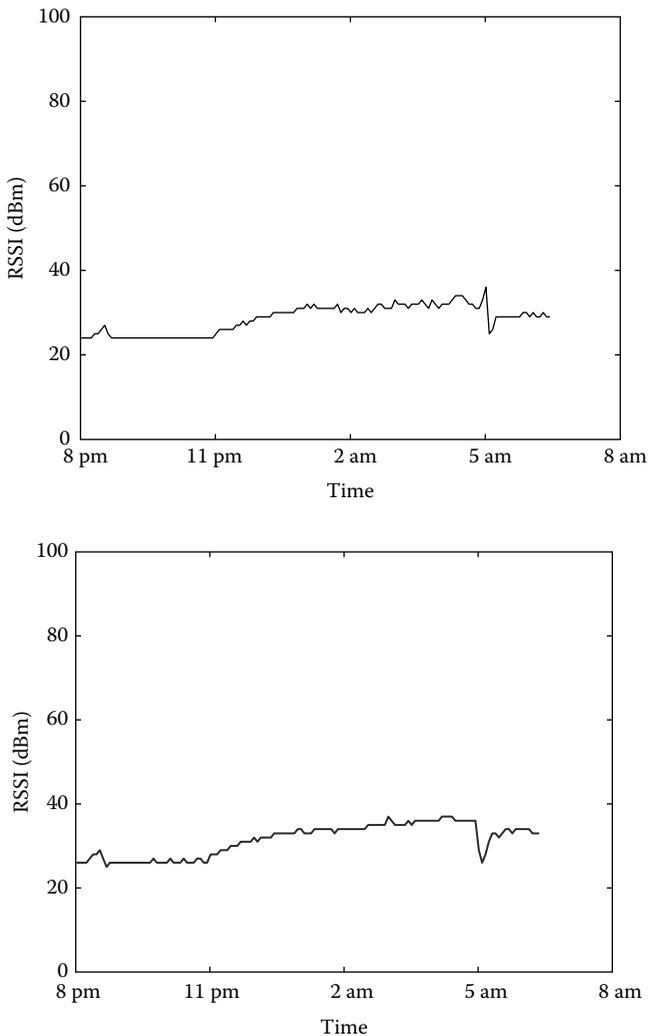


Figure 5.6 (Continued)

with fluid restriction, diuretics, ACE-I, and nitrates. She visits her family physician once a month for follow-ups. She needs to arrange the handi-transit service as well as request her neighbor to accompany her to the doctor appointments. She has been hospitalized three times in the last 18 months due to a aggravation of her symptoms. Ms. Smith lives alone in an apartment 35 miles away from her family doctor's clinic. Ms. Smith's medical condition is slowly declining, and thus it has become difficult for her to manage her appointments.

The smart sensor system will provide Ms. Smith with useful wireless devices that can check her oxygen saturation, pulse rate, and EKG at regular intervals. Ms. Smith will be checking her weight daily and the device will record and forward information about her body weight directly to the system. Ms. Smith will be given a symptom diary to record the related symptoms, including dyspnea, orthopnea, cough, abdominal or ankle swelling, and fatigue. Ms. Smith would rate her symptoms on a scale of one to ten. This symptom scale can be forwarded to the system through a Web browser or a telephone system.

This way, Ms. Smith's health care providers are being updated with the information on a regular basis. The monitoring system is also able to generate alerts for the care providers to intervene. For instance, change in BP, weight, or dyspnea (C1, C2, or C3) could generate an alert to change the monitoring level. Further, based on sampled data and patient's condition, the system can adapt the sampling frequency to avoid unnecessary data transmissions.

An expert system can be used to assist the sensors to identify the emergency situation using patient's vital signs. In case of emergency, the smart sensors will alert the corresponding authority or personnel. Moreover, the continuous and seamless recording of sensor data can be used for future analysis. The physicians or nursing staff would be able to monitor the necessary data for the outpatients using a Web browser or a PDA. The information about the location and the level of the emergency situation will be automatically forwarded to the first responders, if needed.

Similar systems can be developed to assist in management of patients with other chronic conditions like COPD and ischemic heart disease. These systems are able to provide accurate information while using minimal resources.

5.5 Conclusion

This chapter shows a smart monitoring system for CHF patients. After a brief introduction of sensor networks, congestive heart failure is briefly described. Then, some of the patient management and monitoring techniques are investigated. Finally, an architecture and implementation of a smart monitoring system is discussed. The proposed system can assist health care providers in giving high-quality patient care and can reduce the overall health care costs.

References

1. D. Estrin, D. Culler, K. Pister, and G. Sukhatme, Connecting the physical world with pervasive networks, *IEEE Pervasive Computing*, January-March, 2002, 59–69.

2. D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, Next century challenges: Scalable coordination in sensor networks, in *Proceedings of the International Conference on Mobile Computing and Networks (MobiCom)*, 1999.
3. T. Gao, D. Greenspan, M. Welsh, R.R. Juang, and A. Alm, Vital signs monitoring and patient tracking over a wireless network, in *Proceedings of the 27th IEEE EMBS Annual International Conference*, 2005.
4. S.R. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong, TinyDB: An acquisitional query processing system for sensor networks, *ACM Transactions on Database Systems (TODS)*, 30(1): 122–173, 2005, <http://doi.acm.org/10.1145/1061318.1061322>
5. K. Lorincz, D. Malan, T. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, S. Moulton, and M. Welsh, Sensor networks for emergency response: Challenges and opportunities, in *IEEE Pervasive Computing, Special Issue on Pervasive Computing for First Response*, October, 2004, 16–23.
6. C. Toman, M.B. Harrison, and J. Logan, Clinical practice guidelines: Necessary but not sufficient for evidence-based patient education and counseling, *Patient Education and Counseling*, 42(3): 279–287, 2001.
7. M.B. Harrison, G.B. Browne, J. Roberts, P. Tugwell, A. Gafni, and I.D. Graham, Quality of life of individuals with heart failure: A randomized trial of the effectiveness of two models of hospital-to-home transition, *Medical Care*, 40(4): 271–282, 2002.
8. <http://www.uptodate.com>
9. R.E. Hobbs, Using BNP to diagnose, manage, and treat heart failure, *Cleveland Clinic Journal of Medicine*, 70(4): 333–336, 2003.
10. S. Hussain and S.M. Dar, Architecture for smart sensors system for tele-health, in *First IEEE International Workshop on Health Pervasive Systems (HPS'06)*, Lyon, France, 2006.
11. <http://www.moteiv.com>

Chapter 6

Issues in Personal Cardiac Health Monitoring with Sensor Networks

Kathy J. Liszka, Malinda J. Sever, Michael E. Richter, and Sudha Bhattarai

CONTENTS

6.1 I ntroduction.....	104
6.2 T elemedicine for Cardiac Health	105
6.3 F irst Generation Prototype	107
6.4 S econd Generation Prototype	109
6.4.1 P reliminary Results of First Phase Data Collection	111
6.4.2 I nitial Thoughts on Working with Motes	112
6.5 C onclusions and Future Work	113
Acknowledgments	114
References	114

We view telemedicine, particularly for cardiac health, as a proactive model versus the traditional reactive one. Around-the-clock monitoring of vital signs helps identify problems before they become serious or life-threatening emergencies. Our goal

is to define and provide a model for personalized health care for chronic health conditions using patient diagnosis to drive the requirements of the telemedicine solution. We discuss two generations of prototypes using different hardware approaches to determining practicality and robustness. We explore a design using motes for fine-grained sensitive medical data. We find that they are not robust enough without major modifications and work on the communications protocol. However, they may work very well in an overall telemedicine architecture for cardiac monitoring where coarse grain data is to be transmitted on a less frequent basis than ECG signals.

6.1 Introduction

We are an aging society. The U.S. Bureau of Census estimates there are 35 million people over the age of 65 in the United States, or roughly 15% of the population.¹ Baby boomers account for an overwhelming total of 28%. Couple these numbers with advances in medical technology and the simple result is that people are living longer. However, longevity comes at a cost. We want those extra years to enjoy retirement, family, and friends but a certain quality of life must be present. Spending an extra few years bed-ridden, homebound, or confined to a nursing home is not the extended lifestyle people hope for. Quality of life is a very important issue.

Far from enjoying the fountain of youth, the elderly still suffer from many ailments. In the next 35 years, it is projected that the number of people with chronic conditions will increase by an alarming 50%. A chronic illness is defined as a disease or condition that persists over a long period of time, or one that reoccurs frequently. Common examples include cardiovascular disease, chronic heart failure, diabetes, and asthma. While the focus of this chapter is on cardiac health, we remain open to telemedicine in many areas. It is common for the elderly to have multiple illnesses. We also note that chronic conditions are not age discriminating. Indeed, over 45% of Americans have one or more chronic illnesses at some time in their lives.²

The cost of treating these diseases is an astounding \$1 trillion, consuming a total of 75% of health care spending.³ The financial burden for cardiovascular disease alone was estimated to be almost \$370 billion in 2004. By comparison, the estimated cost for cancer treatment, the second leading cause of death in the United States, was approximately \$190 billion in the same year.⁴ The balance to this equation is insurance carriers decrease reimbursements and increase deductibles as they struggle to deal with these numbers.

Telemedicine is the use of computers to monitor patients, diagnose conditions, and report data to clinicians through various means of telecommunications media. Medical telemetry systems are evolving rapidly as communication technology advances, evidenced by the commercial products and research prototypes for remote health monitoring that have appeared on the market. It is one of the fastest growing sectors in the medical industry.⁵ More recently, wireless systems allow

patients to move freely in their home and work environments while being monitored remotely by health care professionals. Patients are being offered more options for independence. Senior citizens can avoid the cost of eldercare. From a practical standpoint, disease management via telemedicine reduces emergency room visits and hospital admissions while increasing quality of life. In the case of chronic heart failure, studies have concluded that almost 75% of emergency room visits and resulting hospital admissions can be avoided, indicating that this approach to disease management is not only appropriate but also very effective.³

6.2 Telemedicine for Cardiac Health

Our primary focus is personalized cardiac monitoring. Cardiovascular disease (CVD) robs us of precious years of life. One in every three adults in the United States is estimated to have some form of CVD. Roughly 2500 people die of CVD every day, or an average of one death every 35 seconds.^{4,6} This sobering statistic includes the following cardiac events:

- Myocardial infarctions, more commonly known as heart attacks
- Congestive heart failure (CHF)
- Cardiac arrhythmia, disorders of the regular beating of the heart:
 - Tachycardia, a heart rate of greater than 100 beats per minute
 - Bradycardia, a heart rate of less than 60 beats per minute
 - Atrial fibrillation, an irregular heartbeat

Electrocardiograms (ECGs) are a well-established and widely accepted method for monitoring the electrical activity of the heart. Figure 6.1 illustrates a theoretical ECG signal. Numerous ECG monitoring devices have been developed and marketed for the sports industry. Athletes' needs have been targeted so that they can be monitored under conditions that are physically stressful. Ambulatory elder patients have significantly different needs. A wider variety of symptoms should be monitored in older adults, particularly as their body chemistry changes. For example, studies suggest that as we age, we are at a higher risk for complications directly related to drug interactions, even if there were no problems in the past.⁷ Among suggestions for management of these risks are regular ECGs tests.

Today, several devices are commercially available for cardiac monitoring. Numerous devices are also available for purchase over the counter (OTC) but their use extends only to personal at-home use and most likely does not involve automatic physician intervention. Commonly used OTC at-home devices for these purposes come in the form of blood pressure monitors, weight scales, pulse oximeters, and three-lead ECG Halter. A remote patient typically wears an ECG Halter that collects data through wires attached to skin-contact biosensors. A home-based ECG Halter uses only three leads. A typical resting diagnostic ECG displays 12 leads

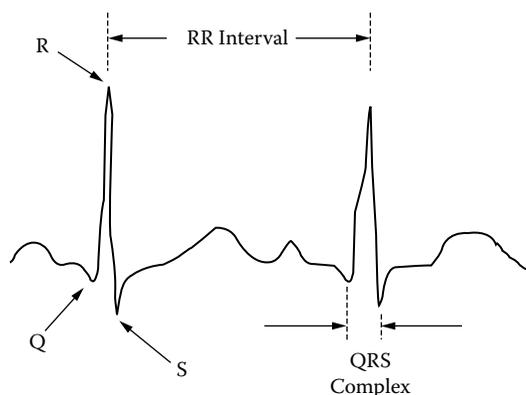


Figure 6.1 QRS Complex. An electrocardiogram represents the electrical activity of the heart muscle as it is recorded from surface sensors placed in standard locations on the body. Current passing toward and away from the positive end of a bipolar electrode causes a large deflection in the waveform of an ECG. Electrical current flowing at an oblique angle to the electrode causes a smaller deflection while current flowing at a perpendicular angle produces a biphasic deflection in the recorder. Each lead “sees” the heart on a different plane. All of this information is plotted as a series of deflections and waves that may be graphically represented, each containing unique information as well as redundant information about the rhythm of the heart. The RR interval is a measurement of the time between heartbeats.

using 10 biosensors. A three-lead configuration supplies the rudimentary beats per minute and QRS interval necessary for a quick assessment of the arrhythmia disorders being monitored. Some of these are capable of sounding a warning if irregular heartbeats are detected.^{8,9} However, current ECG sensor collection methods are not viable for more than several days. Skin contact electrodes and cables are an impediment to long-term monitoring, particularly in the elderly where adhesives used for ECG electrodes are potentially damaging to the skin when worn over long periods of time.

Several cutting-edge technologies are on the horizon. One such device is a patch which, when applied to the skin, continuously monitors a person’s blood pressure, heart rate, and other vital signs, and then forwards that information to their participating physician’s computer via wireless links.¹⁰ Several companies are working on this technology, but are still undergoing testing prior to FDA approval. We predict, based on this and other research being done in the field of biosensors, that a wearable cardiac monitoring shirt or vest device will replace the traditional Halter monitor. Rudimentary biosensor shirts exist mainly in research environments but they approach the problem with a very small number of sensors that are wired together for communication.^{11–13} The sensors are placed fixed in position on a stiff vest with little range of movement that a normal person would be comfortable wearing for

any reasonable length of time. Biosensors are on the inside of the vest and attached directly to the patient restricting a normal range of body movement.

A number of wireless noncontact ECG electrodes are under test and development. They vary in size, capability, and availability but this new technology has been proven and is advancing. The value of this type of electrode is the ability to capture the electrical signal generated by the heart without literally being tethered to an ECG Halter. One particularly large hurdle, especially for the elderly with fragile skin, is the electrode itself. ECG data collection is normally a noninvasive technique, although skin spike and ingestible sensors exist. Skin contact electrodes are applied directly to the skin with a conductive adhesive gel to maintain contact for high-quality traces. These irritate the skin and are not meant for long-term use or extensive movement. Commercially available devices for home monitoring typically use three-lead, hardwired, skin contact ECG sensors that can be worn up to 14 days. Noncontact sensors allow total freedom of patient movement while being monitored. This type of biosensor is not commercially available but researchers in both the United States and Great Britain have proved the science. Patents are currently pending for noncontact electrodes.¹⁴

6.3 First Generation Prototype

The first prototype, the arrhythmia monitoring system (AMS), is a completed, working wireless telemetry system testbed developed at NASA and Case Western Reserve University's Heart & Vascular Center. The AMS system collects real-time electrocardiogram (ECG) signals from a mobile or homebound patient, combines global positioning system (GPS) location data, and transmits it to a remote station for display and monitoring. It was developed as a collection of parts with a commercial ECG Halter, an 8051 development board, Palm Tungsten, Emtac GPS unit, all using Bluetooth, with GPRS (General Packet Radio Service) using AT&T, a 2.5G long-distance wireless service.¹⁵

The end-to-end system architecture consists of a wearable server system, a central server system, and a call center service. The wearable server shown in Figure 6.2 is a small data collection and communication device. ECG signals are collected from a connecting three-lead Halter device worn by the patient, and transmitted over a short-range wireless link to a central server sitting in close proximity. A three-lead configuration supplies the rudimentary beats per minute and QRS interval necessary for a quick assessment of the arrhythmia disorders being monitored.¹⁶

Collected samples are transmitted over a wireless link to the Central Server, which is in close proximity to the person. The digitized ECG data, headers, start and stop bits fill a minimum message size of 9 bytes. The data acquisition rate of 4 milliseconds requires a minimum baud rate of 22.5 Kbps. This sets a bound on the communication requirements for the short-range wireless components. The most widely available and supported COTS components are Bluetooth™ and 802.11b (also known as WiFi).

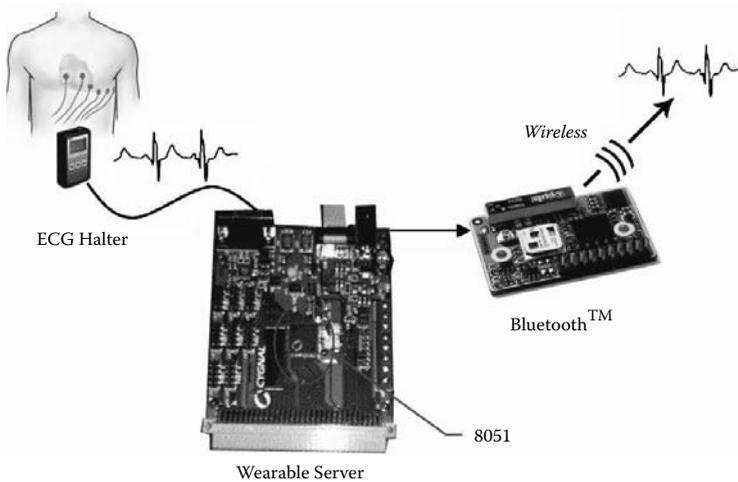


Figure 6.2 The wearable server. This portable device worn on the patient receives analog signals from the ECG sensors, digitizes the signals, and transmits the data over a wireless link to the central server. The prototype system uses an 8051 micro controller on a development board connected to three standard ECG leads. Data is collected at a sampling rate of 250 Hz. One sample contains three digitized ECG readings with 13 bits of resolution. The dynamic range of the data is -9.99 to $+9.99$ mV.

Bluetooth was selected for our first prototype because it is low cost, low power, and robust. The internal antenna transmits in the license-free Industrial, Scientific and Medical (ISM) frequency band ranging from 2.4 to 2.4835 GHz. Class 3 radios can reliably receive signals from other Bluetooth devices within a 10 meter range which more than suffices for our body range ECG monitoring prototype. Frequency hopping reduces signal fading and interference from other nearby devices transmitting in the ISM band. WiFi devices operating in the same 2.4 GHz ISM frequency band would serve as well in the remote arrhythmia monitoring system but at greater cost.

The central server shown in Figure 6.3 performs several functions including data compression, location awareness via GPS signals, and rudimentary arrhythmia detection. The messaging protocol used in the system is influenced by several factors. Device capabilities and wireless service provider drive session management. Our PDA connects to a private network on a long-distance carrier. A private Internet Protocol (IP) address is dynamically assigned for a single communication session. Once connected, data may be sent and received on demand in what is referred to as an “always-on” IP-based service. The central server initiates and establishes a connection and immediately begins streaming data. The interface specification uses a bidirectional, stateless, serial data transmission protocol. Error checking, correction, and retransmission of data is handled by the TCP/IP protocol. This is essential for delivery of a high-integrity ECG signal. The central server buffers data until acknowledged.



Figure 6.3 The central server. The central server is the logical midpoint between a patient and the call center. The prototype component is a Palm Tungsten W personal digital assistant (PDA). ECG data, patient notifications, and optionally GPS location coordinates are multiplexed and continuously transmitted over a single long distance wireless link to the call center via a built-in cellular modem.

The call center is designed to be a facility staffed 24 hours per day, 7 days per week by qualified health care professionals. This staff has the capability of remotely monitoring a set of patients within a given geographical area determined by telephone and Internet infrastructure of the area and by the number of patients within the area. A high-performance, commercially available personal computer collects and displays the three-lead ECG signal in near real-time (~10 to 30 sec latency) using traditional strip chart graphics beside a map showing the most recently acquired patient location. A user-friendly GUI lets the monitoring technicians quickly access patient history and physician contact information.

6.4 Second Generation Prototype

In the second prototype, we are working on a model of mobile sensing that is driven by a patient's personal medical diagnosis to determine the exact architecture and behavior of the system. We have the beginning of a framework for a body area network (BAN) in the form of a biosensor shirt with multiple numbers and types of biosensors.¹⁷ We are designing for next-generation wireless, noncontact ECG sensors with blood pressure and body temperature sensors for diagnostic support. Similar to the AMS system, we use technology with short-range wireless



Figure 6.4 Crossbow MIB 510 base station and Mica2Dot motes. Each mote has an Atmel Atmega128L micro controller and frequency tunable radio with 512 bytes of flash memory.

communication capability but we are using smaller, less developed components for experimentation. For this prototype, we are constructing a system using a Crossbow MIB 510 base station connected to a laptop, a Mica2 mote and up to five Mica2Dot motes, as shown in Figure 6.4.¹⁸ The motes create the communications network and serve limited local processing needs. The radios on both the Mica2 and Mica2Dot operate in the 433 MHz band. Radio transmission power is adjustable.

Introduction of the Mica2 and Mica2Dot motes brings the prototype to a lower level than the commercial products used for the AMS system. Our first challenge is due to the nature of wireless communications. The asymmetry of links makes it difficult to estimate link quality and invalidates many assumptions made in other

environments. Losses due to obstacles and interference compound, decreasing the overall effectiveness of the system. The second challenge we face is the constrained power resource of the motes. It also has very limited computational power and memory space. We must carefully analyze the complexity of the algorithms we select. Finally, the communications bandwidth is narrow.

In medical applications, inaccurate data is simply not an option. Our first goal in this research is to measure robustness of the Mica2Dot motes to determine if they are a reasonable component to work with. Robustness is defined as the degree to which a system can function correctly in the presence of inputs different from those assumed.¹⁹ Here, we are measuring perturbations in communications between motes based on varying parameters of power, samples per packet, and packets per second.

In both the AMS and the mote prototypes, we use the R-peak detection algorithm for detecting QRS complexes.²⁰ With this basic calculation, we monitor for a minimum of the following conditions:

- *Bradycardia*: Heart rate < 60 beats per sec
- *Tachycardia*: Heart rate > 100 beats per sec
- *Sinus arrest*: No atrial electrical activity ≥ 3 sec
- *Ventricular tachycardia with broad QRS complexes*: QRS interval > 120 milliseconds and heart rate > 100 beats per sec²¹
- *Supraventricular tachycardia with narrow QRS complexes*: QRS interval < 120 msec and heart rate > 100 beats per sec

6.4.1 Preliminary Results of First Phase Data Collection

In initial tests, we are using TinyOS 2.0 on Cygwin.²² TinyOS is a small, open source operating system developed by UC Berkeley written to support research in sensor networks. For the interface, we modified a TinyOS application called Oscilloscope and developed test communication programs for the motes. Experiments consisted of a set of up to five Mica2Dots placed at distances from one to five feet from each other, all communicating with the Mica2 mote. Test data created and stored in flash memory were transmitted, checked, and recorded. Identifiers and sequence numbers helped us identify successful communication, dropped packets (or acknowledgments), and corrupted packets. Variables we worked with were 1, 5, and 20 packets/sec (Hz), all at 60 samples per packet. We work with 60 samples per packet based on the block sizes of the ECG data we store on flash for experimentation. Preliminary testing has been done with archived cardiac data available through MIT.^{23,24} This is an open source database that provides a large collection of physiological signals representing a cross-section of many medical conditions. The signals used for testing our algorithms are loaded in nonvolatile storage on the motes.

For the first phase of data collection, we set the power level as low as possible to avoid signal interference among the motes. In the next phase, we plan to study the operational characteristics of different output power levels versus error rates because the sensors are in close proximity to each other. Testing was performed for continuous transmission for 25 min each test. Consolidated results showed an average dropped (not acknowledged) packet ratio from 2 to 5%. We calculated the average of five repetitive test suites on the same sample data for sets of 1 to 5 motes running at 20 Hz. Interestingly, the number of corrupted packets was always less than 1% in all of our runs. Visual inspection of the ECG data signal indicated that the numbers being reported by the MIB base station were not accurate. We selected a normally busy area in a university building with an average amount of traffic, both mobile and sedentary. The building has an active 802.11 wireless network running at 2.4 GHz, which did not interfere with our frequency range of 433 MHz. This is only a sample of the results from the data collected but is representative of the test suite. At one point, we had outlying data for a set of three motes where the number of packets sent dropped dramatically. We reran the tests several times to isolate a malfunctioning mote.

6.4.2 Initial Thoughts on Working with Motes

In our first attempt at testing communication at different sampling rates, we observed some anomalous behavior. With closer inspection, we discovered we were not checking the cyclic redundancy check (CRC) on the command request packet sent out from the MIB 510 base station. Optimally, we need a sign-on protocol to have available motes register with the base station, and then the base station sends out a signal for the motes to synchronize their sampling. Now, if any registered motes do not hear the synchronization signal, the synchronization must be done again. We also made a change to the Mica2Dot to Mica2 communication architecture to make it more error tolerant. Packets received with a bad CRC are now automatically dropped. The sender does not receive an ACK, times out, and resends. While this seems like a normal property of communication, one must remember that we are not using TCP. We discovered two properties of the active message (AM) radio system that were not initially clear. First, a CRC check was being performed on the data but was ignored and valid packets were being ACK'ed. However, our own software check was showing erroneous data both programmatically and visually (the signal was malformed where we knew it should not be knowing the exact data stored in flash). This finally explained why the corrupted packet rate was at less than 1%. The second issue we had with the radios is that the system sends packets based on a 15- to 20-Hz timer. This is most likely caused by the receiver not receiving the packet, or receiving a bad packet and the sender wait times out.

We developed a separate program from the ECG data samples to shed some light on the rate of bit errors. Called our “integrity” program, early results showed

that many single bit errors were occurring in a packet. Given the CRC check is being performed, clearly there must be errors in one or more other bytes as well. The integrity program records and retrieves errors with more information than simply that an error occurred. We had some difficulty communicating over the radio with ActiveMessage. It appears that there is a packet size limit that is not enforced by the compiler or TinyOS. The send command succeeds and the send-Done command returns success, but the message is simply not picked up on the other end. On closer inspection, it appears that the maximum packet payload is 28 bytes. More testing indicates a fairly low limit on the frequency at which packets can actually be sent. Sending packets at a frequency of 18 Hz seems too fast for the motes to sustain without trying to send a second packet while still sending the first. A frequency of 15 Hz appears to be better, but sending issues are still encountered even as low as 10 Hz. We are working on a more sophisticated sending mechanism with the goal of ensuring communication of the data occurs at the full sampling rate.

Based on these observations and the results of the integrity test suite, the second phase of testing is based on the maximum size of a single array in flash. The program will store an initial amplitude and an array of signed 8-bit differences set up to loop continuously. The samples from the MIT-BIH database were recorded at 360 Hz. The sample packets will be reported at 20 Hz with 18 samples in each. We have the loop set to repeat every 90 sec. Because 20 Hz appear to be more frequent than the motes can handle nicely, we are using 20 samples/packet and 18 packets/sec. Our test suite is currently an ongoing process.

6.5 Conclusions and Future Work

We are studying the intricacies of programming the mote radios and finding it to be less straightforward than we originally anticipated. The ECG collection part of the overall system will most likely remain with Bluetooth devices. The AMS prototype has been able to handle that task exceedingly well, and the upgrade from 2.5G to 3G cellular communications from the central server to the call center is an improvement. Mote communication still plays a role in the overall telemedicine architecture for cardiac monitoring. While probably not robust and reliable enough for ECG signals, there is a wide range of vital signs measurements that can be collected and collaborated that is much less fine grain and critical. Bandwidth ceases to be an issue in reporting occasional information for blood pressure, body temperature, blood glucose, O₂ saturation, or weight. For example, weight gain is of extreme importance in CHF patients. Changes in weight caught for patients with CHF can eliminate emergency room visits and increase chances to save a patient's life. Motes may serve better as small, inconspicuous communication nodes in a wireless home network for telemedicine.

Acknowledgments

This work was supported by a 2006 University of Akron Faculty Research Grant, the John Glenn BioEngineering Consortium at the NASA Glenn Research Center, and the Cleveland MetroHealth System.

References

1. C. Lehmann and J. M. Giacini, Pilot study: The impact of technology on home bound congestive heart failure patients, *Home Health Care Technology Report*, 1(4): 50, 59–60, 2004.
2. C. Ashman, What home care agencies should know about telehealth, *Tools for the Trade*, VI, 4, Nov.–Dec. 2004.
3. K. Utterback, Supporting a new model of care with telehealth technology, *Telehealth Practice Report*, 9(6): 3, 11, 2005.
4. Prevention Works: CDC Strategies for a Healthy-Heart and Stroke-Free America, published by Centers for Disease Control and Prevention, Division for Heart Disease and Stroke Prevention, Atlanta, GA, http://www.cdc.gov/DHDSP/library/prevention_works/pdfs/Prevention_works.pdf
5. Telcomed, <http://www.telcomed.ie/technology.html>
6. Heart Disease and Stroke Statistics 2006 Update, American Heart Association, <http://www.heart.org/downloadable/heart/1140534985281Statsupdate06book.pdf>
7. K.E. Brown, Top Ten Drug Interactions Most Dangerous to Seniors in Long-Term Care, <http://www.seniorjournal.com/NEWS/Eldercare/4-12-14TenDrugs.htm>
8. A & D Company, Ltd., <http://www.aandd.jp/products/medical/personal.html>
9. Microlife USA, http://www.microlifeusa.com/prodinfo_bp.asp#studies
10. M. Kanellos, The next thing on the Net: Your cardio system, http://news.com.com/The+next+thing+on+the+Net+Your+cardio+system/2100-11395_3-5865625.html
11. R. Wijesiriwardana, K. Mitcham, and T. Dias, Fibre-Meshed Transducers Based Real Time Wearable Physiological Information Monitoring System, ISWC, Eighth IEEE International Symposium on Wearable Computers (ISWC'04), 2004, 40–47.
12. N. Halín, M. Junnila, P. Loula, and P. Aarnio, Towards the Future OR: LifeShirt™; Wearable Patient Monitoring and AIDA with Network Connection, Internet & Multimedia Systems and Applications, Honolulu, Hawaii, August 2005, 304–309.
13. R. Jafari, F. Dabiri, P. Brisk, and M. Sarrafzadeh, Adaptive and Fault Tolerant Medical Vest for Life-Critical Medical Monitoring, 2005 ACM Symposium on Applied Computing, 272–279.
14. Quasar: <http://www.quasarusa.com/usa/tp.html>
15. K.J. Liszka, M.A. Mackin, M.J. Mackin, D.W. York, D. Pillai, and D.S. Rosenbaum, Keeping a beat on the heart, *IEEE Pervasive Computing, Mobile and Ubiquitous Systems*, 3, 4, 2004, 42–49.
16. D. Durbin, *Rapid Interpretation of EKGs*, Sixth Edition, Cover Publishing Company, Tampa, FL, 2000.
17. K.J. Liszka, A Sensor Network Architecture for Cardiac Monitoring, 4th Consumer Communications and Networking Conference, 737–740, Jan. 2007.

18. Crossbow Technology, Inc. <http://www.xbow.com/>
19. S. Ali, A.A. Maciejewski, H.J. Siegel, and J.K. Kim, Measuring the robustness of a resource allocation, *IEEE Transactions on Parallel Distributed Systems*, 15, 7: 630–641, July 2004.
20. J. Pan and W.J. Tompkins, A real-time QRS detection algorithm, *IEEE Transactions on Biomedical Engineering*, BME-32(3): 230–236, 1985.
- 2 1. The Merck Manuals, <http://www.merck.com/mmpe/sec07/ch075/ch075k.html>
- 2 2. TinyOS, <http://www.tinyos.net/>
23. MIT-BIH Database Distribution, <http://ecg.mit.edu/>
24. PhysioNet , <http://www.physionet.org/>

DIABETES

3

Chapter 7

Automated Blood Glucose Management Techniques Through Micro-Sensors

Fei Hu, Michael Lewis, and Yang Xiao

CONTENTS

7.1 Introduction	120
7.2 Glucose Micro-Sensors.....	121
7.2.1 F inger Sticks	121
7.2.2 I ntravenous Monitoring.....	121
7.2.3 S ubcutaneous Sensors.....	122
7.2.4 D ielectric Spectroscopy	122
7.2.5 Gl ucoWatch.....	122
7.2.6 C ommercial Sensors.....	123
7.3 I nsulin Pump	126
7.4 I njection Methods.....	127
7.4.1 Methods of Injection	127
7.4.2 Comparison of Methods.....	127

- 7.4.3 Subcutaneous Devices128
- 7.4.4 Micro-Needle Array.....128
- 7.4.5 Insulin Selection130
- 7.4.6 Commercial Pumps131
- 7.4.7 Safety.....131
- 7.5 Automated Insulin Injection Control131
 - 7.5.1 Partially Closed Loop Control131
 - 7.5.1.1 Physician-Prescribed Regimens.....131
 - 7.5.1.2 Diabetes Advisory System.....134
 - 7.5.2 Closed Loop Control134
 - 7.5.2.1 Pole-Assignment Control134
 - 7.5.2.2 Self-Tuning Adaptive Control136
 - 7.5.2.3 Model Predictive Control.....137
 - 7.5.2.4 Neural Network Control.....138
- 7.6 Conclusions.....138
- Acknowledgment139
- References139

Different aspects of blood glucose monitoring are explored as well as current technologies used to handle these tasks. In this chapter, we provide a review on the use of micro-sensors for automated blood glucose management. Insulin sensors including finger sticks, intravenous monitoring, subcutaneous sensors, and dielectric spectroscopy are covered. Piezoelectric pumps are described with regard to insulin injection. We explore different types of injection technologies including micro-needle arrays and different preparations of insulin. Several control schemes for blood glucose management are explained, including partial closed loop, pole assignment, self-tuning adaptive control, model predictive control, diabetes advisory system, and neural network system modeling.

7.1 Introduction

Diabetes mellitus is a disease that affects a body’s ability to regulate glucose.¹ Diabetes inhibits a person’s ability to produce or use insulin. Without insulin, the cellular system cannot properly convert carbohydrates such as sugars, starches, or other foods into energy usable by the body. Type 1 diabetes is characterized by the body’s failure to produce insulin, while type 2 diabetes is characterized by the body’s inability to properly use insulin that it has produced. Both conditions are chronic and currently incurable, seemingly linked to genetics. Complications stemming from diabetes are widespread and potentially life threatening. Heart disease, stroke, kidney disease (nephropathy), eye complications (including blindness), diabetic neuropathy and nerve damage, foot complications, skin complications, gastroparesis, and depression are all examples of health issues that can be caused or aggravated by diabetes. Diabetes is a condition that disproportionately affects developed countries. Estimates

place the number of affected Americans at 20.8 million people, accounting for 7.0% of the population. In 2002 alone, direct and indirect expenditures related to diabetes reached \$132 million,¹ a figure that is considered to be an underestimate.

Current treatment for diabetes can include home-administered care under the guidance of a physician. Intensive treatment can mitigate the effects of existing conditions as well as reduce the risk of developing advanced complications, such as those previously listed.

Early methods of home care involved using logs and tables, applying diet and exercise to predetermined doses prescribed by the patient's doctor. Modern micro-controllers, sensors, and pumps now allow for the automated administration of insulin within doctor-prescribed parameters. Further technological advances currently permit these devices to be wearable, acting as an "artificial pancreas." The goals of such a product include being safe, automatic, and noninvasive. These devices must employ an effective control scheme that allows for the blood glucose (BG) level to be kept within a safe range of nominal.

This chapter explores current components of automated blood glucose level management via insulin infusion. The rest of this chapter is organized as follows: Section 7.2 details types of blood glucose level sensors, such as finger sticks, intravenous dialysis, subcutaneous monitoring, and dielectric spectroscopy. Section 7.3 details micro-pumps, focusing on piezoelectric designs. Section 7.4 presents insulin injection technologies such as intravenous injection, subcutaneous injection (detailing micro-needle arrays), and the effects of different preparations of insulin. Section 7.5 covers blood glucose control techniques, spanning partial closed loop, pole-assignment, self-tuning adaptive control, model predictive control, a diabetes advisory system, and neural network system modeling. Section 7.6 concludes the chapter.

7.2 Glucose Micro-Sensors

7.2.1 Finger Sticks

Conventional glucose monitoring entails using finger-stick devices to obtain readings.¹ Finger-stick devices prick the tip of a finger, drawing blood, and measuring the glucose levels in the blood. This is invasive and can be painful. Additionally, measurements of this sort are only taken a few times per day due to the conscious effort required to perform them. Increasingly, different types of sensors are being developed and used, and they provide a wealth of benefits compared to finger sticks.

7.2.2 Intravenous Monitoring

Intravenous systems monitor blood glucose levels by drawing blood through a vascularly embedded needle.² This has the advantage of up-to-date, real-time BG levels, but at the expense of being invasive and painful. Additionally, there are long-term effects associated with prolonged vascular invasion, making it a suboptimal solution for continuous long-term home monitoring.

7.2.3 Subcutaneous Sensors

Subcutaneous glucose sensors are small electrode devices that can be inserted into the skin in the fatty tissues. This includes collecting a blood sample from the dermis layer of the skin,³ which is located about a tenth of a millimeter into the surface of the body.⁴ When the sensors are placed correctly, current proportional to the blood glucose level can be detected and measured. Due to its shallow position, subcutaneous monitoring can be significantly less painful than finger sticking. The greatest boon of subcutaneous monitoring, however, is that it can be performed continuously in a wearable fashion. This quality enables new types of control techniques to be exploited. The finer the measuring time increment is, the more accurate control methods will be. Continuous glucose monitoring permits real-time signal filtering in attempts to regulate closely any changes in glucose due to various factors such as meals, exercise, or sleep patterns.

There are several methods for administering subcutaneous monitoring, including dialysis and open-flow microperfusion. Currently, the most advanced method of subcutaneous monitoring is considered to be microperfusion.³ Microperfusion entails diffusing interstitial fluids from the dermis into a double lumen catheter where it can be monitored by an external sensor without actually drawing blood.

7.2.4 Dielectric Spectroscopy

Dielectric spectroscopy (DS) is a noninvasive, extracorporeal approach to continuous blood glucose monitoring.^{3,4} DS involves analyzing the electrolyte balance across cells, and comparing those results to known behavior for differing BG concentrations. One implementation of DS blood glucose monitoring entails coupling an open resonant circuit to the skin. This acts as an RCL sensor (“R” refers to a general organic molecule and “CL” refers to chlorine), comparing measurement results to derived system models in order to deduct the BG level of the patient.

7.2.5 GlucoWatch

GlucoWatch is a new glucose monitor from Cygnus Corp. The monitor straps to the wrist of the patient and uses a patented electrochemical sensor to measure glucose levels in the patient. The GlucoWatch works both continuously and noninvasively, permitting closed loop blood glucose level control. GlucoWatch displays the most recent blood glucose levels of the patient, updating every 20 minutes, and will sound an alarm if the blood sugar level goes above or below predetermined thresholds. The device stores the previous 4000 readings that can be offloaded for use by a physician. This permits accurate monitoring of long-term insulin dosage regimens and their results.

The GlucoWatch has also recently won FDA approval and is offered for sale in the United States and the U.K. The device is actually based upon technology that is

almost 100 years old. It takes advantage of the observation that an electric current can selectively transport chemicals through human skin. This transport phenomenon, called “iontophoresis,” has historically been seen as a one-way street, a way to get chemicals into the body. Cygnus scientists and engineers saw an untapped opportunity, creating a device that reverses iontophoresis to get the glucose out. “A lot of substances can be measured through reverse iontophoresis, but we felt there was a great unmet need for glucose monitoring,” says Dr. Russell Potts, a biochemist and Cygnus’ vice president of research.⁵

The watch applies a biosensor, called the AutoSensor, against the skin that measures BG levels, producing a current proportional to the BG level. A 20-minute analysis cycle starts as the sensor’s silver–silver chloride iontophoresis electrode applies a 300-microamp current to the skin. For the next three minutes, positive and negative ions travel through the patient’s skin to GlucoWatch’s side-by-side collection disks, which serve as an anode and cathode during glucose extraction. This ion migration acts as a glucose transport, depositing it at the cathode for it to be measured. The onboard microcontroller then interprets the glucose level into a standard unit of mm/dl.

The DirecNet group conducted a six-month randomized trial⁶ to measure the effects of the GlucoWatch continuous sensor on blood glucose control, hypoglycemia, and quality of life as compared to standard care. At the end of six months, there was no measurable difference in blood glucose control between the experimental and control groups, as measured by A1c and mean glucose using the Medtronic retrospective CGMS device. The results also showed that use of the device had no positive or negative psychological impact on the subjects in the experimental group. These results were puzzling until the usage data was reviewed, which tracked the number of times per week subjects actually used the device. During the first month, 64% of subjects used the device at least twice per week (2.1 ± 0.8). However, by the third month, average use was only 1.6 ± 0.7 times per week, and 7 of the 99 subjects had discontinued use altogether. By the sixth month, average use was 1.5 ± 0.6 times per week, and 26 of the original 99 subjects had discontinued use. In summary, differences in clinical outcome failed to materialize because an increasing number of subjects stopped using the device. Data gathered from our questionnaires revealed that families felt the information gained from the device was not worth the discomfort and adhesive problems encountered with its use.

7.2.6 Commercial Sensors

Here we use Table 7.1 to compare four typical commercial glucose sensors (Abbott,⁷ MiniMed Paradigm,⁸ MiniMed Guardian,⁹ and DexCom¹⁰) from the following aspects:

1. Accuracy of measurements
2. Start-up time

Table 7.1 Specification Comparison of Commercial Blood Glucose Sensors

Features	<i>Abbott Freestyle Navigator</i>	<i>MiniMed Paradigm Real-Time Systemb</i>	<i>MiniMed Guardian Real-Time Systemc</i>	<i>DexComd</i>
Accuracy	Varies	Consensus error grid: 98.9% A + B Mard(mean): -19.7%(median): -15.5%	Consensus error grid: 98.9% A + BMard(mean): -19.7%(median): -15.5%	Consensus error grid: 95.4% A + BMard(mean): -49%(median): -15.9%
Startup initiation time	10 h	2 h	2 h	2 h
Sensor life	5 day wear indication	Above 72 h	Above 72 h	Above 72 h
Calibration method	Requires calibration at 10, 12, 24, and 72 h after insertion of sensor	Alarms when calibration value is not entered on time; first and second calibration should be done for 2 h and 6 h after insertion	Alarms when calibration value is not entered on time; first and second calibration should be done for 2 h and 6 h after insertion	First calibration after 30 min, then for every 12 h; manual calibration not possible
Frequency of display	Every 1 min	Every 5 min	Every 5 min	Every 5 min
Transmitter memory	Yes, transmitter stores missed data for up to 40 min	Yes, transmitter stores missed data for up to 40 min	Yes, transmitter stores missed data for up to 40 min	No, transmission lost is data lost
Range of monitor to transmitter	10 ft	6 ft	6 ft	5 ft

Monitor batteries	Uses two AAA batteries with replacement every three months	No separate monitor required; uses insulin pump	Uses two AAA batteries; indication set for chance of battery	Uses rechargeable batteries
Monitor size	3" x 2.5"	Separate monitor not available; uses insulin pump for display	3" x 2.7"	3" x 2.5"
Alarms on user set low and high threshold	Applicable	Applicable	Applicable	Applicable

^a Abbott Diabetes Care, Freestyle Navigator Continuous Glucose Monitoring System, June 19, 2007, http://www.abbottdiabetes-care.com/adc_dotcom/url/content/en_US/10.10:10/general_content/General_Content_0000163.htm

^b Medtronic, MiniMed Inc., Real-Time Continuous Glucose Monitoring, June 19, 2007, <http://www.minimed.com/products/insulinpumps/components/cgm.html>

^c Medtronic, MiniMed Inc., The Guardian Real-Time Continuous Glucose Monitoring System, June 19, 2007, <http://www.minimed.com/products/guardian/index.html>

^d DexCom, DexCom Products, June 19, 2007, http://www.dexcom.com/html/dexcom_products.html

3. Sensor lifetime with batteries
4. How they calibrate values
5. How frequently they display the data
6. Memory size to store the data
7. Transmission distance from the sensor to a monitor
8. Batteries
9. Monitor size
10. Alarm system

7.3 Insulin Pump

Currently, one of the most promising pump technologies is a piezoelectric fluid device.¹¹ Piezoelectric pumps operate by applying voltage to a thin lead zirconate titanate (PZT) film. This distorts the film, causing it to pump fluid through an adjoining silicon nitride membrane, shown in Figure 7.1.

The displacement volume of an unloaded pump can be described as:

$$\Delta V = \frac{3r^4(5+2\mu)(1-\mu)d_{13}U}{4h^2(3+2\mu)} \quad (7.1)$$

$$K = \frac{3r^4(5+2\mu)(1-\mu)d_{13}}{4h^2(3+2\mu)} \quad (7.2)$$

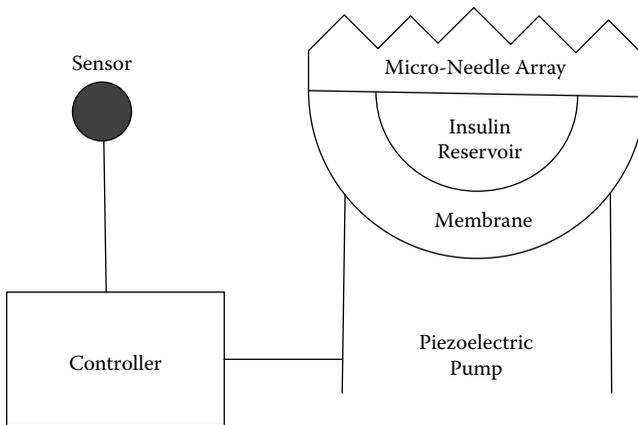


Figure 7.1 Piezoelectric pump and micro-needle device.

$$\Delta V = K \Delta U \quad (7.3)$$

$$Q = Kf \Delta U \quad (7.4)$$

where r is the radius of the membrane in μm , h is the thickness of the membrane in μm , μ is the Poisson's ratio, ΔU is the peak-to-peak voltage applied to the pump every period, d_{13} is the piezoelectric coefficient, and f is the frequency at which voltage is applied to the pump. K represents the pump coefficient described by Equation 7.1. Q is the flow rate of the pump.

Note that ΔV is linearly proportional to ΔU . Likewise, Q is also linear for a fixed f or a fixed ΔU . This permits very simple and intuitive control of the pump by the controller, allowing it to control the flow rate through a variable amplitude signal or a variable frequency signal, depending upon system requirements and the resources available.

7.4 Injection Methods

7.4.1 Methods of Injection

Several methods of injection are available for glucose administration to diabetes patients. The most direct method is intravenous infusion,² while subtler ways can be used for continuous or less painful dispensation, such as subcutaneous injection. Furthermore, the method of injection as well as the control method places a number of requirements on the type of insulin applied to the patient.

Intravenous infusion works by injecting drugs directly into a patient's bloodstream through a needle, which penetrates the skin, and into a vein. This is the most direct method of administration and will have the most immediate effects, but it has disadvantages as well.

Subcutaneous insulin injection is a far less invasive procedure than intravenous injection,² although it presents certain design challenges for blood glucose level control. It can be used to facilitate continuous, noninterrupted glucose management. Due to its relatively pain-free application (compared to intravenous infusion) it is much more agreeable to home monitoring patients who wear artificial pancreas devices at all times.¹¹

7.4.2 Comparison of Methods

Because it is less invasive to the body, subcutaneous injection is generally safer for several reasons.² Persons with an active lifestyle are put at risk by leaving an embedded needle protruding from their body, necessary to bridge the fluids gap

between their bodies and their artificial pancreas. Subcutaneous administration greatly reduces this safety hazard by interfacing no deeper than the skin layers, often distributed over an area, decreasing the invasiveness of the device.¹¹ Furthermore, continuous needle-sticking as well as prolonged embedding poses health risks of infection, clotting, or other sort of body-triggered rejection to an invasive foreign device. Because subcutaneous injection is far less invasive, it is less likely to trigger self-defense mechanisms like this, and if triggered, they tend to be far less severe.

Subcutaneous injection is not without its drawbacks, however. Due to its less invasive nature, insulin takes far longer to permeate into the body than through direct intravenous infusion. It has been reported that when using standard insulin for both injection types, subcutaneous injection can take as much as three times longer to take effect. This makes implementing an accurate control system quite challenging due to the time delays.

7.4.3 Subcutaneous Devices

Several methods for subcutaneous injection have been proposed and implemented in order to take advantage of the benefits of subcutaneous injection.² Manual injection systems typically include a subcutaneous needle or injection pen, which are applied according to a physician-prescribed schedule. Continuous control systems generally deal with its time-to-act shortfalls through appropriate algorithms and choice of insulin.

7.4.4 Micro-Needle Array

One subcutaneous device to be proposed is a micro-needle made out of silicon.¹¹ Micro-needles for insulin injection are designed to penetrate through several layers of skin: the stratum corneum layer, the epidermis layer, and part of the dermis layer. This method of injection punctures through far fewer nerve cells than classic intravenous injection, while still having access to the dermis layer, which is rich in blood vessels. This results in a far less invasive and painful injection experience for the patient.

An expansion upon this idea has been to fabricate an array of micro-needles. This provides many benefits, with few disadvantages. The more micro-needles used, the smaller in diameter each needle needs to be in order to facilitate a load-free flow of insulin into the body. Shrinking the size of the needles translates into an even less painful experience, while maintaining the same operational insulin flow rate. Additionally, an array of needles provides redundancy to the injection system, safeguarding against reduced flow due to channel blocking or clotting. Flow rate Q can be found as follows, given n needles, pressure change ΔP , needle length L , needle radius r , and fluid viscosity μ , where the relationship of R and r can be found in Figure 7.2:

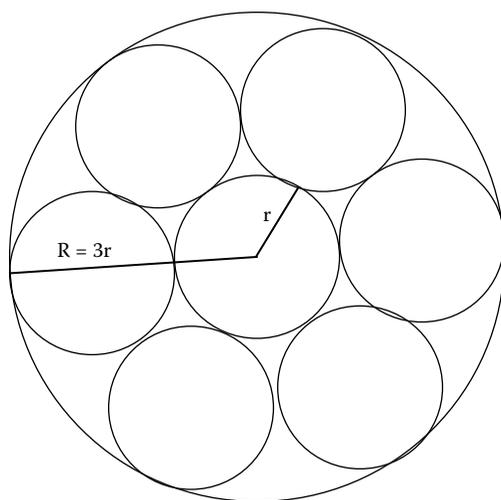


Figure 7.2 Micro-needle honeycomb formation.

$$R = \frac{8\mu L}{\pi r^4} \quad (7) \quad .5)$$

$$Q = n \frac{\Delta P}{R} \quad (7) \quad .6)$$

$$Q = n \Delta P \left(\frac{\pi r^4}{8\mu L} \right) \quad (7) \quad .7)$$

Note that the flow rate is linearly proportional to the number of micro-needles, given a fixed radius. Unfortunately, the flow also decreases by an order of four as radius is reduced.

Replacing a single large needle with a honeycomb formation of seven tightly packed micro-needles (each one third the diameter of the original needle) produces a flow of only 8.6% of the original flow for the same surface area. This observation makes evident the fact that due to the high order impact of needle radius, the skin area required for a micro-needle array increases exponentially as the size of the needles are reduced. In practice, the flow rate required is actually quite low, however, making this exponentially increasing contact area functionally negligible to the end user, due to its still relatively small size. As far as the patient is concerned, the increase in skin surface required is well worth the significantly reduced invasiveness and pain of the device.

7.4.5 Insulin Selection

Several options are available for insulin to be injected, depending upon the situation.² Regular insulin has historically been used to treat diabetes. It can be both intravenously and subcutaneously injected, depending upon the injection interface used. When it is subcutaneously injected its effects take up to three times longer than if it is intravenously injected, posing a difficult control systems problem.

An insulin preparation by the name of Lispro was developed to aid in the rapid absorption of the insulin.² It is designed to take immediate effect, and can be fully absorbed into the body's system significantly faster than regular insulin. Lispro is designed to take effect within 15 minutes and peak about an hour after application. Because of its fast-delivery capabilities Lispro is the most commonly used with subcutaneous injection. Subcutaneous injected Lispro has been proven to take effect within approximately the same time frame as intravenous injected normal insulin. This enables home monitoring and injection patients to control their glucose levels to the same levels as with intravenous management but in a far less invasive and painful way.

Figure 7.3 shows the increased absorption rate of Lispro compared to regular insulin.

An insulin preparation by the name of NPH has also been developed for the exact opposite purpose of Lispro.² NPH is designed to take effect over a much longer period of time, lasting up to 24 hours. The purpose of insulin with this property is to provide the patient with a predictable baseline level of insulin throughout

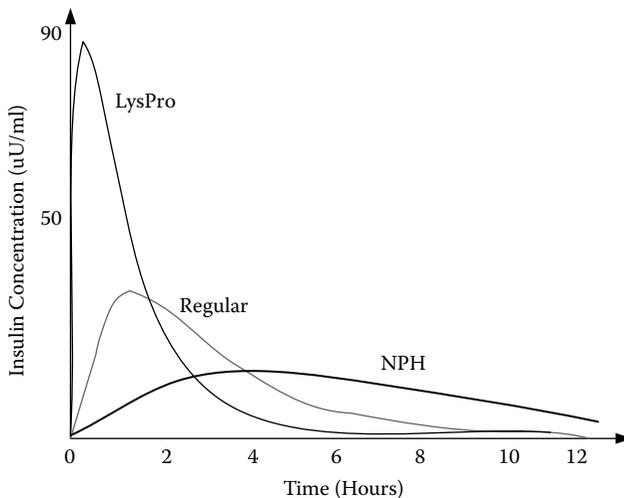


Figure 7.3 Comparison of insulin preparations. (Source: R. Bellazzi, G. Nucci, and C. Cobelli, *The subcutaneous route to insulin-dependent diabetes therapy*, *IEEE Engineering in Medicine and Biology*, Jan./Feb., 2001. With permission.)

the day, allowing the control system to focus its efforts on accounting for and correcting anomalies and disruptions that are encountered, such as meals, exercise, and sleep.

Appropriate injections of a combination of Lispro and NPH has been proven incredibly effective in maintaining short- and long-term glucose stability within patients using a continuous monitoring and control device.¹¹ Lispro is usually used for fast-acting control to counterbalance blood glucose fluctuations due to meals. NPH is usually used to provide a constant trickle of insulin in order to establish a baseline, known as a “basal” level.

7.4.6 Commercial Pumps

We compare five types of popular insulin pumps in Table 7.2.

7.4.7 Safety

Pump reliability is an incredibly crucial factor when considering a blood glucose management system. Pump, sensor, or control failure can lead to incorrect dosages or even a failure to inject. Due to its mechanical nature, pump failure is often the most likely cause of malfunction. This can result from O-ring leaks, air bubbles, bleeding, infection, and clogs. These conditions can easily result in hyperglycemia or hypoglycemia and must be monitored for.

7.5 Automated Insulin Injection Control

Arguably, the most complex component of blood glucose management is the control domain. There are several classes of solutions to this problem, ranging in complexity, prerequisite knowledge, and feedback.

7.5.1 Partially Closed Loop Control

7.5.1.1 Physician-Prescribed Regimens

The insulin regimen prescribed by a doctor, to be administered manually, constitutes partially closed loop control.^{2,12} A physician will dictate an insulin administration routine to a patient, variant upon a patient’s lifestyle. Patients under such a system monitor their blood glucose level several times a day, administering insulin based upon prescribed tables according to their schedule and BG level. Figure 7.4 shows the procedure.

Table 7.2 Specification Comparison of Commercial Insulin Pumps

Company	<i>Animas</i>	<i>Deltec</i>	<i>Disetronic</i>	<i>MiniMed</i>	<i>Insulet</i>
Model	IR-125016a	Cozmob	Spiritc	Paradigm 522/722d	OmniPod
Dimensions	79 × 51 × 19	80 × 47 × 24	80 × 56 × 20	522: 51 × 79 × 20722: 51 × 79 × 20	Pod: 41 × 61 × 18PDA: 66 × 110 × 26
Screen size	992 mm2	870 mm2	Unavailable	774 mm2	1848 mm2 on PDA controller
Basal delivery	Every 3 min	Every 3 min	Every 3 min	Varies	Information unavailable
Basal temperature	Initially -90 to +200%, varies for every 0.5 h	Varies from 0 to 200% with an increment of 5% for every 0.5 h	Varies from 0 to 200% with an increment of 10% for every 0.5 h	0.1 increment as single basal rate for 0.5 to 24 h	Information unavailable
Carb and correction factors	Manual entry and assist from EZ Manager	Manual carbohydrate, BG from attached CoZ monitor	Manual carbohydrate, BG from Accu-Check BG monitor	Manual carbohydrate, BG from BD meter or manual entry	Information unavailable
Battery	AA lithium × 1	AAA × 1	AA × 1 alkaline or rechargeable	AAA	AAA × 2
Motor	DC	DC	DC	DC	Stepper
Memory	Nonvolatile: 600 bolus, 270 basal, 120 daily totals, 30 alarms, 60 primes	Nonvolatile: 90 days of basal, carbohydrates, boluses, correction boluses, alarms	Nonvolatile: 90 days history recall of last 30 boluses, alerts, daily insulin totals, temporary basal rate increase	4000 events volatile: 24 boluses, 7 days totals	90 days of data

Extra features	Clip-on covers, personalized carbohydrate and correction factors, tracks residual bolus insulin.	Carbohydrate and correction factors, tracks residual bolus insulin, detailed records of pump, daily bolus total correction.	Availability of different types of user menus, icon and menu driven programming, backlight display, reversible display screen	Extended bolus, auto off	Integrated freestyle meter, 1000 common foods in PDA
----------------	--	---	---	--------------------------	--

^a Animas Corporation, Animas Corporation IR1250 Specifications, June 19, 2007, http://www.animascorp.com/products/pr_insulinpump_IR1250Spec.shtml

^b Cozmore Insulin Technology System, Deltac Cosmo Insulin Pump, June 19, 2007, <http://www.cozmore.com/default.cfm/PID=1.7>

^c Disetronic USA, ACCU-CHEK Spirit Insulin Pump System, June 19, 2007, http://www.disetronic-usa.com/dstrnc_us/rewrite/content/en_US/3.2.20/article/DCM_general_article_273.htm

^d Medtronic, MiniMed Inc., MiniMed Paradigm 522 or 722 Insulin Pump, June 19, 2007, <http://www.minimed.com/products/insulinpumps/components/insulinpump.html>

^e Insulet Corporation, Product Specifications: OmniPod Insulin Management System, June 19, 2007, <http://www.myomnipod.com/products/section/156>

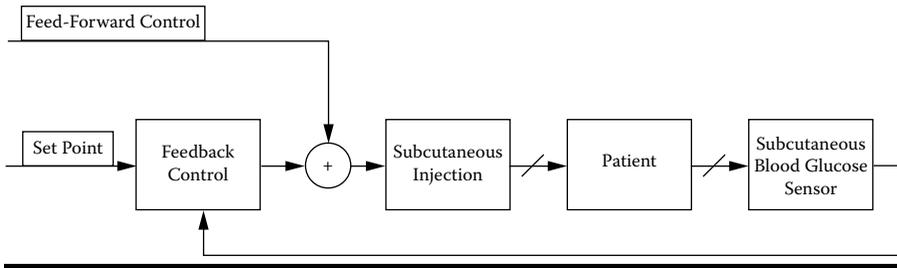


Figure 7.4 Partially closed loop control.

This is the method that has traditionally been used by insulin dependent diabetics, but it performs poorly compared to other methods. Partially closed loop control is far from real-time and only updates its control routine at scheduled physician visits. Furthermore, life style events such as eating, sleeping, or working out must be accounted for by the patient in their interpretation of insulin tables, introducing the very real danger of human error.

7.5.1.2 Diabetes Advisory System

The Diabetes Advisory System (DIAS) is a nonlinear model of the blood glucose–insulin system based upon real-life parameters, versus simply BG measurements.² It incorporates qualitative and quantitative input from the user, including BG levels, meals, and past insulin injections.

The system uses a discrete-time finite-state model of the system based upon user input. The system uses what it understands about the system as a whole, including dormant compartmentalized insulin, predigested carbohydrates, and current BG levels to compute a Bayesian estimate of future BG levels. It uses all known information in order to compute the value of the optimal dosage such that it minimizes an associated cost function (i.e., hypoglycemia is far more “costly” than hyperglycemia, due to its possibility of severe damage). Over iterations it will adjust its system model parameters to better account for patient specific reactions it detects.

7.5.2 Closed Loop Control

The closed loop control uses the feedback from the output, shown in Figure 7.5. It has four types, discussed in the following subsections.

7.5.2.1 Pole-Assignment Control

Pole assignment is a standard control systems technique for designing an infinite impulse response (IIR) filter.^{2,12} This consists of a set of filter coefficients and a feedback loop in order to maintain a stable BG level.

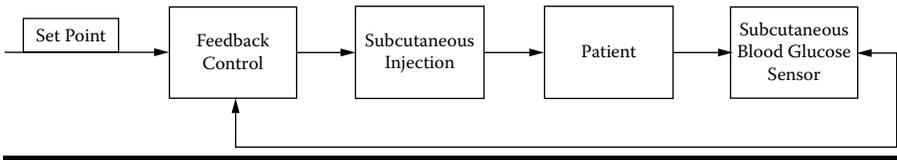


Figure 7.5 Closed loop control.

This is a simplified approach, forgoing adaptive control for ease of characterization and implementation. For most situations, it will perform as desired but if it encounters a situation that it handles poorly, it will handle that situation poorly every time it occurs again in the future. Because of the time-delay difficulties associated with subcutaneously injected insulin, pole-assignment control fares quite poorly when used with normal insulin. Subcutaneously injected Lispro, however, helps mitigate this problem, performing on the same level as intravenously injected normal insulin.

Several important concepts must be understood in order to design a successful coefficient-based closed loop filter for blood glucose level control. The system model relating insulin concentration $[I(t)]$ to blood glucose concentration $[G(t)]$ can be described as

$$I(t) = aG(t) + b \frac{dG(t)}{dt} + c \tag{7.8}$$

where a , b , and c are coefficients relating $G(t)$ to $I(t)$. Given the insulin infusion rate $[IR(t)]$ and plasma volume $[V]$, the model^{2,11} describing the effects of subcutaneously injected insulin is

$$\begin{aligned} \frac{dX(t)}{dt} &= IR(t) - lX(t) \\ \frac{dY(t)}{dt} &= lX(t) - (p + o)Y(t) \\ \frac{dZ(t)}{dt} &= pY(t) - nZ(t) \\ I(t) &= \frac{Z(t)}{V} \end{aligned} \tag{7.9}$$

where X , Y , and Z represent the insulin level in the two subcutaneous compartments and in the plasma, respectively. Figure 7.6 shows such an X/Y/Z three-level model.

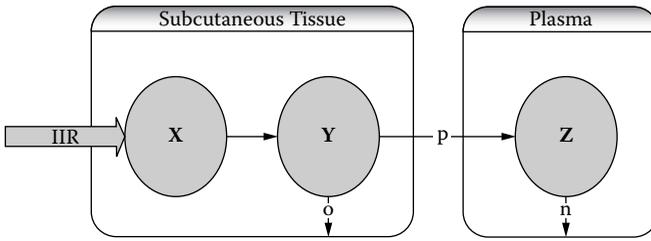


Figure 7.6 Three-level model of subcutaneous insulin absorption.

All model parameters can be determined by nonlinear least squares estimation using experimental data from patients.

The two models represented by equations (7.8) and (7.9) can be combined in order to express the relationship between insulin’s absorption rate and its effect on blood glucose levels. This relationship is required to design a feedback filter for controlling the insulin infusion rate

$$\begin{aligned}
 m &= o + p \\
 K_p &= \frac{amnV}{p} \\
 \frac{K_d}{K_p} &= \frac{1}{l} + \frac{1}{m} + \frac{1}{n} + \frac{b}{a} \\
 K_c &= d + \frac{c}{a} K_p \\
 IR(t) &= K_p G(t) + K_d \frac{dG(t)}{dt} + K_c
 \end{aligned}
 \tag{7.10}$$

where d represents the intravenous basal infusion rate. This design can be implemented simply and inexpensively using a standard control system.

7.5.2.2 Self-Tuning Adaptive Control

Implementation of a self-tuning adaptive control, shown in Figure 7.7, is quite similar to pole-assignment control, as it uses the same system modeling equations in order to compute the insulin infusion rate.²

The primary difference between the two methods is that another controller is used to constantly evaluate the system model, and may “tune” or redesign the PD controller parameters as needed to obtain more accurate results based upon minimum variance. This is accomplished by recursively examining past BG levels and

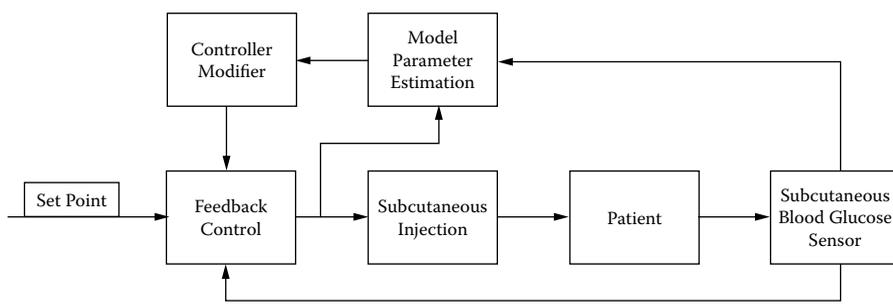


Figure 7.7 Self-tuning adaptive control.

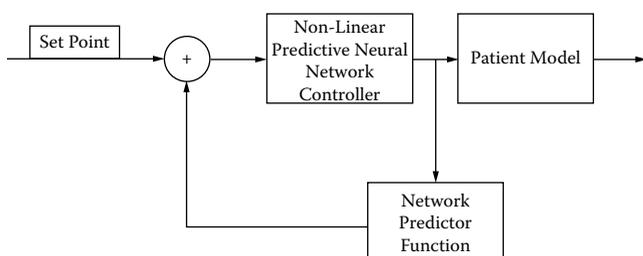


Figure 7.8 Model predictive closed loop control.

insulin doses, and predicting future BG levels. This is used to tune the PD controller actively in an attempt to avoid lagging behind the effects of injected insulin.

7.5.2.3 Model Predictive Control

A model predictive control (MPC), or nonlinear predictive control (NLPC) algorithm attempts to “learn” what nominal means in a system,^{2,3} shown in Figure 7.8. In the case of blood glucose management, a nonlinear MPC algorithm uses sensor data to track glycoregulatory system parameters in order to predict the levels of required insulin infusions. It then uses models of the human glucose metabolism to estimate the effects of the insulin injection. An example of a model used is a nonlinear autoregressive (NARX) model, where previous BG levels and insulin dosage levels are run through a nonlinear function, often obtained through neural network learning.

Bayesian learning is applied using the model-predicted effect of the insulin, and the actual measured effect of it. The learning process adjusts system parameters in order to increase the accuracy of its predictions as more iterations are performed. Using this method, the system will become increasingly accurate, and will begin to “understand” how the patient that it is calibrated to will react to insulin injections of varying compositions and strengths.

7.5.2.4 Neural Network Control

Neural networks, shown in Figure 7.9, approach the problem of blood glucose management without attempting to describe explicitly the exact model of the blood glucose–insulin system.^{13–15}

This is particularly useful in situations where patients have a disease that complicates normal model description, or an abnormality exists which makes prediction difficult using just measured parameters and sensor data.

A feed-forward neural network employing backpropagation can be trained offline using accumulated patient data, including daily BG readings as well as insulin dosages. A neural network will then be able to “learn” based upon experience, much as a human brain learns. This will help it to predict nonlinear behavior, even multiple orders removed, imperceptible to standard data interpretation methods. This capability to be “intuitive” helps to drive a system in which unknowns or immeasurable parameters are still accounted for, and abnormalities are detected and intelligently handled.

7.6 Conclusions

This chapter has systematically reviewed major relevant technologies on diabetes monitoring through glucose micro-sensors, insulin pumps, and control systems between them.

It is evident that reliable solutions to diabetes management are highly sought after and researched. The health benefits associated with intensive diabetes treatment can save untold lives and make life far more comfortable for many others. Additionally, with the number of diabetes sufferers increasing annually, and the

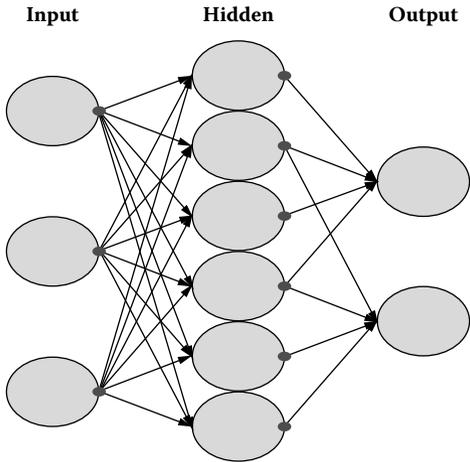


Figure 7.9 Neural network nonlinear predictive controller.

disease already costing well over \$100M per year in health care expenses, it is a very lucrative market. Effective solutions can help mitigate and control the effects and tolls of diabetes, in terms of both health and economic impact.

Personal “artificial pancreas” devices are coming into greater use, spurred on by improvements in technology and an increasing demand. Appropriate technologies must be developed and exploited in order to improve upon today’s methods. Increased miniaturization of sensors, pumps, and controls continuously improve the portability of personal continuous-control devices, thereby increasing their ubiquity among insulin-dependent diabetes patients. Additionally, control methods are constantly being refined in attempts to control more accurately the effects of diabetes in regard to glucose and insulin levels. As more is understood and researched, it is a possibility that the “artificial pancreas” may someday achieve the target of performing functionally equivalent to a real pancreas.

Acknowledgment

This work was partially supported by the U.S. National Science Foundation (NSF) under grants CNS-0716211 and CNS-0716455.

References

1. American Diabetes Association, Diabetes Information, 20 April 2007, <http://www.diabetes.org/about-diabetes.jsp>
2. R. Bellazzi, G. Nucci, and C. Cobelli, The subcutaneous route to insulin-dependent diabetes therapy, *IEEE Engineering in Medicine and Biology*, January/February 2001.
3. R. Dudde, T. Vering, G. Piechotta, and R. Hintsche, Computer-aided continuous drug infusion: Setup and test of a mobile closed loop system for the continuous automated infusion of insulin, *IEEE Transactions on Information Technology in Biomedicine*, 10, 2, 2006.
4. M.S. Talary, F. Dewarrat, A. Caduff, A. Puzenko, Y. Ryabov, and Y. Feldman, An RCL sensor for measuring dielectrically lossy materials in the MHz frequency range, *IEEE Transactions on Dielectrics Electrical Insulation*, 13, 2, 2006.
5. Sensors keep watch on diabetes, *Design News*, June 4, 2001, <http://www.designnews.com/article/CA83168.html>
6. Diabetes Research in Children Network (DirecNet) Study Group, Youth and parent satisfaction with clinical use of the GlucoWatch G2 Biographer in the management of pediatric type 1 diabetes, *Diabetes Care*, 28, 8, 1929–1935, 2005.
7. Abbott Diabetes Care, Freestyle Navigator Continuous Glucose Monitoring System, June 19, 2007, http://www.abbottdiabetescare.com/adc_dotcom/url/content/en_US/10.10:10/general_content/General_Content_0000163.htm
8. Medtronic, MiniMed Inc., Real-Time Continuous Glucose Monitoring, June 19, 2007, <http://www.minimed.com/products/insulinpumps/components/cgm.html>
9. Medtronic, MiniMed Inc., The Guardian Real-Time Continuous Glucose Monitoring System, June 19, 2007, <http://www.minimed.com/products/guardian/index.html>

10. DexCom, DexCom Products, June 19, 2007, http://www.dexcom.com/html/dexcom_products.html
11. R. Yang, M. Zhang, and T. Tarn, Dynamic modeling and control of a micro-needle integrated piezoelectric micro-pump for diabetes care, *IEEE*, 2006.
12. E. Carson and T. Deutsch, A spectrum of approaches for controlling diabetes, *IEEE Control Systems*, 12, 6, 25–31, 1992.
13. M. Alamaireh, A predictive neural network control approach in diabetes management by insulin administration, *IEEE*, 2006.
14. S. Jaafar and D. Ali, Diabetes mellitus forecast using artificial neural network (ANN), *Asian Conference on Sensors and the International Conference on New Technologies in Pharmaceutical and Biomedical Research Proceedings*, September 2005, pp. 5–7.
15. E. Teufel et al., Modeling the glucose metabolism with backpropagation through time trained Elman nets, *IEEE XIII Workshop on Neural Networks for Signal Processing*, 2003.
16. Animas Corporation, Animas Corporation IR1250 Specifications, June 19, 2007, http://www.animascorp.com/products/pr_insulinpump_IR1250Spec.shtml
17. Cozmore Insulin Technology System, Deltec Cosmo Insulin Pump, June 19, 2007, <http://www.cozmore.com/default.cfm/PID=1.7>
18. Disetronic USA, ACCU-CHEK Spirit Insulin Pump System, June 19, 2007, http://www.disetronic-usa.com/dstrnc_us/rewrite/content/en_US/3.2:20/article/DCM_general_article_273.htm
19. Medtronic, MiniMed Inc., MiniMed Paradigm 522 or 722 Insulin Pump, June 19, 2007, <http://www.minimed.com/products/insulinpumps/components/insulinpump.html>
20. Insulet Corporation, Product Specifications: OmniPod Insulin Management System, June 19, 2007, <http://www.myomnipod.com/products/section/156>
21. V. Calhoun, T. Adali, and J. Liu, A feature-based approach to combine functional MRI, structural MRI and EEG brain imaging data, *Proceedings of the 28th IEEE EMBS Annual International Conference*, September 2006.
22. P. Magni and R. Bellazzi, A stochastic model to assess the variability of blood glucose time series in diabetic patients self-monitoring, *IEEE Transactions on Biomedical Engineering*, 53, 6, 2006.
23. S.S. Pruna, R. Dixon, and N.D. Harris, Black Sea TeleDiab: Diabetes computer system with communication technology for Black Sea Region, *IEEE Transactions on Information Technology in Biomedicine* 2, 3, 1998.
24. S. Shea et al., Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) Project, *Journal of the American Medical Informatics Association*, 9, 1, 2002.
25. R. Hovorka et al., Dynamic Updating in DIAS-NIDDM and DIAS Causal Probabilistic Networks, *IEEE Transactions on Biomedical Engineering*, 46, 2, 1999.
26. P. Dua, F.J. Doyle, III, and E.N. Pistikopoulos, Model-based blood glucose control for type 1 diabetes via parametric programming, *IEEE Transactions on Biomedical Engineering*, 53, 8, 2006.
27. National Diabetes Information Clearinghouse (NDIC), National Diabetes Statistics, June 12, 2007, <http://diabetes.niddk.nih.gov/dm/pubs/statistics/index.htm>

28. C. Owens et al., Run-to-run control of blood glucose concentrations for people with type 1 diabetes mellitus, *IEEE Transactions on Biomedical Engineering*, 53, 6, 2006.
29. K. Gerlach, A. Kaeding, S. Kottmair, D. Westphal, G. Henning, and K. Piwernetz, The implementation of a quality-net as a part of the European Project DIABCARE Q-Net, *Transactions on Information Technology in Biomedicine*, 2, 2, June 1998.
30. R. Hovorka, S. Andreassen, J.J. Benn, K.G. Olesen, and E.R. Carson, Causal probabilistic network modeling — An illustration of its role in the management of chronic diseases, *IBM Systems Journal*, 31: 635–648, 1992.
31. O.K. Hejlesen, S. Andreassen, R. Hovorka, and D.A. Cavan, DIAS — The Diabetes Advisory System: An outline of the system and the evaluation results obtained so far, *Computer Methods and Programming in Biomedicine*, 54: 49–58, 1997.
32. D.A. Cavan, R. Hovorka, O.K. Hejlesen, S. Andreassen, and P. Sonksen, Use of the DIAS model to predict unrecognized hypoglycemia in subjects with insulin-dependent diabetes, *Computer Methods and Programming in Biomedicine*, 50: 241–246, 1996.
33. O.K. Hejlesen, S. Andreassen, D.A. Cavan, and R. Hovorka, Analyzing the hypoglycemic counter-regulation: A clinically relevant phenomenon, *Computer Methods and Programming in Biomedicine*, 50: 231–240, 1996.
34. L. Santaso and I.M.Y. Mareels, Markovian framework for diabetes control, *Proceedings of the 40th IEEE Conference on Decision and Control*, December 2001.
35. L. Santaso and I.M.Y. Mareels, A direct adaptive control strategy for managing diabetes mellitus, *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.
36. J.A. Tamada, M. Lesho, and M.J. Tierney, Keeping watch on glucose, *IEEE Spectrum*, 39, 4, 52–57, 2002.
37. T. Iokibe, M. Yoneda, and K. Katika, Chaos based blood glucose prediction and insulin adjustment for diabetes mellitus, *IEEE Bio Medical Engineering MBS Asian-Pacific Conference*, October 2003, 86–87.
38. H.A. Klein, and A.R. Meininger, Self management of medication and diabetes: Cognitive control, *IEEE Transaction on Systems, Man, and Cybernetics — Part A: Systems and Humans*, 34, 6, 2004.
39. T. Katayama, T. Sato, and K. Minato, A Blood Glucose Prediction System by Chaos Approach, *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, September 2004, 750–753.
40. Y. Chen, T. Wu, C. Wu, M. Wu, and S. Jaw, Development of Wireless Blood Glucose Meter and Diabetes Self-Management System, *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, September 2004, 3384–3386.
41. U. Fischer et al., Does physiological blood glucose control require an adaptive control strategy?, *IEEE Transactions on Biomedical Engineering*, 34, 8, 575–582, 1987.
42. S. Kaushik et al., Lack of pain associated with microfabricated microneedles, *Brief Communications of the International Anesthesiology Research Society*, 92, 502–504, 2001.
43. Animas Technologies, LLC, Welcome to GlucoWatch — Consumer, June 20, 2007. http://www.glucowatch.com/us/consumer/frame_set.asp

Chapter 8

Mobile Telemedicine for Diabetes Care

Iñaki Martínez-Sarriegui, Gema García Sáez, M^a.
Elena Hernando, Mercedes Rigla, Eulalia Brugués,
Alberto de Leiva, and Enrique J. Gómez

CONTENTS

8.1 Introduction.....	144
8.2 The Diabetes Care Challenge.....	145
8.2.1 The Medical Problem.....	145
8.2.2 Diabetes Treatment.....	146
8.2.3 Telemedicine and Shared Care Services in Diabetes Management	147
8.2.4 Ambulatory Artificial Pancreas	148
8.3 Technical Requirements of Mobile Telemedicine Systems for Diabetes Care	149
8.4 Building the Mobile Telemedicine System	150
8.4.1 The DIABTel Distributed Architecture	151
8.4.2 Mobile Applications in DIABTel System	152
8.4.2.1 PDA Smart Assistant Application	153
8.4.2.2 WebPDA	154
8.4.2.3 DIABTelMobile	155

8.5 Conclusions 155
Acknowledgments156
References157

Diabetes mellitus is nowadays one of the most frequent noncontagious diseases in the world and remains a major health problem for national health care programs. It is well proved that telemedicine helps diabetic patients control their glucose levels, facilitating their day-to-day therapy management and communication with health care personnel. The rapid growth and development of information technologies in the areas of mobile computing and mobile Internet is shaping a new technological scenario of telemedicine and shared care systems. In this chapter we will show one approach to mobile telemedicine for diabetes care.

8.1 Introduction

According to the International Diabetes Federation,¹ 5.1% of the world’s population suffers from diabetes and the forecast for 2025 is that it will increase to 6.3%. In Western societies the numbers rise to 7.8% now and 9.5% by 2025. The situation of diabetes will get worse due to the number of people with alteration of glucose tolerance that affects 8.2% of the world population and by 2025 it will affect 9%. Another alarming fact is that for each case of diagnosed diabetes there is another that is not diagnosed.

Diabetes remains a major health problem being responsible for up to 8 % of national health care expenditure.² Diabetes mellitus is a chronic disease characterized by a sustained elevated blood glucose level, caused by a reduction in the action of insulin secretion where related metabolic disturbances generate severe, acute, and long-term complications that are responsible for premature death and disability.³ The World Health Organization projects that diabetes deaths will increase by more than 50% in the next ten years without urgent action.⁴ Most notably, diabetes deaths are projected to increase by over 80% in upper-middle income countries between 2006 and 2015.

Due to its multifactorial and systemic character, diabetes mellitus has been considered a paradigm of chronic disorders which has led to an extensive application of information technologies in diabetes care.⁵ Nowadays telemedicine provides an integrated approach to information technology tools, which enhances cooperation between users, information, and knowledge sharing.

Over the last three decades, diabetes has been a major focus for biomedical engineering efforts to improve the diagnosis, monitoring, and treatment of diabetic people. Earlier experiences aimed to facilitate the remote monitoring of patients from home by the transmission of computerized blood glucose profiles to the hospital.^{6–8} Most interactive telemedicine services were delivered using a distributed approach

integrating “patient units” (PU), implemented on a personal computer or a palm device and used by patients during their daily living; and “medical workstations” (MW), used by physicians and nurses at the hospital.^{9–12} Technology evolution has enabled the development of advanced systems based on more powerful, portable, and easy-to-use terminals and applications, such as hand-held electronic diaries with touch-screen,¹³ video telephones,¹⁴ or Web-based prototype systems.^{9,15}

The rapid growth and development of information technologies during recent years in the areas of mobile computing, computer-telephony integration (CTI), and mobile Internet are changing the way people access common services. Today people can buy cinema tickets, check the traffic situation, pay their bills, and perform many other tasks with their mobile telephone or PDA. And a growing sector of citizens demands more services available in a mobile way.

As more and more people use mobile handsets, a revolution is taking place in computing and telecommunications. Two extraordinary industries—the Internet and mobile communications—are converging. But this is just the beginning. As a third industry—consumer electronics—and a fourth—media and entertainment—join in, changes in consumer markets are inevitable, as evidenced by the explosive growth of mobile media, games, and entertainment.¹⁶

Telemedicine and shared care systems should not obviate the technology and market evolution. Today the seamless integration of available mobile and wireless technologies allows building a new scenario for telemedicine and shared care in diabetes in which the traditional concepts of PU and MU are blurred, suffering a profound transformation evolving to a new multi-access, mobile, universal, and ubiquitous workspace concept for diabetes care.¹⁷

In this chapter we present the telemedicine services requisites and the technical requirements of mobile telemedicine systems for diabetes care. These considerations were taken into account to build a multi-access mobile telemedicine workspace (Figure 8.1), implemented over the sustenance of a multi-agent architecture. This workspace provides universal access to a wide range of services both for patients and professionals. The chapter finishes by drawing some conclusions about current telemedicine and shared care experiences and the possibilities of mobile technologies in diabetes care in the future.

8.2 The Diabetes Care Challenge

8.2.1 The Medical Problem

Diabetes mellitus is a group of metabolic diseases characterized by a sustained elevated blood glucose level (hyperglycemia), caused by defects in insulin secretion, insulin action, or both. Insulin facilitates entry of glucose into muscle, adipose, and several other tissues, and stimulates the liver to store glucose in the form of glycogen. The chronic hyperglycemia generates severe, acute, and long-term complications that

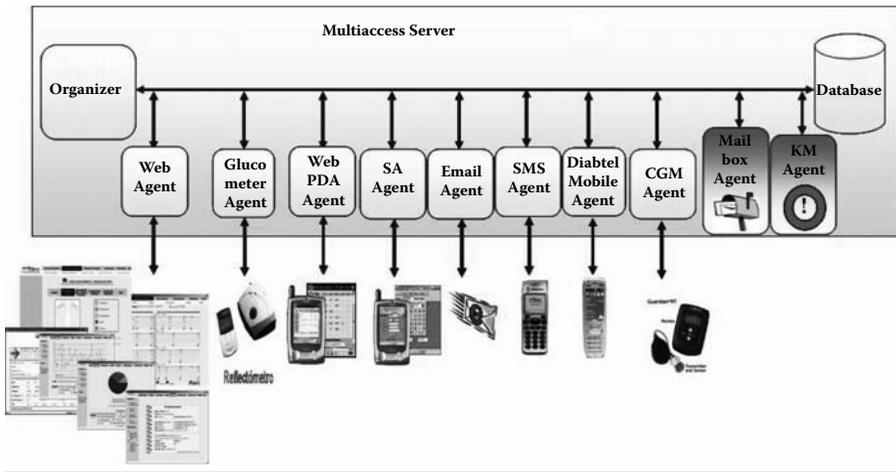


Figure 8.1 The mobile multi-access telemedicine workspace.

are responsible for premature death and disability.³ The cause of diabetes continues to be a mystery, although both genetics and environmental factors such as obesity and lack of exercise appear to play roles.¹⁸

The vast majority of cases of diabetes fall into three broad categories: type 1, type 2, and gestational diabetes.¹⁹ Type 1 diabetes comprises 10% of all cases of diabetes. The cause is an absolute deficiency of insulin secretion due to beta cells destruction, and the glucose cannot be metabolized. People with type 1 diabetes need insulin administration to use glucose from meals. By taking good control of blood glucose levels, type 1 diabetic patients can expect to have an almost normal life span as the benefits of intensive therapy management have been well established and include reduced long-term complications.²⁰

Type 2 is the most prevalent form of diabetes, comprising 90% of cases. It results from a combination of resistance to insulin action and an inadequate compensatory insulin secretor response. People with type 2 need diabetes oral medication or insulin injections to control their blood glucose levels. The inability of a large proportion of the population in the more affluent Western societies to cope with the excess caloric supply together with a lack of physical exercise results in a greater prevalence of type 2 diabetes.

Gestational diabetes may give rise to several adverse outcomes, including congenital malformations, increased birth weight, and an elevated risk of prenatal mortality, and affects about 4% of all pregnant women. Strict metabolic control may reduce these risks to the level of those of nondiabetic expectant mothers.

8.2.2 Diabetes Treatment

Effective control of patients' blood glucose level minimizes the progression of the disease and reduces the risk of long-term neurological, renal, and cardiovascular

complications. The treatment of diabetic patients attempts to achieve normoglycemia by maintaining a careful balance between diet, physical exercise, and insulin therapy.

Patients monitor their own blood glucose levels (self-monitoring) daily to predict and avoid hypoglycemia (low blood sugar level) and hyperglycemia (high blood glucose level) and to make decisions regarding the adjustment of changes of insulin doses, meals, and physical activity.

Insulin therapy aims to mimic the physiological insulin patterns delivered by the pancreas in nondiabetic people with the injection of short-acting insulin given under the skin (subcutaneously) just before a meal and long-acting insulin given in the morning and evening to simulate the basal supply.

Insulin cannot be taken orally as it would be broken down during digestion. Therefore, insulin must be injected and its efficiency depends on when glucose from food starts to enter the blood¹⁸ and the site of injection.

Scientific prospective studies have provided evidence on the use of intensive insulin therapy, based on multiple insulin injections or continuous insulin delivery, as the most appropriate method to equilibrate a patient's blood glucose and to delay the onset and slow the progression of complications later in life.

To obtain the best care outcomes, diabetic people should receive medical care from a physician-coordinated team.²¹ Such teams may include physicians, nurses, dietitians, pharmacists, and mental health professionals with expertise and interest in diabetes. They assess the patients' glycemic control on the basis of patient-reported monitoring blood glucose data as well as by the glycated hemoglobin (HbA1c) test, which provides an indication of the average blood glucose level over the previous two to three months. This collaborative and integrated shared care approach requires that individuals with diabetes assume an active role in their care, requiring in many cases better patient empowerment and education.

However, fulfilment of the current guidelines in diabetes management²² implies a significant increase in the amount of patient data to be monitored, increasing physicians' and nurses' workload, and raising immediate health care costs.

8.2.3 Telemedicine and Shared Care Services in Diabetes Management

A telemedicine system for chronic care has to help physicians' decision making. In the case of diabetes mellitus the responsibility of data collection is shared between patients, who record data during daily life (self-monitoring data) and health care personnel, who capture data during hospital visits (hospital patient record). This supposes a huge quantity of data and increases physicians' workload in analyzing them in order to make therapeutic decisions. The telemedicine system must provide enough information to enable assessment of the patient's condition and must present the relevant patient clinical data to define a therapeutic change. It should analyze patient's data,

filter the information, and generate automatic intelligent alarms to focus doctors' attention on those points where undesired deviations in the metabolic control are found.

A telemedicine system for diabetic people has to provide telemonitoring of the patients' monitoring data in order to allow physicians to know the state of the patient. It should also allow teleconsultation between patients and doctors through an asynchronous message exchange. Another service that it must supply is supervised care, providing patients with a "supervised autonomy." Patients carry out day-to-day therapy changes and send this information to the physician who reviews the data and validates patients' decisions or proposes further adjustments of the treatment.

To obtain the best care outcomes, diabetic people should receive medical care from a physician-coordinated team. The information technologies provide helpful tools for this coordination. One example is the application of computer-supported collaborative work (CSCW), current theories and tools to create new information technologies, and telecommunications workspaces to bear diabetes shared care.²³

The services to be provided by a shared care workspace for diabetes management are

- Multi-channel messaging services, to allow users to send and receive messages in any format despite the original message's format
- Shared information dissemination and request, to send data automatically to interested users using both pull and push methods
- Group and activity awareness, to integrate an event and activities notification service; the aim of this service is to alert every user of the most relevant events occurring at the lowest cost
- Shared agenda (scheduling meeting/visits), to let users share an agenda complementing the above services

8.2.4 Ambulatory Artificial Pancreas

The combination of insulin pumps and continuous glucose monitoring systems seems to be the best current solution to achieve good metabolic control for insulin-dependent diabetic people. Continuous glucose monitoring can detect glycemia patterns that cannot be discovered only by means of a few diary measurements.²⁴ The possibility of obtaining a complete blood glucose profile allows monitoring the suitability of an intensive therapy and can be used to fine-tune the therapy. The old concept of the "artificial pancreas" is starting to be a reality supported by the availability of these technologies and the integration of control systems able to close the loop modifying the pump parameters.²⁵

A closed-loop device capable of maintaining normoglycemia over extended periods of time could dramatically improve the quality of metabolic control of insulin-dependent diabetic patients.²⁶ An artificial pancreas contains three primary components: (1) an insulin pump, (2) a continuous glucose sensor, and (3) a closed-loop mathematical algorithm to regulate the pump given a sensor measurement.

Over the last decade, the development of an artificial pancreas has been a huge challenge to the application of biomedical technologies to diabetes therapy. The evolution of continuous glucose monitoring and insulin pump technologies is creating a very promising situation in the near future. However, the current reliability constraints of the continuous glucose sensor, the nonlinearity of the glucoregulatory system, and the inherent complexity of the design of a glucose controller for a subcutaneous–subcutaneous (SC–SC) setup are still some of the problems to be faced before obtaining a portable artificial pancreas.

8.3 Technical Requirements of Mobile Telemedicine Systems for Diabetes Care

We have seen some of the requirements of telemedicine and shared care systems for diabetes care. These requirements are general but for mobile systems there are some considerations and other requirements that we should consider.

The applications on a mobile device, PDA, or mobile telephone have to take into account the following considerations^{27–29}:

- *Concise content*: Only show relevant information using understandable abbreviations.
- *Current process recuperation*: The user's task could be interrupted by a phone call or a loss of coverage. The application must recover the process after the interruption.
- *Writing options*: Include all predictable texts in order to avoid user inserting it because of the text input limitations of mobile devices.
- *Navigation*: It is recommended that the application has a lot of categories with low depth. The information introduced should be stored in order to guarantee backward navigation.
- *Reduce the number of keystrokes*: Especially in mobile telephones where the text input is made with several keystrokes for each letter.
- *Velocity*: When the application makes use of the Internet it has to be considered that the user pays for time or transmitted bytes, so the connection should be fast enough.
- *Device limitations*: Mobile devices have limitations of battery, velocity of process, storage capacity, and graphics. These limitations must be considered at the design phase.
- *Easy-to-use interface*: The system should provide simple interfaces.
- *Media adaptability*: The system should provide support for different data, i.e., images, video, and text.
- *Modular design*: The system should have a modular design so that it allows for the development of a road map for growth that can accommodate future generations of functionality.

The implementation of an ambulatory artificial pancreas requires additional features:

- *Always-on technology:* The closed-loop algorithms have to run 24 hours a day, reading continuous glucose devices and operating insulin pumps in real-time. Robust devices and long-duration batteries are crucial for this critical task.
- *Services for remote supervision:* PDA applications usually work as client processes unable to provide information to external agents. The artificial pancreas application requires server functionalities (e.g., RPC or distributed CORBA objects) to allow remote supervision and even remote operation of medical devices.

Moreover, any mobile telemedicine application must fulfill some other requirements:

- *Security:* Health care data is protected by official privacy laws and data protection. Therefore mobile telemedicine applications must be very careful with security at all levels, i.e., communication, storage.
- *Communication with medical devices:* This communication can be wireless (i.e., Bluetooth, IrDA, WiFi) or by cable. In any case the security and integrity of the data transferred from and to the medical device should be guaranteed.
- *Controllability:* The system should support remote control functions. The health care provider should be empowered with the ability to control the media content according to medical specialties or his or her personal preference. For example, the health care provider is allowed to control ECG sample rate, video frame rate, and image quality, etc.

8.4 Building the Mobile Telemedicine System

Actually there are different mobile telemedicine applications for diabetes care. Some mobile companies are commercializing mobile solutions that allow patients to download data to a mobile telephone from a glucometer using a cable,³⁰ via Bluetooth,³¹ or IrDA.³² Besides, some other systems download data from the glucometer to a private network via a long-range wireless connection.³³

Patients and doctors can access the data in the network using a Web-based application or a mobile telephone application. Some applications provide more facilities like manual data recollection or graphic visualization of previous data in addition to data analysis and alarm generation.³⁴

There are some noncommercial experiences that allow glucometer data downloading.^{35,36} Another mobile application for diabetes care uses a mobile telephone to develop text messages with specific behavioral health strategies to young diabetes patients.³⁷

In this section we discuss the DIABTel telemedicine system for type 1 diabetic patients developed by the authors. DIABTel is a multi-access mobile telemedicine workspace (Figure 8.1), implemented over the sustenance of a multi-agent architecture. This workspace provides universal access to a wide range of services both for patients and professionals.

8.4.1 The DIABTel Distributed Architecture

The complexity and flexibility required to build the mobile workspace is achieved with the definition of a middleware multi-agent architecture that comprises a full range of nonexpensive and widely accepted information technologies offering users a universal, easy-to-use, online, and cost-effective access to telemedicine and information services. This architecture allows users to access the information from whatever access terminal or combination of terminals they choose. It facilitates the use of heterogeneous software and/or hardware solutions making possible that each health care organization configures the number and the kind of services they want to offer to their own users. Moreover, it creates a continuous record of the users' actions to monitor all the collaboration processes involved in shared care to improve the quality of care.

This integrated platform embeds several technical implementations to allow the access to a wider range of services both for patients and for professionals: consultation of patients' records, text and voice mailing, printing of reports, alarm management, visit management, telecare, tele-education, and intelligent therapy advisement.

The architecture works on three components: the agents, the multi-access organizer, and the database. The agents collaborate to guarantee homogeneous access of the users to the system services allowing users to access the information whatever access terminal or combination of terminals they choose. The number of agents is not limited initially and it only depends on the kind of terminals that will be used at each site.

There are two types of agents:

1. Communication server agents are in charge of communications with the different user terminals. Each communication agent is in charge of managing the communication process between the system and one specific user terminal. Their main functionality is to receive data from users updating it into the database and to retrieve data from the database to be presented at the user's terminal. They have the responsibility of performing the security policies for user access control, data confidentiality, and data integrity during data transfers.
2. Application server agents are in charge of data analysis and data processing.

The multi-access organizer is a message-oriented middleware module that is in charge of coordinating the interoperability between all the agents integrated into the architecture. The communication between the organizer and the agents is done using “event messages” encapsulated into TCP/IP messages. The organizer receives a message from each agent whenever an event occurs and activates the next action to be done. Additionally the organizer includes event registry services that enable the establishment of an event registry and event monitoring services, and smart routing services that ensure the message gets delivered to the appropriate recipients in the correct sequence.

The number and nature of the events is not fixed and can be defined at any system implementation according to their preferences or specific care protocols. During the exploitation of the system, the multi-agent architecture allows the inclusion of new events without affecting previous service performance. Likewise, the action the organizer triggers when a certain event is received is fully configurable.

The benefit of the multi-access multi-agent architecture is that it isolates the different communication processes from the rest of the system, so if some communication server needs to change its communication protocol, the system is not affected and only this server has to be modified. The definition of independent agents allows each site to have its own preferred communication servers working and, even more, the addition of new communication servers is transparent to the existing ones.

8.4.2 Mobile Applications in DIABTel System

Users can interact with the common workspace using several terminals such as Web, PDAs, and mobile phones. The user applications running on these terminals have a set of common functions both for patients and the health care professionals that enable the provision of the diabetes management services described above.

The main access method to reach the information in the multi-access system is the Web graphical interface, offering complete functionality of the telemedicine/shared care services with minimal human factors limitations. Through the DIABTel Web access, the users, patients, and doctors can consult the electronic patient logbook, insert new data, consult or create new treatments, access cooperative electronic mail, download data from a glucometer, and see graphical information such as clinical and system use information.

The “mobile applications” implements some of these functionalities, adapting them to the limitations of the mobile devices and following the considerations exposed in Section 8.3.

8.4.2.1 PDA Smart Assistant Application

The Smart Assistant (SA) application configures one of the most advanced approaches to mobile telemedicine close-loop systems in diabetes management.³⁸ It provides patients with several closed-loop control strategies (personal and remote) that offer an augmented information, self-management environment to increase patient empowerment. The SA has an independent executable algorithm implemented, based on a nonlinear MPC with Bayesian learning. The algorithm allows real-time control of the insulin pump based on glucose data, meal intake, and patient's personal data. The algorithm can work in a continuous mode, providing calculation for basal insulin rates every 15 minutes.

The PDA application is a portable and mobile solution which provides the patient with flexible tools for data visualization and with local and remote communication services. The Smart Assistant is able to manage a local database, allowing the patient to work isolated from the telemedicine server without the need of a permanent active communication link. On demand by the user, the patient unit synchronizes the local database with the server database, causing the automatic updating of all the new information in both directions.

The most characteristic functionality developed for the PDA application is the “patient electronic logbook” (Figure 8.2), where the patient could perform the following tasks:

- *Manage his or her monitoring data.* It includes visualization and edition of the patient logbook. It is possible to select the period or the single time slice for the data to be introduced or visualized. The patient can visualize or insert meal data, glucose data, insulin data, device events (needle change, catheter



Figure 8.2 The logbook and pump communication scenarios for the PDA application and iPAQ h2210 PDA.

change, etc.), in individual activities (exercise), additional data (medication, stress, etc.)

- *Visualize graphics and statistical data analyses.* Three types of graphics are available: plot, pie, or temporal.
- *Consult his or her active therapy.* The last therapy (insulin and diet) proposed by the physician is updated in the application after each data synchronization.
- *Download data from the insulin pump or the glucometer.* Communication with the insulin pump is wireless (Bluetooth or IrDA); a serial cable is used to communicate with the glucometer.

The PDA application can be configured by the patient. Some of the options of configuration are preferred language, preferred blood glucose units (mg/dl or mmol/l) for the visualization scenarios, and preferred starting hour for the logbook representation. It is also possible to enter additional patient information (weight, date of birth, age, birth date, etc.). It also allows introducing authentication data of the patient that uses the application, such as login and password and modifying the configuration settings for each patient.

The PDA application is developed in Java in order to be portable on multiple platforms. The portability of the application has been successfully tested in conventional PCs and in the iPAQ hp2210 and TSM500 PDAs, both with Windows CE operating system. It is necessary to install CrEme v3.24 JVM technology for embedded platforms in order to run the application. The TSM500 has GPRS mobile connection integrated and for the iPAQ the AudioVox RTM 8000 card was used to provide the GPRS communication capabilities.

The Smart Assistant was evaluated in a clinical experiment and demonstrated an improvement in the glycemic control in pump-treated patients with type 1 diabetes.³⁹

8.4.2.2 WebPDA

The WebPDA is a Web application that was specifically designed taking into account PDA limitations. The DIABTel Web application is optimized for a high-resolution display (1024 × 768) in order to present as much information as possible. When accessing with a lower-resolution device, scrolling is allowed but the PDA resolution is much lower, so much scrolling is necessary to access Web applications. The “y” dimension in PC screens is bigger than the “x” dimension but the opposite is true in PDAs, so a Web application could not be seen properly in a PDA screen (Figure 8.3).

The WebPDA can be considered a lighter version of the Web application. The functionalities of the WebPDA application are

- *Doctors:* Patient selection, creation and visualization of treatments, electronic logbook visualization, visualization and creation of electronic mails, visualization, and creation of news.



Figure 8.3 The WebPDA application.

- *Patients*: Electronic logbook visualization, visualization of treatments, visualization of news, visualization and creation of electronic mails.

8.4.2.3 DIABTelMobile

DIABTelMobile is a mobile telephone application. The functionalities are introduction, storing, sending, and visualization of patient monitoring data, visualization of treatment, transmission and reception of messages, and data downloading from glucometer via Bluetooth.

The DIABTelMobile (Figure 8.4) has been developed in Java Microedition, J2ME, using the CLDC 1.0 configuration and the MIDP 2.0 specification. Theoretically, it can work in any mobile telephone that implements the MIDP 2.0 specification and also the Bluetooth JSR82 specification. But the practice shows that not all mobile telephones that are supposed to implement the MIDP 2.0 and the JSR82 specifications completely make it. The application was tested in the TSM520 mobile telephone with a Windows mobile operating system.

8.5 Conclusions

Wireless technologies and the further integration of medical devices into mobile scenarios are the pillars for new mobile services that operate in the personal area network, optimizing patient intervention for data retrieval and communication.

Moving in this direction, mobile telemedicine increases the quality and the quantity of the information collected by patients, improving physician–patient communication and enhancing the decision-making process to achieve effective therapy adjustments in a collaborative and supervised way.



Figure 8.4 The DIABTelMobile application.

The implementation of these mobile services requires the definition of distributed architectures including strict strategies for security to guarantee patient privacy, data confidentiality and integrity, and auditing of the clinical acts are mandatory. The limitations of mobile devices, such as data processing and batteries duration, have to be taken into account in order to develop the users' applications.

The current challenge in mobile diabetes care is to combine several control strategies adapted to patient conditions and glycemic control state supported on communications and decision aid tools. The augmented reliability and availability of continuous glucose sensors, insulin pumps, mobile computing technologies, and telemedicine services is making the implementation of a feasible and reliable ambulatory closed-loop glucose control system a reality as the best current solution to achieve a good metabolic control for insulin-dependent diabetes mellitus patients. The research efforts on mobile technologies are moving in this direction looking for new solutions to better the quality of care and quality of life of people with diabetes.

Acknowledgments

This work has been partially funded by the Spanish FIS Projects PARIS (PI042466) and ADVISING (I060437). The authors wish to acknowledge the valuable work of all researchers from the Grupo de Bioingeniería y Telemedicina, specifically to V. Torralba.

References

1. International Diabetes Federation, *Diabetes Atlas*, Second Edition, Brussels, 2003.
2. W.K. Waldhäusl, Finally we have arrived in a new millennium, *Diabetologia*, 44, 1, 2001.
3. A. de Leiva, P. Lefèbvre, and J. Nerup, European dimension of diabetes research, *Diabetologia*, 39: 5–11, 1995.
4. World Health Organization, Fact Sheet #312, September 2006.
5. S. Andreassen, E.J. Gómez, and E.R. Carson, Computers in diabetes 2000, *Computer Methods and Programming in Biomedicine*, 62: 93–95, 2002.
6. D.G. Marrero, K.K. Kronz, and M.P. Golden, Clinical evaluation of a computer-assisted self-monitoring of the blood glucose, *Diabetes Care*, 12: 345–350, 1989.
7. K.K. Ahring, C. Joyce, J.P.K. Ahring, and N.R. Farid, Telephone modem access improves diabetes control in those with insulin-requiring diabetes, *Diabetes Care*, 15: 971–975, 1992.
8. A. Billiard, V. Rohmer, M.A. Roques, M.G. Joseph, S. Suraniti, P. Giraud, J.M. Limal, P. Fressinaud, and M. Marre, Telematic transmission of computerised blood glucose profiles for IDDM patients, *Diabetes Care*, 14: 130–134, 1991.
9. R. Bellazzi, C. Larizza, S. Montani, A. Riva, M. Stefanelli, G. d'Annunzio, R. Lorini, E.J. Gómez, E. Hernando, E. Brugués, J. Cermenon, R. Corcoy, A. de Leiva, C. Cobelli, G. Nucci, S. Del Prato, A. Maran, E. Kilkki, and J. Tuominen, A telemedicine support for diabetes management: The T-IDDM Project, *Computer Methods and Programming in Biomedicine*, 69: 147–162, 2002.
10. F. del Pozo, E.J. Gómez, and M.T. Arredondo, A telemedicine approach to diabetes management, *Diabetes, Nutrition & Metabolism*, 4(Suppl. 1): 149–153, 1991.
11. E.J. Gómez, F. del Pozo, and M.E. Hernando, Telemedicine for diabetes care: The DIAB-Tel approach towards diabetes telecare, *Medical Informatics*, 21(4): 283–295, 1996.
12. E.D. Lehmann, Application of information technology in clinical diabetes care. Part 1: Databases, algorithms and decision support, *Medical Informatics*, 21(4): 255–374, 1996.
13. M.W. Tsang, M. Mok, G. Kam, M. Jung, A. Tang, U. Chan, C.M. Chu, I. Li, and J. Chan, Improvement in diabetes control with a monitoring system based on a handheld, touch-screen electronic diary, *Journal of Telemedicine and Telecare*, 7(1), 2001.
14. M. Edmonds, M. Bauer, S. Osborn, H. Lutfiyya, J. Mahon, G. Doig, P. Grundy, C. Gittens, G. Molenkamp, and D. Fenlon, Using the Vista 350 telephone to communicate the results of home monitoring of diabetes mellitus to a central database and to provide feedback, *International Journal of Medical Informatics*, 51(2), 1998.
15. D.J. Nigrin and I.S. Kohane, Glucoweb: A Case Study of Secure, Remote Biomonitoring and Communication, *Proceedings of the AMIA 2000*, Los Angeles, November 2000.
16. D. Steinbock, *The Mobile Revolution: The Making of Worldwide Mobile Markets*, Kogan Page (Business and Management), 2005.
17. M.E. Hernando, E.J. Gomez, A. Garcia, and F. del Pozo, A multi-access server for the virtual management of diabetes, *Proceedings of ESEM 99* (fifth conference of the European Society for Engineering and Medicine), 309–310, 1999.
18. American Diabetes Association, 2007, <http://www.diabetes.org/>

19. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, *Diabetes Care*, 25(Suppl. 1): S5–S20, 2002.
20. The Diabetes Control and Complications Trial Research Group, The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus, *New England Journal of Medicine*, 329: 977–986, 1993.
21. American Diabetes Association, Implications of the diabetes control and complications trial, *Diabetes Care*, 26(Suppl. 1): S25–S27, 2003.
22. American Diabetes Association, Standards of medical care for patients with diabetes mellitus, *Diabetes Care*, 26(Suppl. 1), S33–S50, 2003.
23. A. García-Olaya, E.J. Gómez, M.E. Hernando, F. del Pozo, A middleware CSCW architecture for diabetes shared care, *IFMBE Proceedings of EMBEC'02*, 3(2): 1376–1377, 2002.
24. B.W. Bode, T.M. Gross, K.R. Thornton, and J. Mastrototaro, Continuous glucose monitoring used to adjust diabetes therapy improves glycosylated hemoglobin: A pilot study, *Diabetes Research & Clinical Practice*, 46: 183–90, 1999.
25. R. Bellazzi, G. Nucci, and C. Cobelli, The subcutaneous route to insulin-dependent diabetes therapy, *IEEE Engineering in Medicine and Biology*, 20(1): 54–64, 2001.
26. R.S. Parker, F.S. Doyle, and N. Peppas, The intravenous route to blood glucose control: A review of control algorithms for noninvasive monitoring and regulation in type 1 diabetic patients, *IEEE Engineering in Medicine and Biology*, 20(1): 65–73, 2001.
27. S.S. Chan, X. Fang, J. Brzezinski, Y. Zhou, S. Xu, and J. Lam, Usability for mobile commerce across multiple form factors, *Journal of Electronic Commerce Research*, 3, 3, 187, 2002.
28. J. Hobart, Designing Mobile Applications: Principles and Guidelines for Successful UI Design, <http://www.classicsys.com/>
29. Y. Chu and A. Ganz, Mobile telemedicine systems using 3G wireless networks, *Business Briefings: US Healthcare Strategies*, 2005.
30. MedicalGuard, <https://www.medicalguard.net/modules.php?name=whereis>
31. Think Positive Diabetes, <https://www.thinkdiabetes.com/tpdiabetes/>
32. Emminems, <http://www.emminens.com/>
33. GlucoMON, <http://www.diabetech.net/glucomon.html>
34. SiDiary, <http://www.sidiary.org/>
35. D. Gammon, E. Arsand, O.A. Walseth, N. Adersson, M. Jenssen, and T. Taylor, Parent–child interaction using a mobile and wireless system for blood glucose monitoring, *Journal of Medical Internet Research*, 7, 5, 57, 2005.
36. A.J. Farmer, O.J. Gibson, C. Dudley, K. Bryden, P.M. Hayton, L. Tarassenko, and A. Neil, A randomized controlled trial of the effect of real-time telemedicine support on glycemic control in young adults with type 1 diabetes, *Diabetes Care*, 28, 11, 2697–2702, 2005.
37. V. Franklin, A. Waller, C. Pagliari, and S. Green, “Sweet talk”: Text messaging support for intensive insulin therapy for young people with diabetes, *Diabetes Technology and Therapeutics*, 56: 991–996, 2003.

38. E.J. Gómez, M.E. Hernando, T. Vering, M. Rigla, O. Bott, G. García-Sáez, P. Pletschner, E. Brugués, O. Schnell, C. Patte, J. Bergmann, R. Dudde, and A. de Leiva, The INCA System: A further step towards a telemedical artificial pancreas, *IEEE Transactions on Technology in Medicine*, 13: 24, 2007.
39. M. Rigla, M.E. Hernando, E.J. Gómez, E. Brugués, G. García-Sáez, V. Torralba, A. Prados, L. Erdozain, J. Vilaverde, and A. de Leiva, A telemedicine system that includes a personal assistant improves glycemic control in pump treated patients with type 1 diabetes, *Journal of Diabetes Science and Technology*, 1: 4, 2007.

Chapter 9

Telemedicine: A Way to Improve Glycemic Control among Elderly Diabetics

Sheila Black

CONTENTS

9.1 Introduction.....	162
9.2 Technological Reminders and Sensors	163
9.3 Cognitive Orthotics	164
9.4 Virtual Medical Offices.....	167
References	171

This contribution discusses the efficacy of telemedicine for elderly diabetics. The bulk of the evidence indicates that telemedicine is effective in improving glycemic control among older diabetics and in reducing the medical complications associated with diabetes. Furthermore, most studies indicate that older diabetics report high satisfaction with various aspects of telemedicine.

9.1 Introduction

Diabetes mellitus has wide ranging consequences including amputations, blindness, and renal failure, etc. In an effort to decrease the untoward medical complications associated with diabetes, recently, the Diabetes Control and Complications Trial Research Group (DCCT) (1993) has emphasized tight control through patient self-management.¹ According to experts, diabetics who are able to maintain tight glycemic control gain an extra five years of life, an extra six years free of kidney damage, an extra eight years of sight, and an extra six years free of nerve damage.² Tight control as defined by the American Diabetic Association refers to hemoglobin A1c levels below 7%. To explain further, hemoglobin A1c levels measure the proportion of glycated hemoglobin cells. A cell is glycated if glucose is attached to it. As individuals process sugar, glucose bonds with hemoglobin and it does so in proportion to the glucose in the bloodstream. It takes two to three months for sugar to be removed from these cells; thus, health care professionals reason that A1c levels provide a good estimate of the amount of sugar that has remained in the bloodstream over a three-month period.

Maintaining tight control of diabetes is not easy. It requires that patients keep track of their food and insulin intake to avoid hypo- or hyperglycemia. Hyperglycemia (i.e., high blood sugar) can lead to diabetes-related complications described earlier, primarily because diabetes damages blood vessels, which ultimately leads to the constriction of blood flow. These macro- and micro-vascular complications often result in heart damage, kidney damage, neuropathy, etc. Hypoglycemia (i.e., low blood sugar), if severe enough and not corrected, can lead to a coma or even death.³⁻⁴

Older adults are particularly susceptible to diabetes; in fact, close to 20% of older adults suffer from diabetes.⁵ Many of these older adults have difficulty achieving tight control because of the high degree of cognitive resources needed to manage diabetes. For example, managing diabetes requires that patients use good problem-solving strategies and that they remember to balance their glucose and insulin intake on a given day. Diabetics also need to remember to check their glucose levels and the nutritional value of various foods. There are age-related changes in both memory retrieval and in problem-solving skills.² Telemedicine might be a way to compensate for such declines.

Telemedicine includes using any form of technology that allows health care professionals to communicate with patients, including computers, telephones, electronic handheld devices, etc. Often telemedicine is divided into three components: (1) synchronized videoconferencing, (2) remote monitoring, and (3) education through Websites.⁶ Cognitive aids can also serve as reminders to diabetics that fall under the rubric "telemedicine."

This chapter will focus on various forms of telemedicine including technological reminders for the healthy older adult and for the memory impaired. These

technological advances include sensors to monitor cognitively impaired older adults, and virtual offices through sophisticated computer units. The aforementioned technological advances are ideal for diabetics because diabetics have complex medical regimens to follow and often forget to perform tasks required to manage diabetes adequately. As well, diabetes accelerates the aging of the brain. Therefore, older diabetics probably have more cognitive impairments than the typical older adult⁷ and could benefit from the extra cognitive support provided through telemedicine.

9.2 Technological Reminders and Sensors

Technological cues and reminders are particularly apt for older adults suffering from memory impairment, but these technological reminders are also appropriate for older adults without memory impairment.

A number of studies have provided evidence that medication compliance is particularly problematic for older adults.^{2,8} Health care professionals have tried to improve medication compliance by providing educational programs for diabetics; however, educational programs have not been successful in increasing compliance rates.⁷

The most effective intervention to increase compliance rates appears to be memory aids or cues.⁷ Several studies have in fact investigated the efficacy of telephone reminders. For example, Fulmer et al.⁷ conducted a study to examine the efficacy of telephone reminders versus education in increasing compliance and to determine if certain memory aids were more efficacious than others. The target participants for this study were frail individuals over 65 with multiple health problems. Fulmer et al. assigned participants to one of three conditions: videophone, telephone, and control. The videophone group received audio-video reminders to take their medication and the telephone group received telephone reminders. The control group received the usual information presented in the doctor's office.

The Medication Event Monitoring System caps (MEMs) was used to assess compliance. MEMs are computerized caps placed on medication bottles to record removal of the cap from the bottle. Interestingly, all participants took their medication over 80% of the time initially; however, over time, compliance dropped drastically for the control group but remained high in the two intervention groups. Thus, there was a time × group interaction. Surprisingly, there was no advantage of a videophone over a telephone. This study indicates that even fairly simple interventions such as telephone call reminders increase compliance among older adults. Other studies provide evidence that fairly simple reminders increase compliance. For example, a study conducted by LaVigne and Tapper⁸ found that automatized voice interactive systems that reminded patients to take their medication increased compliance among the individuals who received

automated messages relative to the control group. The fact that automated voice reminders can be effective in improving compliance is important. Health care professionals may not have the time or staff available to make personal calls to patients. Nevertheless, they might be able to provide support by sending out automated voice messages.

In another study, Piette and Mah⁹ found that 98% of diabetic patients found automated voice messages to be helpful. In the Piette and Mah study, upon receiving an automated call, participants answered questions about glycemic control and diabetic complications by pressing various number options on a Touch-Tone phone. If a patient indicated that she was experiencing a complication (e.g., foot ulcers or chest pains), then the patient received a call from a human professional. Patients also received follow-up calls if they reported a glucose reading above 250. This was important because patients could receive personal attention immediately when there was evidence that patients were exhibiting poor glycemic control.

As indicated earlier, older adults often have difficulty with the problem-solving aspect of diabetes.^{10,11} Telephone reminders can provide solutions in “real-time.” For example, Long et al.¹² conducted a study examining the efficacy of providing problem-solving support for older diabetics over the phone. The number of calls that a patient received was determined by the patient’s HbA1c levels. Each call was 20 minutes long and patients could query the health care professional about illness and blood sugar levels or the effect of exercise on blood sugar, etc. The patients were overwhelmingly pleased with the service.

Telephone reminders can be very effective in helping diabetics to remember to take their medication and in providing assistance with everyday problem-solving dilemmas. However, there are groups of diabetics who require more than daily reminders in order to manage their diabetes. Those groups of diabetics may benefit from cognitive orthotics and/or sensors.

9.3 Cognitive Orthotics

A cognitive orthotic is a computerized memory aid that is often used for people who have difficulty remembering to perform activities of daily living.^{13,14} It is particularly useful with respect to prospective memory, or remembering to do something. There are age-related changes in prospective memory and there is evidence that older diabetics have more difficulty with this type of memory than nondiabetics.¹⁵

One of the most difficult aspects of diabetes is balancing food and insulin intake. Many dietitians recommend that diabetics count carbohydrates and restrict fat. However, many older adults are unfamiliar with the nutritional value of food. One research group¹⁶ examined the likelihood that an electronic aid would assist patients in maintaining their diet. There was an experimental group that received the handheld dietary device for 12 weeks and a control group that did not receive

such a device. The experimental group was provided with an electronic diary (i.e., hospital-based monitoring system [DMS]) that would immediately access the nutritional value of food, including calories, carbohydrates, protein, etc., of each item in the database. Patients in the experimental group actually reduced their HbA1c levels from 8.56 to 7.55 as a result of this monitoring system.

More recently, Agarwala, Greenberg, and Ho¹⁷ have developed a new innovative orthotic, a context-aware pill bottle that allows both clients and caregivers to monitor medication management. The device consists of a pill bottle stand that is connected to a computer. The device audibly reminds elderly individuals to take their medication. The plastic pill bottles are augmented with a radio frequency identification (RFID) tag that provides information about when the pill should be taken and possible side effects.

This device provides reminders to patients in an interesting way. If the client does not move the pill box at the assigned time, the patient receives increasingly obtrusive alerts. A medication monitor, which informs the caregiver when the pill bottle has been removed from the stand, is placed in the caregiver's home.

If the client fails to take the medication despite the increasingly obtrusive alerts, then the caregiver is contacted. The caregiver is alerted by the blinking of the medication monitor device. To stop the blinking, the caregiver must touch the device and can act upon the alert by sending a message that will be delivered back to the client. Obviously, this reminder could greatly assist in reminding diabetic patients to take their medication, check their blood sugar, etc.

Rhodes¹⁸ has also designed a device that would be helpful for diabetics—a remembrance agent, basically an electronic reminder. A remembrance agent would be helpful for diabetics because of their long list of daily prospective memory tasks such as remembering to monitor glucose levels, to take medications, to exercise, etc. This remembrance agent resembles a personal digital assistant (PDA), the popular device that provides reminder aids and schedules for busy professionals. However, the device is different from a PDA in that it is equipped with sensors that determine the location of the user. The sensors provide notes and reminders to the user based on the location. Thus, if the client is in the supermarket, then the remembrance agent lists appropriate grocery items to buy. However, if the client is in the kitchen, then the remembrance agent provides reminders appropriate for that location (e.g., reminders about cooking).

The devices discussed heretofore are available for clients who may need reminders in terms of their medication, but who are for the most part mentally and physically healthy enough to live independently without much assistance. However, a number of older diabetics need more than fairly simple cognitive orthotics to live independently. These groups of older adults need fairly intensive monitoring. This next subsection focuses on new devices that allow continual, 24-hour monitoring. These devices are ideal for older adults who have severe enough memory impairments that their families worry about their ability to live independently but do not think that the severity of the impairment warrants placement in a health care facility.

Examples of such devices are activity monitoring systems. These systems track the movements of an older person and even send alerts if anything is awry (e.g., a fall). These systems rely on motion detectors mounted in a nonobscure location in the house (e.g., on a ceiling). Interestingly, these motion detectors use a computer algorithm to track and ultimately learn a person's usual behavior. After learning the person's behaviors, the computer sends an alert when the person deviates from expected behavior. Thus, the monitoring system can send an alert to a caregiver if a person is in the bathroom too long, etc.¹⁹

One study²⁰ examined how such sensors would actually operate in the home. This study took place in the "smart house" that was created by the Medical Automation Research Center at the University of Virginia as a prototype of living environments for individuals who are physically and/or mentally impaired. The sensors were placed in every room and responded to any movement; each time the sensor responded, the date, time, and duration of the sensor activation was forwarded to health care professionals and caregivers via a computer. After examining the pattern of sensor activation over a number of days, investigators were able to make reasonable predictions about regular activities such as cooking and eating dinner based on the pattern of sensor activation.

The fact that investigators can infer the activity of clients in "smart homes" is important. One of the pitfalls of diabetes is hypo- or hyperglycemia, which could render a patient delirious. If a patient deviated too much from her daily activities, concerned caregivers could be alerted. More recently, Yang and Rhee²¹ developed a sensor that could be worn as a ring. This "ring" measures oxygenated blood flow continuously and the signals are sent to a computer for observation. Along the same line, Gatzoulis and Lymberis²² have developed unobtrusive sensors that can be worn by individuals who require constant monitoring. One of the more innovative systems includes wireless garments embedded with textile sensors that continuously monitor biomedical signs such as respiration. Again, all of the medical information gleaned via the wearable sensors is forwarded back to health care professionals or caregivers.

One of the most sophisticated sensor/automated reminding devices is PEARL, an automatic robot that can be used by frail and cognitively impaired individuals.^{13,14,23,24} Pearl was recently constructed via a multidisciplinary team project. The researchers involved with this project wanted to create a robot that would assist elderly individuals who were in their home. PEARL is equipped with two Pentium computers, sonar sensors, microphones, a speech recognition system, and stereo camera systems. Most importantly, PEARL is equipped with the Autominder software that provides reminders to cognitively impaired elderly individuals.

Autominder provides reminders to people about their daily activities. There have been reminder systems since the 1960s.²³ However, the new Autominder system takes advantage of artificial intelligence (AI). It models the client's daily activities such that it becomes aware of the length of time that clients spend on various activities and it becomes aware of the sequence of activities. Autominder has three components: a plan manager (which stores a client's scheduled activities), a client

modeler (uses information about a client's observable behavior to update his or her schedule), and personal cognitive orthotics (reminds patients about what they are supposed to do). Due to the artificial intelligence component, Autominder can be programmed to estimate how long activities should take. Reminders are generated if an activity takes longer than normal or if an individual is not in a room where the activity usually takes place. Sensors for Autominder are deployed via PEARL. Sensor information can be sent to the caregiver throughout the day.

The reminders have to be initiated by the caregiver; however, the artificial intelligence allows Autominder to make intelligent judgments about when a reminder is needed. Thus, if an elderly person has to be reminded to go to the bathroom every three hours and, based on the client's regular routine, the system "knows" the individual watches a two-hour television program on a particular day, the system might remind the client before the television program begins and thereby avoid interrupting the client's television program. Thus, due to artificial intelligence, autominder can issue reminders at a time most convenient for the client. PEARL was piloted at Longwood Health Care Center¹³ and was a very valuable addition to the facility.

Thus, the bulk of the evidence indicates that cognitive orthotics can be very helpful for older diabetics who are mildly forgetful or very impaired and actually require constant monitoring through technology. However, what about diabetics who may be cognitively intact but have difficulty traveling to a doctor's office or have a very complex medical regimen because of multiple health problems? Is there a way that technology could bring the medical office to their home?

9.4 Virtual Medical Offices

A number of complex technological systems are currently being tested with older adults. Initially, there was some concern about designing complex technology for the elderly because the elderly were viewed as being less technological savvy than younger adults. Furthermore, due to age-related changes in working memory, and processing speed,^{25,26} older adults grasp information at a slower rate than younger adults. In fact, as Stronge, Rogers, and Fisk²⁷ point out, some age-related changes in motor dexterity might impede older adults' ability to utilize the technology associated with complex telemedicine systems.

In fact, Kaufman et al.²⁸ reviewed a complex telemedicine program that involved creating a virtual office in the patient's home. Their main concern was the extent to which elderly patients would adapt to sophisticated new computer technology. They examined the usability of this telemedicine system in 25 diabetic patients' homes and discovered that successful usage of the program required the patient to complete 15 subtasks. They found that participants could readily use the glucometers attached to the unit and enjoyed the video visits with health care professionals. However, participants had some difficulties due in part to age-related changes in cognition and

lack of experience with computers. For example, the participants had difficulty using the mouse (perceptual–motoric skills), understanding how the entire system worked (mental models) and health literacy (which goes back to age-related changes in reading comprehension). To explain further, several of the older adults were hampered by health literacy because they did not understand the significance of blood pressure or glucometer readings. Also, with regard to mental models, older adults had difficulty understanding how the Internet worked (e.g., the user needs to click hyperlinked words to find out more information about a particular topic).

Thus, the Kaufman et al.²⁸ study indicates that telemedicine affords older diabetics the opportunity to become more proficient in managing their disease. However, the study also indicates that a minimal amount of proficiency may be necessary for patients to take full advantage of telemedicine.

Studies have been conducted in which memory-impaired younger adults received medical care through telemedicine. These studies indicate that participants with memory impairments relative to healthy younger adults²⁹ were capable of mastering the complex technology needed to access a “virtual office.” However these individuals require more training than a typical younger adult. Thus, it is probably true that older adults (who as a group learn at a slower rate than younger adults) will require significantly more training to master the technology associated with “virtual offices.” However, studies indicate that with the proper training, older adults can master the technology associated with complex telemedicine units.³⁰

Probably the most publicized telemedicine study focusing on telemedicine for older adults was the one conducted by Columbia University, titled the Informatics for Diabetes Education and Telemedicine Project (IDEAT). The investigators in the IDEAT project tried to improve glycemic control among a wide cross section of diabetics by designing a very user-friendly computer/electronic system that included all three facets of telemedicine.

All of the diabetics were Medicare recipients who lived in underserved areas, with some of the recipients living in poor areas of New York City and some recipients living in rural counties within New York state.³¹ The focal point of this intervention was the home telemedicine unit (HTU). The inventors of this product wanted to integrate all of the components so that elderly patients, with no computer experience, would have no difficulty using the technology. The patients are able to operate the computer for the most part without using a mouse or a keyboard. Instead, they had “four launch buttons”: answer video call, send data from glucose and blood pressure meters, connect to Website, and reboot computer. In fact, upon turning on the computer, the patient was immediately connected to the Web. The installation of the computer and the training was performed by home care nurses. The home care nurse received a constant stream of data informing them of patients’ health status. The nurses had guidelines to follow in terms of reporting “worrisome values” to doctors and other health care specialists.

Recently, several studies have examined the effectiveness of the IDEAT program in helping diabetics achieve glycemic control. One of the primary reasons

that such a program should improve glycemic control is that it reduces the extent to which diabetics have to rely on episodic memory (i.e., memory for episodes in one's life) and engage in problem solving. If they have a question that involves problem solving, they could consult one of the professionals through the IDEAT program.

One of the first studies to evaluate the IDEAT program was conducted by Shea et al.³⁰ In this study, participants were either in the usual care condition or involved in the IDEAT program described earlier. Patients in the usual care condition received information about diabetes, similar to the educational information given to diabetics in doctors' offices. As indicated earlier, those in the IDEAT program consulted with health care professionals and their A1c levels were forwarded to health care professionals on a regular basis. Those in the IDEAT program reduced their cholesterol, A1c levels, and lipid levels significantly more than those in the control group. Specifically, A1c levels decreased from 7.35 to 6.97% in the intervention group. This study indicated that even individuals who were controlling their diabetes fairly well could improve glycemic control with the extra support provided by the IDEAT program.

Most of the work investigating telemedicine among the elderly has not been as sophisticated as the IDEAT project at Columbia University. For example, several VA hospitals received a grant to implement a telemedicine program called the Care Coordination approach. In the Care Coordination approach, the nurse coordinator maintains frequent contact with a patient through telemedicine and feeds information back to a doctor. One of the reasons that health care providers have been interested in this approach is that it might actually prevent complications from complex diseases such as diabetes. In one study,³² conducted by Chumbler et al., participants were older veteran diabetics who were already experiencing health problems in that they had been to the emergency room or had been admitted to the hospital twice within 12 months. The diabetics were divided into two groups with one group receiving intense monitoring on a weekly basis and the other group receiving less intense monitoring on a daily basis.

Patients and caregivers in the weekly monitoring condition were trained to use an instant camera to take snapshots of wounds; the pictures were mailed weekly to the care coordinator. Upon receiving the pictures, the care coordinator would decide if further treatment was necessary. Patients were also followed through the diabetic clinic.

The daily monitoring group³² was not followed by a diabetic clinic nor was this group required to send pictures of wounds. However, this group received general information about managing diabetes via a home messaging system. They were equipped with a telemonitoring system that allowed weekly glucose monitoring and that provided two-way audio-video connectivity. However, the main difference between the daily monitoring group and the weekly monitoring group was that patients in this group answered questions daily about their glycemic control and individuals in the weekly monitoring condition did not. The answers to the questions were forwarded to their care coordinator. The questions were designed to assess the patient's knowledge, symptoms, and behavior in regard to chronic disease

management. The care coordinator or health buddy reviewed the answers to the questions and made decisions about the necessity of further intervention.

The investigators found that hospital admission rates were 52% lower in the daily versus weekly monitoring group. The diabetes-related emergency room visits were decreased 15%. Interestingly, even though there was a decline in hospitalizations among the daily monitoring group relative to the weekly monitoring group, there was no evidence that glycemic control was better in the daily monitoring group. That is, the hemoglobin A1c levels were not better from pre-post intervention among individuals in the daily monitoring group relative to the weekly monitoring group. Barnett et al.³³ also found that diabetics who were monitored daily were hospitalized less than controls (matched for severity of diabetic symptoms).

Mease² conducted a study using technology and environmental support similar to that used in the IDEAT program. Mease and colleagues used equipment approved by the Federal Drug Administration for the monitoring of diabetes, the Avia 20/20 and the Avia 10/10. Only diabetics with hemoglobin A1c levels above 8% were included in the study. It should be noted that individuals who have levels above 8% are significantly more likely to experience complications than diabetics who have A1c levels in the 6 to 8% range. It should also be noted that Mease excluded individuals who were not able to use the equipment properly after training and individuals who had a psychiatric history. The average age of participants in this study was 61.

On a weekly basis the patients in the treatment group received telemonitoring visits from the case manager, which included both voice and video interaction. The case manager counseled the patients about their nutritional intake and medication compliance. The case manager also reviewed each patient's well-being and glycemic episodes each week.

Patients in the telemedicine program also used technology to maintain contact with their physicians. In fact, the two physicians visited the patients once a month through telemonitoring. The outcome measures included the HbA1c values, microalbumin, creatinine, and lipid panel. The results indicated that individuals who received the home telemedicine system reduced A1c levels by 16%, and their weight by 4%. The reduction in A1c levels was significant in the telemedicine group. Although the control group reduced A1c levels, the reduction was not significant.

One of the reasons that the Mease² study and the Shea et al.³⁰ study were successful in lowering A1c levels might be that these interventions provided regular feedback and counseling for patients about diet and management of blood sugar levels. The Chumbler et al.³² study appeared to focus more on serious complications rather than day-to-day management of diabetes.

The consensus from all of the data is that telemedicine can definitely improve the medication management and quality of life of diabetics. Interestingly, the Chumbler et al.³² and the Shea et al.³⁰ studies indicate that older adults can even take advantage of fairly complex telemedicine systems. However, it is crucial that designers of such systems be aware of age-related changes in learning new information and cohort differences in experience with technology.

There are a number of advantages of using telemedicine to improve the quality of life of older diabetics and older adults in general. With advances in computer technology, concerned caregivers and health care professionals can now keep track of an elderly person's vital signs and ability to maintain glycemic control. Hopefully, telemedicine will become more widespread in the near future. If telemedicine were readily available, older adults would be able to live in their own home longer, which would improve their quality of life. A number of studies indicate that older adults fare better emotionally when allowed to age in place. Furthermore, telemedicine would allow doctors in rural areas to obtain information via teleconferences, Email, etc., that otherwise would be inaccessible to them. Telemedicine appears to be the wave of the future. It has the capacity to greatly enrich the lives of older diabetics.

References

1. Y.M. Po, *Journal of Telemedicine and Telecare*, 263, 2000.
2. A. Mease, *Military Medicine*, 579, 2000.
3. S. Black and F. Scogin, *Educational Gerontology*, 553, 1998.
4. C. Blaum, M. Ofstedal, K. Langa, and L. Wray, *Journal of the American Geriatrics Society*, 745, 2003.
5. M. Williams, J. Lacson, M. Teng, N. Ofsthun, and J. Lazarus, *Kidney International*, 1503, 2006.
6. R.L. Reed and A.D. Mooradian, *American Family Physician*, 915, 1991.
7. T.T. Fulmer, P.H. Feldman, T.S. Kim, B. Carty, M. Beers, M. Molina, and M. Putnam, *Journal of Gerontological Nursing*, 6, 1999.
8. M. LaVigne and K.A. Tapper, *Disease Management and Health Outcomes*, 1, 1999.
9. J.D. Piette and C.A. Mah, *Diabetes Care*, 15, 1997.
10. R.E. Glasgow, J. Mullan, L. Fisher, D.J. Toobert, and M. Skaff, *Diabetes Care*, 33, 2007.
11. F. Hills-Briggs, *Annals of Behavioral Medicine*, 182, 2003.
12. A.F. Long, J. Taylor, T. Gambling, J.M. Mason, and R.J. Young, *Diabetes Care*, 283, 2005.
13. M.E. Pollack, Planning Technology for Intelligent Cognitive Orthotics, retrieved December 27, 2006, from <http://www.cs.cmu.edu/~flo/papers/umich/AIPS-02Pollack.pdf>
14. M.E. Pollack, L. Brown, D. Colbry, C.E. McCarthy, C. Oroz, B. Peintner, S. Ramakrishnan, and I. Tsamardinos, *Robotics and Autonomous Systems*, 273, 2003.
15. F. Grodstein, J. Chen, R.S. Wilson, and J.E. Manson, *Diabetes Care*, 24, 1060, 2001.
16. M.W. Tsang, M. Mok, G. Kam, M. Jung, A. Tang, U. Chan, C.M. Chu, I. Li, and J. Chan, *Journal of Telemedicine and Telecare*, 47, 2001.
17. A. Agarwala, S. Greenberg, and G. Ho, The Context-Aware Pill Bottle and Medication Monitor, retrieved March 13, 2007, from http://pharos.cpsc.ucalgary.ca/Dienst/UI/2.0/Describe/ncstrl.ucalgary_cs/2004-752-17

18. B.J. Rhodes, The wearable remembrance agent: A system for augmented memory, *Proceedings of the First International Symposium on Wearable Computers (ISWC '97)*, Cambridge, Massachusetts, October 1997, pp. 123–128.
19. M.D. Cantor, *Generations*, 46, 2006.
20. T.S. Barger, D.E. Brown, and M. Alwan, *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 22, 2005.
21. B. Yang and S. Rhee, *Robotics and Autonomous Systems*, 373, 2000.
22. K.M. Gatzoulis and A.L. Lymberis, *Wearable Health Systems: From Smart Technologies to Real Applications*, retrieved March 12, 2007, from embc2006.njit.edu/pdf/1049_Lymberis.pdf
23. M.E. Pollack, S. Engberg, J.T. Matthews, S. Thrun, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, J. Dunbar-Jacob, C. McCarthy, M. Montemerlo, J. Pineau, and N. Roy, *Proceedings of the AAAI Workshop on Automation as a Caregiver (2002)*, retrieved March 12, 2007, from www.ai.sri.com/~peintner/papers/aaai02wkshp.pdf
24. M. Rudary, S. Singh, and M.E. Pollack, Adaptive cognitive orthotics: Combining reinforcement learning and constraint-based temporal reasoning, *Proceedings of the 21st International Conference on Machine Learning*, Banff, Canada, 2004, retrieved March 5, 2007, from <http://www.eecs.umich.edu/~mrudary/PAPERS/aaai04-cogorth.pdf>
25. T.A. Salthouse and R.L. Babcock, Decomposing a adult age differences in working memory, *Developmental Psychology*, 27: 763–777, 1991.
26. T.A. Salthouse and V.E. Coon, *Journal of Gerontology*, 245, 1993.
27. A.J. Stronge, W.A. Rogers, and A.D. Fisk, *Journal of Telemedicine and Telecare*, 13: 1–3, 2007.
28. D.R. Kaufman, J. Starren, V.L. Patel, P. Morin, C. Hillman, J. Pevzner, R.S. Weinstein, R. Goland, and S. Shea, *AMIA Annual Symposium Proceedings*, 2003.
29. B.J. Diamond, G.M. Shreve, J.M. Bonilla, M. V. Johnston, J. Moordan, and R. Branneck, *NeuroRehabilitation*, 171, 2003.
30. S. Shea, R.S. Weinstein, J. Starren, J. Teresi, W. Palmas, L. Fields, P. Morin, R. Goland, R.E., Izquierdo, L.T. Wolfe, M. Ashraf, C. Hillman, S. Silver, S. Meyer, S.D. Holmes, E. Petkova, L. Capps, and R. Lantigua, *Journal of the American Medical Informatics Association*, 40, 2006.
31. S.J. Starren, G. Hripcsak, S. Sengupta, C.R. Abbruscato, P.E. Knudson, R.S. Weinstein, and S. Shea, *Journal of the American Medical Informatics Association*, 25, 2002.
32. N.R. Chumbler, B. Neugaard, R. Kobb, P. Ryan, Q. Haijing, and Y. Joo, *Journal of Telemedicine and Telecare*, 150, 2005.
33. T.E. Barnett, N.R. Chumbler, B. Vogel, R.J. Beyth, H. Qin, and R. Kobb, *The American Journal of Managed Care*, 467, 2006.

**SECURITY AND
PRIVACY IN
TELEMEDICINE**

4

Chapter 10

Security and Privacy in Mobile Telemedicine

Jungwoo Ryoo, Young B. Choi, and Tae Hwan Oh

CONTENTS

10.1 Introduction.....	176
10.2 Types of Mobile Telemedicine Technologies	178
10.2.1 WLAN, MANET, WPAN, and 3G	178
10.2.2 RFID.....	180
10.2.3 Wireless Sensor Networks.....	181
10.2.3.1 Pulse Oximetry	181
10.2.3.2 pH Monitoring.....	181
10.2.3.3 E chocardiogram	182
10.2.4 RFID and Micro-Sensors	182
10.3 Mobile Telemedicine Technologies and Their Security Implications.....	183
10.3.1 Confidentiality and Integrity	183
10.3.1.1 WLAN	183
10.3.1.2 MANET.....	185
10.3.1.3 WPAN.....	185
10.3.1.4 3G	186
10.3.1.5 RFID	186
10.3.1.6 Wireless Sensor Networks.....	187

- 10.3.2 Availability..... 187
- 10.4 Privacy in Telemedicine 188
- 10.5 Security and Privacy Regulations Impacting Mobile Telemedicine 189
 - 10.5.1 HIPAA Security Rules 189
 - 10.5.2 HIPAA Privacy Rules..... 190
- 10.6 Outstanding Issues..... 190
 - 10.6.1 Interoperability 190
 - 10.6.2 Balance..... 191
 - 10.6.3 Coordination 191
 - 10.6.4 Compliance 191
- References 191

Due to its heavy reliance on wireless technologies, mobile telemedicine suffers from the same set of security and privacy problems as its underlying technical standards (such as WLAN, MANET, WPAN, 3G, RFID, and WSN) face. This chapter looks into how each of these wireless technologies is being used in mobile telemedicine (see Table 10.1 for a summary), and discusses the security and privacy measures they provide (see Table 10.2). In addition, the current regulatory environment is considered to offer an in-depth view of how the security and privacy safeguards (made available by the state-of-the-art technologies) are being enforced in real-life situations. The chapter is concluded by a discussion of outstanding security and privacy issues in mobile telemedicine.

10.1 Introduction

Telemedicine is a newly emerging branch of medicine that takes advantage of telecommunications technologies to treat patients in remote locations.¹ In the early stage of telemedicine, very basic telecommunications technology such as telephony was used to enhance the quality of medical services for people to whom large-scale, advanced medical care was not easily accessible. However, due to the shift in the telecommunications industry from analog (voice-oriented) to digital (data-oriented), telemedicine today is much more sophisticated than what it used to be. Thanks to the technical advances in telemedicine and robotics, doctors can now operate on a patient in an operating room thousands of miles away from his or her physical location.

With the advent of cellular phones and other mobile telecommunications technologies, telemedicine is going through another major set of changes. More and more conventional telemedicine devices are relying on cutting-edge mobile telecommunications technologies. This latest phenomenon of telemedicine is collectively referred to as mobile telemedicine.^{2,3}

The popularity of mobile telemedicine is growing fast mainly due to the convenience associated with the technology. Instead of carrying bulky folders and worrying about loose paper, in the mobile, paperless world medical professionals can instantly access and modify data residing on a central server at a remote location through portable devices such as personal digital assistants (PDAs), laptops, or cellular phones. This ubiquitous use of mobile devices can potentially result in a dramatic increase in productivity, but the productivity gain is achieved at the cost of increased security and privacy concerns. The information technology (IT) industry is already witnessing an unprecedented number of computer crimes, and mobile telemedicine is not immune from these threats.

In mobile telemedicine, ensuring security and privacy is particularly important due to the following reasons:

- The nature of data being exchanged is fundamentally different. If not handled properly, some of the data could even be life-threatening because the wrong data can easily lead to a fatal misdiagnosis.
- Although not fatal, stolen, or revealed, medical information can be used to commit crimes and discriminate against a person for the purposes of insurance sales, hiring, etc.

Compared to the paper-based and other traditional forms of medical data, the digitized information utilized in mobile telemedicine is much more vulnerable to theft, duplication, and forging, which serves as another justification for extra security and privacy precautions.⁴⁻⁷

In this chapter, security and privacy issues in mobile telemedicine will be discussed. In-depth discussions of various types of mobile telemedicine technologies and outstanding issues also follow.

There are several types of mobile telemedicine technologies, including wireless local area networks (WLANs), mobile ad-hoc networks (MANETs), wireless personal/patient area networks (PANs), third generation (3G) cellular phone technologies, radio frequency identification (RFID) technologies, and wireless sensor networks (WSNs).

Security in mobile telemedicine is explained by mobile telemedicine technologies and their security implications in terms of confidentiality, integrity, and availability.

Privacy in mobile telemedicine is not tied to a specific technology. In fact, the question of how privacy is protected is largely answered in the security section. Therefore, in this chapter the privacy section focuses more on how privacy is dealt with in the context of mobile telemedicine, mainly concerning better awareness and education for patients and medical staff.

In addition, we look into regulations governing the security and privacy aspects of mobile telemedicine. The Health Insurance Portability and Accountability Act (HIPAA) of 1996 and legal safeguards it provides will therefore be discussed.

10.2 Types of Mobile Telemedicine Technologies

10.2.1 WLAN, MANET, WPAN, and 3G (Figure 10.1)

WLANs enable transmission of data between network elements using radio frequencies through the open air. IEEE governs most of the popular WLAN standards in use today. These standards include IEEE 802.11a/b/g and others. WLANs are attractive both for consumers and businesses because they do not require expensive cabling and enhance mobility and accessibility to the network.

Already large portions of health care information management systems rely on the WLAN technology. For instance, doctors and nurses casually carry personal digital assistants (PDAs) and laptops that are used either to view or to submit information via WLAN connections.

Mobile ad hoc networks (MANETs) are a special form of WLAN that consists of mobile network nodes and does not require any centralized administration to switch/route data packets. In a MANET, each individual network node can play the role of a switch and a router to send data packets to a final destination.

Wireless personal area networks (WPANs) are also a specialized type of WLAN, which features low-power connections whose range is only a few meters. The technology is used mainly to eliminate cables that connect network devices in the

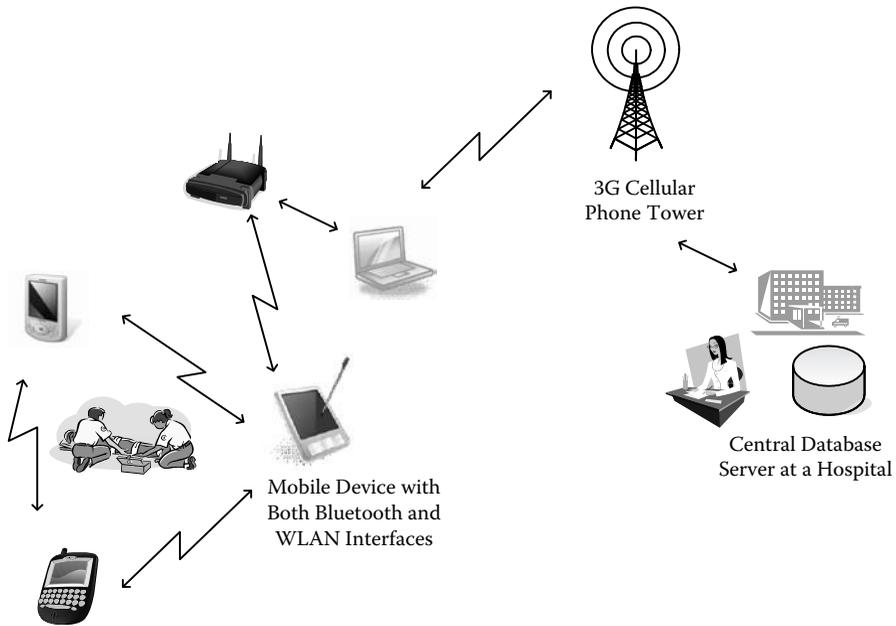


Figure 10.1 Applications of mobile telecommunications technologies in telemedicine

Table 10.1 Mobile Telemedicine Technologies

<i>Technologies</i>	<i>Major Features</i>
Wireless local area network (WLAN)	A network in which data transmission is performed by using radio frequencies through the open air IEEE 802.11a/b/g and other protocols are used.
Mobile ad hoc network (MANET)	A special form of WLAN Each individual network node can play the role of switches and routers.
Wireless personal area network (WPAN)	A special form of WLAN Low power, short range wireless network connections are provided to cover the distance of only a few meters.
Third generation (3G) cellular communications	Third generation cellular phone technologies that can handle high speed data connections as well as voice
Radio frequency identification (RFID)	A wireless identification system consisting of tags (transponders) and readers (interrogators) Used to identify a person or an object Can be combined with micro-sensors and a global positioning system (GPS)
Wireless sensor network (WSN)	A wireless network consisting of spatially distributed devices to monitor a physical environment

vicinity. As in general WLAN technologies (i.e., IEEE 802.11), a WPAN can also be configured to operate in an ad hoc mode (therefore, forming its own MANET). Bluetooth is one of the best-known standards implementing the WPAN concept. Especially in the medical field, PAN often stands for patient area network (PAN). General medical applications of the WPAN technology focus largely on the elimination of wires from medical devices, including:

- Data collection devices for measuring vital signs such as body temperature and pulse, making patients more mobile
- Devices needing to exchange information (e.g., between a doctor's PDA and an electrocardiogram machine)
- Devices requiring constant and seamless uploading of information to a centralized database

In mobile telemedicine, combined with third generation (3G) cellular phone technologies, MANETs⁸ and WPANs are particularly relevant to disaster response

or rescue situations.⁹ As one may expect in a disaster situation, severe damage to the existing telecommunications infrastructures is most likely. First responders often need to create an entirely new network from scratch. For these purposes, emergency workers in the field can use a WPAN technology such as Bluetooth to form a lowest-tier network both for voice and data communications among people and devices around them. However, to forward data packets to network elements that are farther away, a designated device equipped with both Bluetooth and an IEEE 802.11 network adapter may be necessary. The device can then collect the outbound packets and send them through the WLAN interface that has a longer range than Bluetooth. If meant to be sent to a location completely outside the disaster area, the packets will eventually have to arrive at a destination that has a gateway to the desired site. The gateway can be implemented by satellite links or a connection to an existing 3G cellular network.

A teletrauma system¹⁰ is one of the real-life applications of the approach described so far. First responders in an emergency situation can reliably provide prehospital care by relaying the video and vital signs of a patient to the medical staff in a hospital and by receiving instructions directly from the doctors.

10.2.2 RFID

RFID is a system that wirelessly identifies an object or a person. An RFID system consists of tags (also called transponders) and readers (also called interrogators). A tag attaches to a person or an object, and when the person or the object passes through an electromagnetic field generated by the reader, the tag is energized and emits a wireless signal that is, in turn, detected by the reader. The information transmitted by the tag could be a serial number, a model number, an employee number, a Universal Product Code (UPC), etc.

RFID systems are becoming very popular because the tags they use are relatively inexpensive and have limited but useful storage and computing capabilities. They are frequently found in supply chain management (SCM) applications and in many other situations that require the identification of products and people.

In mobile telemedicine, RFID systems are also beginning to be used more widely. For instance, in 2003 the Alexandra Hospital in Singapore used an RFID tracking system to track a severe acute respiratory syndrome (SARS) outbreak.¹¹ The RFID tracking system kept records of patients, visitors, and staff members entering the hospital using RFID-based ID cards. If a person was diagnosed with SARS within the hospital, all individuals who entered the hospital building within a certain time frame could immediately be identified.

In another example, in 2003, Intel established the Center for Aging Services Technologies (CAST). In 2004, CAST demonstrated a monitoring system that keeps track of medicine bottles, teacups, and other objects used regularly at home by affixing RFID tags on the items and by attaching an RFID reader to the back of

a patient's hand. When patients grabbed the bottles or teacups, the reader scanned the items and recorded which containers had been used.

In addition to the anecdotal uses described above, RFID systems can be adopted for helping manage medical equipment and medicine.

10.2.3 Wireless Sensor Networks

A wireless sensor network consists of spatially distributed devices monitoring physical and environmental characteristics such as temperature, pressure, and noise. Sensor networks were originally developed for military applications like battlefield surveillance but recently they have also been used for several private-sector applications as in mobile telemedicine.

For example, sensor networks can monitor patients who need long-term care and allow doctors and nurses to check patients' health remotely. More specifically, micro-sensors can be attached to a patient's body to collect electrocardiogram, pulse rate, body temperature, and other vital signs. Because the sensors are wireless, the patient can freely move around. The sensors transmit data wirelessly to a cluster head to perform special processing (including compression) before the data is sent to a mobile device in the vicinity like a PC, a PDA, or a cell phone. Additionally, instead of the local storage methods, the sensors can forward the data to a remote medical specialist over a wireless connection. The specialist can then diagnose the patient. A sleep feature in the sensor architecture extends the battery life of each sensor constituting the network. The sleep function activates and deactivates the entire sensor network depending on the current needs.

Sensor networks can also monitor the health of first responders such as soldiers, firefighters, and police officers while they are performing their duties during an emergency. More sensor usage examples in mobile telemedicine are listed in the following subsections.

10.2.3.1 Pulse Oximetry

The oximetry is a device that measures the amount of oxygen in blood. To measure the oxygen level, a small sensor is attached to a finger or a toe. The sensor tap is painlessly applied and is typically accompanied by an indicator (e.g., a small red light). Whenever a physician needs the oximetry data, he or she can query the sensor to transmit the real-time data to a base station.

10.2.3.2 pH Monitoring

This technology can measure the acidity in the esophagus in cases of gastroesophageal reflux disease (GERD). The acidity is measured by a thin plastic tube inserted into a nostril and guided down the throat into the esophagus. The tube is inserted

until it reaches the lower esophageal sphincter (located between the esophagus and the stomach). A sensor is located at the end of the tube to measure the acidity. The other end of the tube is connected to a wireless monitoring device that calculates the acidity. The patient can perform normal activities while the acidity is measured periodically. The entire device is now available in a capsule that is swallowed and can perform the same function.

10.2.3.3 Echocardiogram

An echocardiogram is an ultrasound scan of one's heart. A sensor is placed on the chest of a patient. Then a sound image is taken through the chest wall. This is noninvasive and provides a relatively accurate assessment of a patient's heart health. One peculiar application of the echocardiogram technology is evaluating the structure and function of the heart for babies with Down syndrome by a pediatric cardiologist. The current statistics show that 40 to 50% of Down syndrome babies have heart problems. The sensors can record the moving pictures of heart and heart valves especially during the first two months after birth so that they can detect any heart problems.

10.2.4 RFID and Micro-Sensors

Another interesting trend in mobile telemedicine is the integration of RFID and micro-sensors, which are miniaturized sensors designed to be less expensive, more fault-tolerant, and easier to maintain than regular sensors. They are prefabricated to sense measured parameters and to ignore other parameters.

Working with other mobile telemedicine infrastructures, the combination of a micro-sensor and an RFID tag in a single form factor can offer critical, pseudo-real-time data from patients to medical personnel monitoring the patients at remote locations. Therefore, the RFID–micro-sensor technology can provide high quality medical services to patients both inside and outside a hospital around the clock.

For example, Intel's Caregiver's Assistant and Georgia Tech's Memory Mirror utilize RFID tags combined with micro-sensors to monitor the health conditions of the elderly at home and to constantly inform caregivers of health status.¹²

Additionally, RFID tags equipped with micro-sensors can be used to monitor patients in hospitals, disaster areas, and ambulances in conjunction with real-time patient monitoring systems. The monitoring systems can collect the patients' vital signs in real-time so that remote diagnoses can be performed anywhere and any time. When combined with a global positioning system (GPS), the monitoring system can also be used to locate mobile patients. This significantly reduces the response time during medical emergencies.

10.3 Mobile Telemedicine Technologies and Their Security Implications

10.3.1 Confidentiality and Integrity

In mobile telemedicine, weak security measures cannot be tolerated because the data being exchanged tend to be extremely sensitive. In the worst-case scenario, a person's entire medical history can be at risk. Especially in a man-made disaster situation such as a terrorist attack, communications among first responders need to be protected to prevent the terrorists from jeopardizing the rescue efforts and inflicting further damage by using the stolen information. This type of security control that restricts access to secretive data is, in general, referred to as confidentiality.

In addition to the information-theft scenarios, even more sinister are attacks that violate the integrity of data. For instance, any malicious manipulations in the patient's medical information (i.e., what medication he or she is currently taking, any known health problems, current measurements of vital signs, etc.) may be highly detrimental to diagnosing and treating the patient correctly.

This section discusses how the mobile telemedicine technologies explained in Section 10.2 fare against each other in terms of ensuring confidentiality and integrity.

10.3.1.1 WLAN

WLAN is inherently vulnerable to security attacks due to its use of radio signals that are readily accessible to the public.⁶ Various exploits have been developed to bypass basic WLAN security measures. For example, wired equivalent privacy (WEP) is an encryption method still used today to encrypt wireless data on many wireless routers but a hacking tool can crack the encrypted messages within minutes.

Wi-Fi protected access (WPA) is a newer standard developed to provide more comprehensive security protection for WLAN. One of the salient features of WPA is its use of the IEEE 802.1x standard, which requires authentication through remote authentication dial-in user service (RADIUS). As its name suggests, RADIUS is a third-party server dedicated to verifying the identity of every WLAN user before granting access to the network. In a mobile telemedicine setting, RADIUS may not always be available. An alternative for RADIUS in WPA is a preshared key that is a passphrase stored once and shared among multiple users to gain access to the network over and over again. Temporal Key Integrity Protocol (TKIP) defines an encryption algorithm for WPA. One of the features that make TKIP more secure than WEP is its ability to change encryption keys for every data frame. Advanced encryption standard (AES) is an even more sophisticated encryption scheme than TKIP. AES is also supported by WPA but often requires a hardware upgrade in addition to the firmware upgrade. Although much stronger, WPA can also be cracked if a weak preshared key is used.

Table 10.2 Confidentiality and Integrity Safeguards in Mobile Telemedicine

<i>Technologies</i>	<i>Security Features</i>
Wireless local area network (WLAN)	Several security protocols are used: Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA), Temporal Key Integrity Protocol (TKIP), WPA2 (also referred to as IEEE 802.11i), etc. Mechanisms for end-to-end security are provided.
Mobile ad hoc network (MANET)	MANET presents new challenges due to its open network architecture, use of a shared wireless medium, stringent resource constraints, and highly dynamic network topology. Security is provided at the network and data link layers.
Wireless personal area network (WPAN)	Three security modes are provided: no security, service level security, and data link-level security. Bluetooth does not provide end-to-end security.
Third generation (3G) cellular communications	Subscriber identity module (SIM) for global system for mobile communications (GSM) and re-usable identification module (RUIM) for code division multiple access (CDMA) are primary authentication methods for 3G terminals. 3G networks are still vulnerable to malicious software (malware).
Radio frequency identification (RFID)	Tag deactivation and tag shielding techniques for security control are used but proved to be ineffective. An anonymous tag hiding a permanent identifier allows users to assign their own private IDs.
Wireless sensor network (WSN)	Compact and abbreviated versions of the existing algorithms are emerging. Authentication for WSNs can be done through either direct or indirect validations.

WPA is a temporary solution for the WLAN industry to cope with immediate security challenges before the introduction of a full-blown, next-generation WLAN security standard called IEEE 802.11i (also referred to as WPA2). Unlike WPA, WPA2 mandates AES, making hacking almost impossible with the existing computing technologies.

10.3.1.2 MANET

Although security has been an active research area in wireless networking for some time, the unique characteristics of MANET present a new set of security design problems due to its open network architecture, use of a shared wireless medium, stringent resource constraints, and a highly dynamic network topology. As a result, most of the existing security solutions for other types of networking do not work well with a MANET environment.

MANET security can be implemented at a network layer or a data link layer. The network layer security measures are concerned with protecting the network functionality of MANET to deliver packets between mobile nodes through multiple hops. These security measures ensure

- *Protocol specification consistency*: The secure ad hoc routing protocols in MANET use a proactive approach and enhance the existing ad hoc routing protocols such as distance source routing (DSR) and ad hoc on-demand distance vector (AODV) routing with a security extension.^{13–16} In these protocols, each mobile node proactively utilizes cryptographic authentication to encrypt messages before transmission. The collaborating nodes can efficiently distinguish legitimate traffic from unauthenticated packets inserted by outside attackers.
- *Secure packet forwarding*: Protecting message exchanges is only a partial solution because the attackers can still participate in route discoveries and disrupt the forwarding table. The secure packet forwarding solution minimizes this possibility.

At the data link layer level, MANET protects one-hop communications between direct neighbors through secure media access control (MAC) protocols.

10.3.1.3 WPAN

Although WPAN can be implemented using various technologies, Bluetooth is the most common form of WPAN in use today. Bluetooth security is enforced in one of the following three modes¹⁷:

1. *Security mode 1*: No security
2. *Security mode 2*: Service-level security
3. *Security mode 3*: Data link-level security

Obviously, no security does not provide any security protection at all. This mode is used for applications that do not require secure connections (e.g., wireless mice). Any Bluetooth devices operating in this mode can freely connect to another Bluetooth device. The second mode provides security once a data link layer connection is established. A security manager provides access control for computing resources available on a Bluetooth-enabled device. Fine-grained access control (i.e., offering different permissions for multiple roles) is possible, too. In the third mode, a pairing process includes authentication before an actual communication channel establishment. The authentication step requires a personal identification number (PIN) that is used as a basis for building a key for encrypting data sent through a Bluetooth channel. A weak PIN is therefore dangerous because it may lead to the revelation of the encryption key. Unlike WLAN, Bluetooth does not provide any mechanisms for end-to-end security.

10.3.1.4 3G

Subscriber identity module (SIM) and re-usable identification module (RUIM) are primary authentication methods currently used for 3G terminals. These cards hold information uniquely identifying a subscriber and plug into any compatible devices, effectively decoupling the concept of ownership and the corresponding hardware. Convenience is the biggest advantage of using the SIM and RUIM cards because they free their users from having to memorize PINs. SIM cards are used for a global system for mobile communication (GSM) while RUIM cards are for code division multiple access (CDMA). Both GSM and CDMA have built-in air interface encryption mechanisms, making it very difficult for hackers to intercept messages. Despite these relatively strong security measures, 3G networks are still vulnerable to malware (malicious software). In addition to connectivity to cellular networks, most of the 3G terminals have multiple communications interfaces such as WLAN and Bluetooth. Once injected through these secondary interfaces, malware can wreak a havoc not only to the device itself but also the entire network the device is connected to by bringing it down for an extended period of time.

10.3.1.5 RFID

Amid growing security concerns on RFID tags, researchers are developing a wide range of solutions.¹⁸ One of the most basic approaches is deactivating RFID tags that are no longer used to minimize the possibility of information theft. When necessary, the deactivated tags can be reactivated. Another way of protecting information residing in RFID tags is shielding the tags so that unwanted scanning does not occur. The effectiveness of this approach is questionable because the tags eventually need to be unshielded to be useful. A more practical way is using a locking mechanism. Because the lock requires a key, this approach is also vulnerable to

many traditional attack methods exploiting key-based systems, including spoofing, replay attacks, etc. Conventional authentication algorithms turn out to be inappropriate for RFIDs due to their resource-intensive nature. Lightweight authentication is necessary for practicality. Encryption can offer added security as long as the processing burden of decryption is kept negligible. Finally, there are attempts to hide the permanent identifiers on RFID tags and to allow users to assign their own private IDs, which results in an anonymous tag.

10.3.1.6 Wireless Sensor Networks

To ensure confidentiality and integrity, wireless sensor networks (WSNs) also employ encryption technologies although their adoption is challenging due to limited CPU and memory capacity. Particularly, more advanced encryption algorithms such as asymmetric key encryption (using two different keys for encryption and decryption) presents a high hurdle. However, more compact and abbreviated versions of the existing encryption algorithms are emerging. To offload some of the significant computational burden, sensors often process only the upfront part of a cipher (i.e., an encryption task with a public key), and send more resource-intensive calculations (decryption with a matching private key) to a base station with more computing resources. Authentication for WSNs can be done through either direct or indirect validations. In a direct validation, an already validated node assumes the authentication responsibility, while in an indirect validation, a third party conducts the authentication task. The actual mechanisms for validation range from the use of a limited number of radio frequency channels to applying tight control over the number of available keys for any given node to use to initiate communication.

10.3.2 Availability

In addition to information theft and unauthorized modification, denial of service (DoS) attacks can also be deadly. Interruption of data exchanges between first responders onsite and medical personnel at a remote location can make a difference between life and death. Suppose that instructions from a doctor to inject a certain type of medication to resuscitate a patient are lost due to a DoS attack. One can easily see the consequences. Anybody with basic equipment can launch a DoS attack to all the mobile telemedicine technologies discussed so far by jamming certain radio frequencies.

Another common type of DoS attack is flooding, overwhelming stations with an enormous number of connection requests or misleading data streams.

Although the low-level specific defense mechanisms for disparate mobile telemedicine technologies may differ significantly, commonly used solutions include (1) adding redundancies in the network so that legitimate traffic can be rerouted

around the disabled region of a network and (2) incorporating more intelligence into network nodes to be able to discern invalid traffic from valid traffic.

10.4 Privacy in Telemedicine

Incorporating computer networks into health care organizations has been a huge part of the medical revolution in recent years. Additionally, mobile telemedicine opens up a unique opportunity to treat patients especially in remote areas or in emergency situations. Despite these benefits, mobile telemedicine has its own disadvantages because it involves exchanging the detailed medical records of patients among medical staffs and third parties through wireless and wired network connections to understand the patients' current health conditions and eventually to provide further medical treatment. Therefore, providing quality medical care without jeopardizing the privacy of a patient is very crucial in mobile telemedicine.^{4,5}

Modern medical records have much more detailed and sensitive information about an individual than other types of records (e.g., driver's licenses and credit cards) do. The records could include dietary habits, sexual orientation, sexual activities, employment status, income, eligibility for public assistance, history of diseases, treatments rendered, medications taken, diagnostic information, psychological profiles, genetic testing, family history, doctors' and nurses' subjective notes about patient personality and mental state, and much more.¹⁹ Mercuri characterized a medical record as "an entire work-up of your being."²⁰ In addition to storing the potentially damaging information when misused, the mobile telemedicine technology has made the sensitive personal information readily available to too many people who have the ability to distribute it through the Internet with just a few mouse clicks. Because the Internet is built upon an open architecture, any traffic passing through it could easily be intercepted by someone who is reasonably familiar with the technology. If medical records are leaked to a wrong person, the incident could have a devastating effect on every area of the patient's life.

To circumvent these privacy problems, all the people involved in mobile telemedicine must be fully aware of the security vulnerabilities of the system they use daily at an operational level. Ideally, a voluntary scheme can be put in place, but this is most probably inefficient. Therefore, it is crucial for the management to establish a clear set of policies, standards, procedures, and guidelines for employees who have to tackle privacy issues constantly. Awareness campaigns and education must follow the creation of rules for effective enforcement. It is also highly important to let patients know about the potential dangers of privacy breaches and what the organization is doing about them to minimize the possibility.

Most importantly, tangible security safeguards should be used to implement privacy rules. For instance, only a patient ID number must be stored in a mobile

telemedicine device to curtail any malicious attempts to steal additional information. Database access must be limited to people who have more layers of credentials. Added security should be applied to the computer/communications systems in which the database is located to prevent local/remote access from any unauthorized users.

10.5 Security and Privacy Regulations Impacting Mobile Telemedicine

Until recently, little has been done to regulate the handling of security and privacy matters in the medical industry. This trend has radically been changed since the introduction of the Health Insurance Portability and Accountability Act (HIPAA) of 1996 (Public Law 104-191).^{4,21} HIPAA “specifies the privacy, security, and electronic transaction standards with regard to patient information for all health care providers.”^{21,22} It provides standardization in the health care industry with the strong consideration of rapidly expanding technologies. HIPAA contains major provisions on insurance reform and electronic data interchange (EDI). The provisions on EDI have particular significance in mobile telemedicine because they address the security requirements for high-volume patient information exchanges within and across different organizations. Other elements of HIPAA include standard code sets for diagnoses and procedures, privacy standards for mandating the use of protected health information (PHI), unique identifiers such as a national provider identifier, and security standards. HIPAA seeks to validate and assist with the inevitability of electronic data transactions while also addressing privacy and security issues.

10.5.1 HIPAA Security Rules

The security aspect of HIPAA specifically addresses PHI in an electronic form. The rules require that PHI, whether in storage or in transmission, must be kept confidential and protected against unauthorized users and other threats. While establishing behavioral standards, the rules do not regulate the functional capabilities of computer applications.

The HIPAA security rules enforce compliance on the following categories:²¹

- *Administrative safeguards*: Formal practices to manage security and personnel
- *Physical safeguards*: Protection of computer systems and the facilities within which they reside
- *Technical safeguards*: Control and monitoring over information access
- *Organizational requirements*: Business associate contracts
- *Policies, procedures, and documentation requirements*

Compliance with the HIPAA security rules generally demands identifying those areas that must be changed to support the standards. All of the above safeguards are important but, in mobile telemedicine, the technical safeguards are especially important due to its heavy reliance on wireless technologies that inherently have more vulnerabilities than their wired counterparts, as already discussed in the previous sections.

10.5.2 HIPAA Privacy Rules

The purpose of the HIPAA privacy rules is to “meet the pressing need for national standards to control the flow of sensitive health information and to establish real penalties for the misuse or improper disclosure of this information.”^{21,23} In HIPAA, the privacy rules apply to the oral, written, and electronic forms of PHI. The general rules within the privacy regulations control the use and disclosure of PHI for treatment, payment, and health care operations, the minimum necessary use and disclosure of information, creation of de-identified information or a limited data set, application of standards to business partners through contracts, application to information about the deceased, adherence to the notice of privacy practices, and “application as covered entities components of organizations that are not covered entities.”^{21,23}

Generally, the HIPAA privacy rules protect individuals’ PHI by managing when PHI is disclosed and the reasons for disclosure. It grants individuals more involvement by allowing them to have specific rights to access their medical records and to request amendments, to authorize or restrict the disclosure of their information in certain circumstances, to be informed of the way in which their information is shared with others, and to be informed of their rights relating to privacy.^{21,23}

10.6 Outstanding Issues

There are several outstanding issues in ensuring the security and privacy of mobile telemedicine. They are strongly related to the inherent vulnerabilities of wireless communications.

10.6.1 Interoperability

Many new wireless security protocols are under development, and various types of existing wireless security protocols are used together. Interoperability among the protocols and the overhead of protocol translations to achieve interoperability make the implementation of security and privacy measures in mobile telemedicine more difficult.

10.6.2 Balance

Better balanced security and privacy policies based on agreements between a health care provider and a patient should be considered. The policies are too often one-sided (that is, dominated by the health care provider's interests). A comprehensive information protection scheme taking into account both the health care provider's interests and the patient's security and privacy concerns is critical for the successful adoption of mobile telemedicine technologies.

10.6.3 Coordination

Mobile telemedicine is heavily dependent upon the telecommunications technologies (especially mobile and wireless communications technologies). Patient information is transmitted across many telecommunications boundaries such as local access and transport areas (LATAs), telecommunications service providers (TSPs), and different states. Disparate regulations regarding security and privacy protection for the senders and the receivers should be aligned to each other to guarantee the maximum security and privacy protection of patients' health information.

10.6.4 Compliance

The HIPAA compliance requirements for credentialing, accreditation, and reimbursement make it challenging to implement security and privacy policies in mobile telemedicine.⁴ There are many potential loopholes to close in compliance with HIPAA regulations. True and meaningful compliance will occur only when health care providers, medical staffs, patients, insurance companies, mobile device manufacturers, and mobile telecommunications service providers work together and agree to what the critical vulnerabilities are, who is responsible, and how the weaknesses can be mitigated.

Therefore, to guarantee security and privacy in mobile telemedicine, all the players in a telemedicine environment should be aware of the importance of security and privacy through education, training, research and development, system deployment, and maintenance activities.

References

1. D. Perendina and A. Allen, Telemedicine technology and clinical applications, *JAMA*, 273(6), 483–488, 1995.
2. E. Rosen, Mobile telemedicine arrives, *Telemedicine Today*, 5(5), 14–42, 1997.
3. R. Istepanian, Integrated mobile telemedical systems: Current status and future prospects, *Virtual Medical Worlds Monthly*, June 1999, <http://www.hoise.com/vmw/99/articles/vmw/RI-VM-07-99-1.html>.

4. Y.B. Choi, K.E. Capitan, J.S. Krause, and M.M. Streeper, Challenges associated with privacy in health care industry: Implementation of HIPAA and the security rules, *Journal of Medical Systems*, 30(1): 57–64, 2006.
5. Y.B. Choi, J.S. Krause, H. Seo, K.E. Capitan, and K. Chung, Telemedicine in the U.S.A.: Standardization through information management and technical applications, *IEEE Communications Magazine*, 44(4): 41–48, 2006.
6. Y.B. Choi, J. Muller, C.V. Kopek, and J.M. Makarsky, Corporate wireless LAN security: Threats and an effective security assessment framework for wireless information assurance, *International Journal of Mobile Communications*, 4(3): 266–290, 2006.
7. K. Chung, Y.B. Choi, and S. Moon, Toward efficient medication error reduction: Error-reducing information management systems, *Journal of Medical Systems*, 27(6): 553–560, 2003.
8. E.M. Husni, Y. Heryadi, W.T.H. Woon, M.S. Arifianto, D.V. Viswacheda, and L. Barukang, Mobile ad hoc network and mobile IP for future mobile telemedicine systems, *Proceedings of the 2006 IFIP International Conference on Wireless and Optical Communications Networks*, April 2006, p. 5.
9. V. Singh, Telemedicine and Mobile Telemedicine Systems, 2006, http://works.bepress.com/vikas_singh/2/.
10. Y. Chu and A. Ganz, A mobile teletrauma system using 3G networks, *IEEE Transactions on Information Technology in Biomedicine*, 8(4): 456–462, 2004.
11. F.C.Y. Lee, W. Keong Wee, and A. Johan, Public hospital preparations for SARS outbreak: Experience of Alexandra Hospital, *Prehospital and Disaster Medicine*, 20(1): 24–31, 2005.
12. M. B aard, R FID Keeps Track of Seniors, March 2004, <http://www.wired.com/medtech/health/news/2004/03/62723>.
13. H. Deng, W. Li, and D.P. Agrawal, Routing security in wireless ad hoc networks, *IEEE Communications Magazine*, 40(10): 70–75, 2002.
14. M. Conti and S. Giordano, Multihop ad hoc networking: The reality, *IEEE Communications Magazine*, 45(4), 88–95, 2007.
15. H. Yang, H. Luo, F. Ye, S. Lu, and L. Zhang, Security in mobile ad hoc networks: Challenges and solutions, *IEEE Wireless Communications*, 11(1): 38–47, 2004.
16. N. Milanovic, M. Malek, A. Davidson, and V. Milutinovic, Routing and security in mobile ad hoc networks, *IEEE Computer*, 37(2): 61–65, 2004.
17. T. Karygiannis and L. Owens, Wireless network security 802.11: Bluetooth and handheld devices, NIST Special Publication 800–48, 2002.
18. Y. Xiao, X. Shen, B. Sun, and L. Cai, Security and privacy in RFID and applications in telemedicine, quality assurance and devices in telemedicine, *IEEE Communications Magazine*, 64–72, 2006.
19. L. Goldberg, Electronic medical records and patient privacy, *The Health Care Manager*, 18(3): 63–69, 2003.
20. R.T. Mercuri, The hipaa-potamus in health care data security, *Communications of the ACM*, 47(7), 25–28, 2004.
21. 104th Congress, Health Insurance Portability and Accountability Act of 1996.
22. L. Volonino and S.R. Robinson, *Principles and Practice of Information Security*, Pearson Higher Education, Upper Saddle River, NJ, 2003.
23. SNIP, Security and Privacy Workgroup, Introduction, January 2004.

24. Y. Chu and A. Ganz, Mobile telemedicine systems using 3G wireless networks, *U.S. Health Care Strategies*, 2005.
25. T.J. Owens, S. Tachakra, K.A. Banitsas, and R.S.H. Istepanian, Securing a medical wireless LAN system, *Proceedings of the 23rd Annual International Conference of IEEE*, October 2001, vol. 4, pp. 3552–3555.
26. K.A. Banitsas, S. Tachakra, and R.S.H. Istepanian, Operational parameters of a medical wireless LAN: Security, range, and interference issues, *Engineering in Medicine and Biology*, 2002, in *Proceedings of the 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*, October 2002, 3, 1889–1890.
27. E. Kyriacou, S. Voskarides, C. Pattichis, R. Istepanian, M. Pattichis, and C. Schizas, Wireless telemedicine systems: An overview, *IEEE Antennas and Propagation Magazine*, 44(2): 143–153, 2002.
28. E. Kyriacou, S. Pavlopoulos, A. Berler, M. Neophytou, A. Bourka, A. Georgoulas, A. Anagnostaki, D. Karayiannis, C. Schizas, C. Pattichis, A. Andreou, and D. Koutsouris, Multi-purpose health care telemedicine systems with mobile communication link support, *Biomedical Engineering Online*, 2(7), 2003.
29. S. Tachakra, X. Wang, S.R. Istepanian, and Y. Song, Mobile e-health: The unwired evolution of telemedicine, *Telemedicine Journal and e-Health*, 9(3): 247–257, 2003.
30. K. Hung and Y.-T. Zhang, *Wireless Internet in Telemedicine*, CRC Press, Boca Raton, FL, 2003.
31. NIST, MANET and sensor network security, <http://csrc.nist.gov/manet/index.html>
32. W.J. Song, B.H. Ahn, and W.H. Kim. Health care information systems using digital signature and synchronized smart cards via the Internet, *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'02)*, April 2002, 177–182.

Chapter 11

Security of Body Sensor Networks

Shu-Di Bao, Carmen C.Y. Poon, and Yuan-Ting Zhang

CONTENTS

11.1 I ntroduction	196
11.1.1 B io-Channels in BSN.....	197
11.1.2 N etwork Topology	197
11.1.3 S ecurity Challenges.....	198
11.2 C ryptographic Primitives	198
11.2.1 B lock Ciphers.....	199
11.2.1.1 C BC Mode of Operation	199
11.2.1.2 I nitialization Vector.....	200
11.2.2 M essage Authentication Code.....	201
11.2.3 R andom Number Generator	201
11.3 K ey Distribution	202
11.3.1 K ey Pre-Distribution Scheme	202
11.3.2 B iometrics Method-Based Key Distribution	203
11.4 C onclusions.....	206
References	206

It has been widely accepted that body sensor networks (BSN) will take an important role in mobile telemedicine systems as the basic interface for biological data collection, fusion, and drug delivery. To protect the vital medical information, efficient security mechanisms must be properly deployed in a resource-constrained wireless BSN, which faces more serious security challenges compared with a wired BSN, e.g., one that uses e-textile materials to connect the various sensors. This chapter introduces a variety of security techniques that are applicable to wireless BSNs, with emphasis on a novel biometrics method that utilizes the biological channels (bio-channels) to assist secure information transmission.

11.1 Introduction

Recent advances in embedded systems, mobile computing, and communication technologies have led to the emergence of body sensor networks (BSN) as one of the main research trends to facilitate the joint processing of spatially and temporally collected biological data from different parts of the body for resource optimization and systematic health monitoring and diagnosis.^{1,2} Consisting of a sensing unit, microprocessor, transceiver, and battery, each node of the BSN ensures accurate capture of biological data, carries out low level processing of the data, and transmits them to a patient terminal or local station. Data are then further processed and fused before being sent to a central server either via wireless personal area networks, wireless local area networks, or cellular networks, as depicted in Figure 11.1. The personalized BSN is an elementary unit in mobile telemedicine systems, and should be designed for seamless integration in homes, workplaces, and hospital environments.

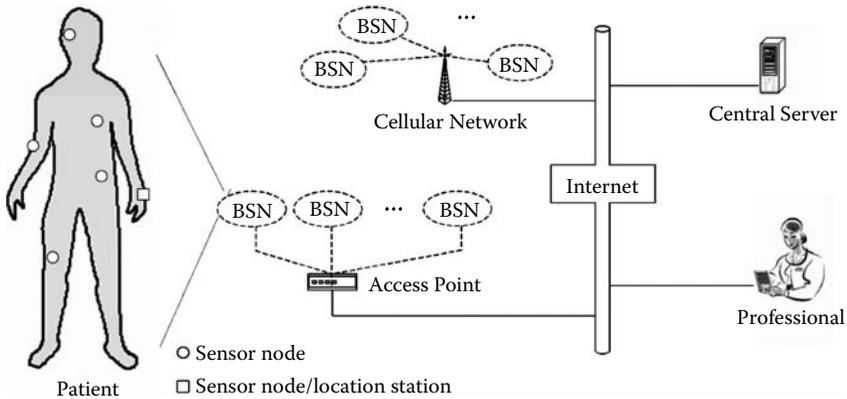


Figure 11.1 Integration of personalized BSN in a mobile telemedicine system.

11.1.1 Bio-Channels in BSN

Different from large-scale sensor networks, sensors of BSN are placed on or in a human body, an area which is comparatively small and unique to each BSN. Consequently, BSN possesses a number of characteristics which researchers could make use of in their development.

For example, because the sensors are concentrated in a small area, radio frequency (RF) communication techniques need not be the only option. Nodes that are close to each other could be connected through “wires,” provided that the appearance of the user is not too adversely affected and his mobility not impaired. In this respect, Park, Mackenzie, and Jayaraman proposed the concept of wearable motherboard, where fabric made of e-textiles is developed into a computer and served as a framework for personalized mobile information processing.³

Even more uniquely, because nodes of the BSN are placed in or on the human body, they are inherently linked by pathways that we name biological channels (bio-channels). Bio-channels are commonly referred to the voltage-gated channels that allow the exchange of selected ions across the otherwise impermeable cell membrane. In this context, we use bio-channel to denote any biological conduit that is part of the human body and enables the transfer of information. Signals transmitted via bio-channels could be either processed information or biological data. As some of the biological data are unique to individuals, they could potentially be an identifier of the owner of the BSN. Thus, a biometrics approach using bio-channels could be used to secure wireless communications in a BSN. We will provide more details of the approach in a later section of this chapter.

11.1.2 Network Topology

Because the scale of the network is relatively small and direct communications between each pair of sensors might not be necessary, star topologies are often suitable for a BSN.⁴ The star topology implies a centralized architecture where the intelligence of the system is concentrated in a central node (master), which is superior to other nodes (slaves) in terms of resources such as processing, memory, and power. The master node also acts as a wireless relay that serves to connect the BSN to the outside world. Advantages of star-based topologies include simple network setup, low power consumption of slave nodes, low latency, and avoidance of routing.

In cases where peer-to-peer communications are required, mesh-based topologies are preferred. As a consequence of shifting intelligence towards the sensors, the BSN consists of smart and self-contained sensors that communicate with one another. Because the peer-to-peer networks are not dependent on any particular component, they are failure-tolerant, i.e., even if one component fails the remaining parts of the system continue to operate. However, nodes in such BSNs must have routing capability, and high latency is also typically a problem. Moreover,

a role negotiation process needs to be carried out among nodes that can act as masters.

As a compromise between star- and mesh-based topologies, a star–mesh hybrid topology combines the simplicity of the single-hop star topology with the extensibility and flexibility of the multi-hop mesh topology. It can be formed by connecting a mesh network with one or more star networks. However, the network setup will be of high complexity where all nodes can act as masters.

11.1.3 Security Challenges

Security issues of the BSN are particularly important because sensitive medical information must be protected from unauthorized usage for personal advantage or fraudulent acts that might be hazardous to the owner's life (e.g., by alteration of system settings, drug dosage, or treatment procedure). The basic security goals to be achieved for the BSN include data confidentiality against eavesdropping, data authentication against message injection and interference, and data freshness against message replay attacks.

Compared to nodes that are connected by wires, e.g., e-textile materials, nodes that communicate through RF techniques face more serious security challenges because of the properties of the deployment:

- Sensor nodes use wireless links, which are particularly easy to eavesdrop on. Similarly, an attacker can easily inject malicious messages into the wireless network.
- Neighboring wireless BSNs interact closely, which increases security vulnerabilities in cases of improperly managed security.
- Because wireless BSN nodes usually have severely constrained resources, asymmetric cryptography is too expensive in terms of system overhead. Thus, a promising approach is to use more efficient symmetric cryptographic alternatives. However, symmetric cryptography is not as versatile as a symmetric cryptographic techniques, which complicates the design of secure applications.

Therefore, the remainder of this chapter will present security solutions for wireless BSN with focus on key distribution, which is always a crucial and difficult problem in symmetric key systems.

11.2 Cryptographic Primitives

Symmetric key algorithms can be classified into stream cipher and block cipher algorithms. A stream cipher is one that encrypts plaintext one byte or one bit at a

time. Stream ciphers typically execute at a higher speed than block ciphers and have lower hardware complexity. However, they can be susceptible to serious security problems if used incorrectly. For example, using the same key for two different messages can open the messages up to attacks. Consequently, the design goal for a stream cipher is the efficient generation and sharing of pseudo-random bit strings that are long enough and indistinguishable from truly random ones. This is still an area of much research. Far more effort has gone into analyzing block ciphers. A block cipher transforms a fixed-length block of plaintext data into a block of ciphertext data of the same length. Typically, a block size of 64 or 128 bits is used. Except for encrypting data, block ciphers can also be designed to produce message authentication codes and pseudo-random numbers.

11.2.1 Block Ciphers

A block cipher effectively provides a permutation of the set of all possible messages. Examples of block ciphers include DES, AES, RC5, and Skipjack. To encrypt a message of arbitrary length, techniques known as modes of operation are used for the block cipher, e.g., electronic codebook (ECB), cipher block chaining (CBC), cipher feedback (CFB), output feedback (OFB), and counter (CTR).⁵ As operation modes such as CFB, OFB, and CTR turn a block cipher into a stream cipher, they share the problems as any stream cipher. Therefore, those modes are not considered for securing a wireless BSN. Neither is the ECB method, which encrypts each plaintext block using the same key and thus produces the same ciphertext for the same plaintext block. This mode is particularly insecure for transmitting lengthy messages. To this end, CBC mode may be the most appropriate method for securing a wireless BSN.

11.2.1.1 CBC Mode of Operation

In this scheme, plaintext blocks are linked together in the encryption operation with an initialization vector (IV). Let $E_K(X)$ denote the encipherment of a plaintext block X using a key K and a block cipher E , P_1, P_2, \dots, P_N represent a sequence of plaintext blocks, and C_1, C_2, \dots, C_N represent the corresponding sequence of ciphertext blocks. The scheme is depicted in Figure 11.2. For encryption, $C_i = E_K[P_i \oplus C_{i-1}]$ and $C_0 = IV$; for decryption, $P_i = D_K[C_i] \oplus C_{i-1}$. The IV must be known to both the sender and receiver.

Any block-oriented modes may result in message expansion because the size of produced ciphertexts is exactly multiples of the block size. As a result, the power consumption of transmitting the ciphertexts will be higher. A technique known as ciphertext stealing⁶ can be used to ensure the ciphertext has the same length as the underlying plaintext that is longer than one block.

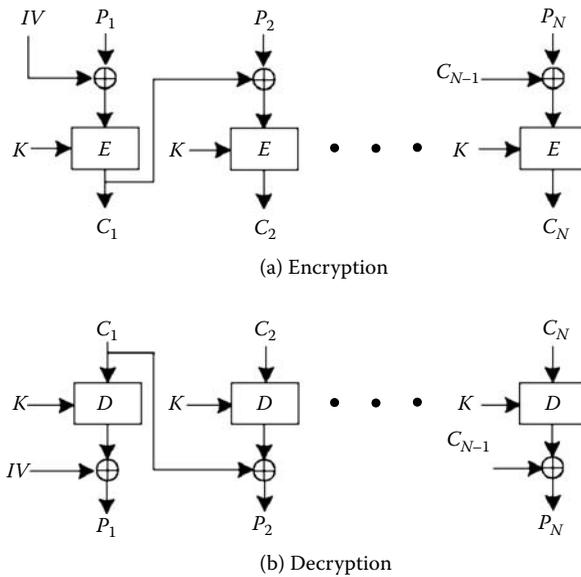


Figure 11.2 Cipher block chaining mode.

11.2.1.2 Initialization Vector

The IV can be used to achieve semantic security. Semantic security ensures that an eavesdropper has no information about the plaintext, even if it sees multiple ciphertexts of the same plaintext with the same key. Consider application messages with low entropy, such as YES or NO messages that are sent periodically to indicate the patient’s overall health status. Without semantic security, all encryptions of YES messages are identical. Once an adversary determines what a YES message looks like, it can determine the contents of every YES/NO message by simply looking at its encryption. The IV is also intended to ensure data freshness against message replay. Typically, the value of IV is set to increase monotonically with every message, e.g., $IV = IV + 1$.

For maximum security, the IV should be known only to the sender and receiver of the message. In practice, the IV need not always be secret, as long as values of IV are different for the encryption process with the same secret key. Generally, there are two ways for both ends to share the IV. One way is to send the IV along with each message. In such cases the packet format must be carefully designed due to an unavoidable increase of energy consumption. An example can be found in TinySec.⁷ Another way is to maintain the IV at both ends. Every receiver must maintain a table of the last value of IV from every sender it receives, like maintaining counters at each node designed by SPINS.⁸ This method may be applicable to a wireless BSN with a star topology because every slave node may only need to maintain two counters, i.e., one for two-party communications and another for

broadcast communications. However, in mesh topology-based networks it would become problematic due to stringent memory constraints.

11.2.2 Message Authentication Code

Message authentication code (MAC) is a public function of the message and a secret key that produces a fixed-length value that serves as the authenticator to achieve message authentication and integrity. It is well known that using encryption without authentication is insecure. For example, flipping bits in unauthenticated encrypted messages can cause predictable changes in the plaintext, and without an authentication mechanism to guarantee integrity, receivers are unable to detect the changes. Moreover, because an adversary can easily inject messages, the receiver needs to ensure that incoming data originates from a trusted sender.

MAC can be viewed as a cryptographically secure checksum of a message. Computing a MAC requires authorized senders and receivers to share a secret key, and this key is part of the input to a MAC computation. The sender computes a MAC over the packet with a secret key and includes the MAC with the packet. A receiver sharing the same secret key recomputes the MAC and compares it with the received MAC value. If they are equal, the receiver accepts the packet, otherwise it rejects the packet.

There are three types of MACs for practical use, i.e., stream cipher based, block cipher based, and hash function based. Hash function-based MACs (often called HMACs) use a key or keys in conjunction with a hash function to produce a checksum. Among the previously described block cipher modes of operation, CBC can be used to convert a block cipher into a hash function. To do this, CBC is run repeatedly on the input data blocks, and all the ciphertext is discarded except for the last block, which will depend on all the data blocks in the message. This last block becomes the output of the hash function.

11.2.3 Random Number Generator

Random numbers are always indispensable for a number of network security algorithms based on cryptography, such as key generation and entity authentication. Generating a true random number with a true random number generator is a difficult task to perform using digital hardware.

Therefore, for cryptographic applications, it makes some sense to take advantage of the encryption logic available to produce a sequence of pseudo-random numbers that has all the appearance of a random sequence. For example, CBC mode can be used to convert a block cipher into a pseudo-random number generator. The limitation of such schemes is that the generated bit stream is completely determined by the cipher algorithm, the key, and the IV. Moreover, random seeds are necessary as the

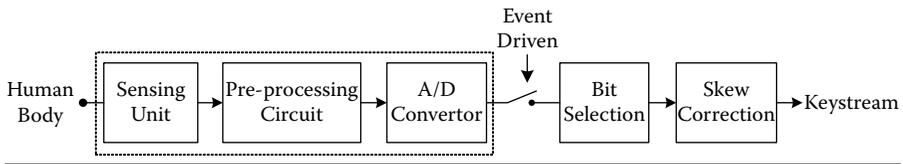


Figure 11.3 A random number generation scheme specific for body sensors.

input of the pseudo-random number generator. The generation of random seeds is a problem to be solved.

Consider body sensors that have abilities to collect time-variable biological signals. An easily implemented scheme for generating random numbers is depicted in Figure 11.3,⁹ where the sensing unit, preprocessing circuit, and A/D converter are essential parts in each sensor device. A binary sequence can be easily generated by chaining the least significant bits of quantified values. Affected by the chaotic body system and random thermal noise, each generated bit is expected to take on the value 0 or 1 with almost the same probability. To ensure that there is an approximately even distribution of 0s and 1s, a skew correction can be performed on the bit stream.

11.3 Key Distribution

Appropriate and successful key distribution is critical to the secure use of every symmetric cryptosystem without exception, where the encryption algorithm requires the same secret key to be used for both decryption and encryption. Communicating parties must possess a shared secret key prior to using any encryption. However, the distribution of secret keys has always been vulnerable to man-in-the-middle attacks.

11.3.1 Key Pre-Distribution Scheme

Because the body area is comparatively small, using a trusted device to distribute an initial key to each sensor of a BSN before deployment is a straightforward scheme. The initial key is randomly generated, and session keys are consequently established through challenge–response communications secured by the initial key. It can be considered as a simple key pre-distribution scheme, compared to various key pre-distribution schemes^{10,11} originally designed for large-scale sensor networks. The process of distribution of initial keys must be carried out in a secure environment, e.g., an IT administration department of a clinical center, where intruders cannot access by any means.

The main disadvantage of this key distribution scheme is that whenever there is a need to add or change a body sensor, the user has to visit the clinical center for

configuring a new initial key. Otherwise, the security of communications between the new sensor and other sensors cannot be ensured. This obviously discourages sharing sensors (mainly wearable sensors) among a group of people, such as family members. Furthermore, thorough key updates can only be made by obtaining a new initial key, because all the communications of consequent keys are secured initially by the initial key. This requires frequent visits to the clinical center, which would be troublesome for disabled users. Therefore, a new kind of scheme without involvement of an extra device is desirable for wireless BSNs.

11.3.2 Biometrics Method-Based Key Distribution

The concept behind the biometrics method is to make use of the bio-channels, over which biological signals are transferred around the human body, for securing communications in a wireless BSN. For example, Cherukuri, Venkatasubramanian, and Gupta¹² proposed using a group of similar random numbers generated from the properties of the human body at different sites (i.e., a biometric trait) to protect the transmission of symmetric keys between communicating parties. The transmitting node binds a cryptographic key with a locally captured biometric trait. At the receiving node, a binding-off process is preceded using the biometric trait captured by itself to recover the cryptographic key.

Because the biometric traits captured at different locations of the body should have slight variations, a fuzzy commitment scheme¹³ is employed to ensure that the difference between two biometric values captured at different locations can be tolerable to a certain degree. The cryptographic key used in the fuzzy commitment scheme needs to be constructed as an error-correcting code, the original goal of which is to enable transmission of a message intact over a noisy communication channel. To use an error-correcting code, we require functions for encoding and decoding of messages. Let $M = \{0,1\}^k$ represent the space of messages. The function $g : M \rightarrow C$, known as a translation function, represents a one-to-one mapping of messages to codewords. In other words, g is the mapping used prior to message transfer. Conversely, g^{-1} is used upon message receipt to retrieve the transmitted message from a reconstructed codeword. The function $f : \{0,1\}^n \rightarrow C \cup \{\phi\}$, known as a decoding function, is used to map arbitrary n -bit strings to codewords. When successful, f maps a given n -bit string x to the nearest codeword in C . Otherwise, f fails, and outputs ϕ . We say that the decoding function f has a correction threshold of size t if it can correct any set of up to t bit errors. More precisely, for any codeword $c \in C$ and any error term $e \in \{0,1\}^n$ with $\|e\| \leq t$, it is the case that $f(c+e) = c$.

Let $K \in \{0,1\}^k$ and $\bar{K} \in \{0,1\}^n$ represent a cryptographic key and the corresponding codeword after a translation function, respectively; let $b \in \{0,1\}^n$ represent the biometric value (also called the original encrypting witness) used for securing

cryptographic keys; let $h: \{0,1\}^n \rightarrow \{0,1\}^l$ be a one-way hash function. The fuzzy commitment scheme $F: (\{0,1\}^n, \{0,1\}^n) \rightarrow (\{0,1\}^l, \{0,1\}^n)$ is formally defined as

$$F(\bar{K}, b) = (h(\bar{K}), \bar{K} \oplus b) \quad (1.1)$$

where \oplus is the bitwise XOR operation. To decommit $F(\bar{K}, b)$ using witness b' , the receiver computes $\bar{K}' = f(b' \oplus (\bar{K} \oplus b))$. If $h(\bar{K}') = h(\bar{K})$, then the decommitment is successful. Otherwise, b' is an incorrect witness that is not close enough to the original encrypting witness in a suitable metric.

To simplify the security analysis, it is assumed that the witness b is drawn uniformly at random from $\{0,1\}^n$. From the construction of F , determining b in its entirety is clearly as hard as determining \bar{K} . Because $|\bar{K}| = 2^k, |b| = 2^n$ ($k < n$), k is a security parameter governing the concealment of the construction. For most applications, a value of about $k = 80$ should provide an adequate level of security. Under common assumptions about hash functions, this security level will require an average of 2^{k-1} hash function computations from an attacker seeking to open a commitment under F .

The most important practical problem of the biometrics method is what kind of biometric traits can be used. Poon, Zhang, and Bao¹⁴ further studied the characteristics of biometric traits potentially to be used as a witness for securing transmission of keying materials. Compared to the traditional biometric traits,¹⁵ the most peculiar feature of these new biometric traits is that they must be random in nature and preferably change with time. A detailed discussion on the nature of these kinds of biometric traits is given elsewhere, and therefore will not be repeated here. Table 11.1 summarizes the common and different properties of the two kinds of traits.

Besides, it is suggested that by solving the question of how nodes of a BSN know that they belong to the same individual, node interference can also be easily avoided. In other words, biometric traits that satisfy the above requirements can be used for dynamic identification of grouped nodes.

To evaluate the performance of biometric traits for such purposes, the false rejection rate (FRR) is defined as the rate of which b and b' measured from the same person during the same period of time was unmatched (i.e., corresponding to a node in the same wireless BSN being rejected by the judging node), and the false acceptance rate (FAR) is defined as the rate of which b matched b' measured from a different person or at a different time (i.e., corresponding to a node of another wireless BSN or an impostor being accepted as a genuine node).

The biometric solution was tested on experimental data, i.e., by using the inter-pulse interval (IPI) of heartbeats as the biometric trait calculated from cardiovascular signals, a minimum half total error rate (HTER), which is equal to $1/2(\text{FRR} + \text{FAR})$, of 2.58% was achieved when the IPIs were coded into a 128-bit binary sequence.

Table 11.1 Comparison of the Properties of Traditional and Newly Proposed Biometric Traits

	<i>Traditional Biometric Trait</i>	<i>Biometric Traits Used for Securing BSN</i>
Common properties	<p><i>Universal</i>: Possessed by the majority, if not the entire population</p> <p><i>Collectable</i>: Easily collected and measured quantitatively</p> <p><i>Effective</i>: Yield a biometric system with good performance; that is, given limited resources in terms of power consumption, computation complexity, and memory storage, the characteristic should be able to be processed at a fast speed with recognized accuracy</p> <p><i>Acceptable</i>: Willingness of the general public to use as an identifier</p>	
Different properties	<p><i>Distinctive</i>: Sufficiently different in any two individuals</p> <p><i>Permanent</i>: Sufficiently invariant, with respect to the matching criterion, over a reasonable period of time</p> <p><i>Invulnerable</i>: Relatively difficult to reproduce such that the biometric system would not be easily circumvented by fraudulent acts</p>	<p><i>Distinctive</i>: Sufficiently different on any two individuals when copies of it are captured simultaneously, even if the copies are captured by different types of biosensors and at different locations of the body</p> <p><i>Time variant but invulnerable</i>: Change with time and have a high level of randomness so that biometric traits captured at different times would not match even if they are obtained from the same individual</p>

In view of threshold decision making, it is quite important to select the threshold value δ because it is not only responsible for FRR and FAR but also the security performance of the fuzzy commitment scheme. Given the original encrypting witness b , any witness b' that satisfies $\|b \oplus b'\| \leq \delta$ should successfully decommit F . It is easy to understand from the structure of F that the threshold value δ should equal t , which is the correction threshold of the decoding function f . Therefore, FA/FR performance of biometric traits must be taken into account when we select the translation function \mathcal{G} .

To this end, another major concern is whether the randomness degree of b is sufficient for cryptographic purposes. Insufficient randomness would open up the possibility for attackers to guess the coded trait and thus obtain the cryptographic key for decrypting the confidential medical data.

11.4 Conclusions

To address the security issues of BSNs, a variety of cryptographic primitives that may be applicable to secure wireless communications within BSNs have been presented in this chapter. Due to severe resource constraints, it is suggested to reuse code by producing ciphertext data, message authentication codes, and even pseudo-random numbers from a single block cipher.

The problem and possible solution to the symmetric key distribution in wireless BSNs has been discussed. To free users from frequent visits to clinical centers or hospitals, a newly proposed biometrics method, rather than key pre-distribution schemes which indispensably need a trusted device for distributing initial keys to network nodes, is suggested for securing transmission of keying materials. Via bio-channels that are unique to each wireless BSN, nodes can connect to each other and securely share symmetric keys while blocking other parties from knowing them.

Though it is well known that one of the most basic conundrums in network security is the constant trade-off between security and usability, it is possible to come up with breakthrough ideas by utilizing characteristics of the target network for security purposes, as shown in the case of wireless BSNs.

References

1. R.S.H. Istepanian, E. Jovanov, and Y.T. Zhang, Introduction to the special section on m-health: Beyond seamless mobility and global wireless health-care connectivity, *IEEE Transactions on Information Technology in Medicine*, 84(4): 405–414, 2004.
2. G.Z. Yang, *Body Sensor Network*, Springer-Verlag, London, 4–12, 2006.
3. S. Park, K. Mackenzie, and S. Jayaraman, The wearable motherboard: A framework for personalized mobile information processing (PMIP), *Proc. 39th Design Automation Conference*, pp. 170–174, 2002.
4. S.D. Bao, L.F. Shen, and Y.T. Zhang, A novel key distribution of body area networks for telemedicine, *Proceedings of the IEEE International Workshop on Biomedical Circuits and Systems* 17–20a, 2004.
5. W. Stallings, *Cryptography and Network Security*, Prentice Hall, 90–98, 2003.
6. B. Schneier, *Applied Cryptography, Second Edition*, John Wiley & Sons, 191–195, 1996.
7. C. Karlof, N. Sastry, and D. Wagner, TinySec: A link layer security architecture for wireless sensor networks, *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, 162–175, 2004.

8. A. Perrig, R. Szewczyk, J.D. Tygar et al., SPINS: Security protocols for sensor networks, *Wireless Networks*, 8: 521–534, 2002.
9. S.D. Bao, Y.T. Zhang, and L.F. Shen, A new symmetric cryptosystem of body area sensor networks for telemedicine, *Proceedings of the 6th Asian-Pacific Conference on Medical and Biological Engineering*, 2005.
10. L. Eschenauer and V. Gligor, A key-management scheme for distributed sensor networks, *Proceedings of the 9th ACM Conf. on Computer and Communication Security*, 41–47, 2002.
11. H. Chan, A. Perrig, and D. Song, Random key predistribution for sensor networks, *Proceedings of the IEEE Symposium on Security and Privacy*, 197–213, 2003.
12. S. Cherukuri, K.K. Venkatasubramanian, and S.K.S. Gupta, Bio Sec: A biometric based approach for securing communication in wireless networks of biosensors implanted in the human body, *Proceedings of the IEEE International Conference on Parallel Processing Workshops*, pp. 432–439, 2003.
13. A. Juels and M. Wattenberg, A fuzzy commitment scheme, *Proc. 6th ACM Conference on Computer and Communications Security*, 28–36, 1999.
14. C.C.Y. Poon, Y.T. Zhang, and S.D. Bao, A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health, *IEEE Communications Magazine*, 44(4): 73–81, 2006.
15. U. Uludag, S. Pankanti, S. Prabhakar, and A.K. Jain, Biometric cryptosystems: Issues and challenges, *Proceedings of IEEE*, 92(6): 948–960, 2004.

Chapter 12

A Survey of Security in Telemedicine with Wireless Sensor Networks

Daisuke Takahashi, Yang Xiao, and Fei Hu

CONTENTS

12.1 Introduction.....	211
12.1.1 Cost-Effective Portable Telemedicine Kit	211
12.1.2 Wireless Telemedicine Models	212
12.2 Biomedical Sensor Network Hardware.....	213
12.2.1 Motes.....	213
12.2.2 Pulse Oximeter.....	214
12.2.3 ECG Sensor.....	215
12.2.4 Motion Sensors.....	215
12.3 Medical Wireless Sensor Networks Deployment Scenarios.....	216
12.3.1 RFID-Based Wireless Telemedicine Systems	216
12.3.2 Health Care on 802.15.4 Beacon-Enabled Clusters	217
12.3.3 Central Trusted Security Servers.....	218
12.4 Attacks on Medical Wireless Sensor Networks.....	218
12.4.1 General Attacks	219

12.4.2	Attacks on Biomedical Sensors	220
12.4.3	Attacks on PAN Coordinators	220
12.5	Security Policies	221
12.5.1	Threats to Clinical Confidentiality and Integrity	222
12.5.2	Designing Security Policies	222
12.6	Security Algorithms	225
12.6.1	Hierarchical Security Architecture	225
12.6.2	Key Generation Enforcing Security Policies	226
12.6.3	Hierarchical Security Architecture	227
12.6.4	Digital Signature Schemes	228
12.6.5	Zigbee Security Application Programming Interfaces (APIs)	229
12.6.6	Security Protocol for Wireless Medical Networks with Multi-Hopping	230
12.6.7	Inter-Cluster Session Key Generation	230
12.6.8	Initialization Vectors (IVs) and SkipJack Algorithm	231
12.7	Conclusions	232
	Acknowledgment	232
	References	233

Telemedicine is a promising technology that can reduce both physical and monetary burdens from patients traveling to distant hospitals in order to have medical consultations. Telemedicine is simply an application of information technology to medical examinations. More precisely, in telemedicine doctors remotely conduct medical examinations with patients by using several forms of information technology, including videoconferencing or digital imaging. In the meantime, wireless sensor networks are progressing rapidly and already benefit other related fields, e.g., cell phones, PDAs, or other Bluetooth and Zigbee devices. Likewise, along with miniaturization of medical devices, integrating wireless sensor networks in telemedicine should provide more portability and unobtrusiveness to the medical devices facilitating comparatively long-term monitoring. Regardless of these benefits, however, due to the nature of radio broadcasting, wireless telemedicine poses risks to expose patient care records (PCRs) to third parties more easily than connected networks, which may cause loss of confidentiality and integrity, and PCRs are basically very sensitive to every person. Thus, this chapter mostly concerns a security aspect of telemedicine with wireless sensor networks. However, before exploring several security protocols, the chapter introduces basic ideas of telemedicine and wireless telemedicine. Then it discusses scenarios of wireless telemedicine, and possible attacks and threats. To address the possible wireless network attacks, nine security policies and security algorithms to protect patients from data leakage are presented in the later sections.

12.1 Introduction

Telemedicine is an emerging technology that largely benefits patient health care areas. Simply, telemedicine is a medical application of information technology enabling patients to have medical consultations outside hospitals by using videoconferencing or digital imaging systems. In its history, it was only a closed circuit TV connection of two remote psychiatric institutes in Nebraska in the United States.¹ Through this connection, doctors on both sides could initiate an audio-visual conversation and discuss each patient's case to improve their medical skills. In some cases, digital imaging and communication were used by the doctors to examine psychiatric patients. Currently, telemedicine is used in teleradiology, teler dermatology, or telepsychiatry.

Because of the recent progress of both electronics and information technology, telemedicine is currently not only a technology to facilitate remote medical conversations but also to utilize a variety of biomedical sensors to capture several critical vital signs. Usually these vital signs are sent to doctors very securely through the Internet, e.g., using IPsec.

For example, the process of medical examinations generally follows these steps.¹ At a telemedicine facility which deploys one side of a two-way videoconferencing system along with computers and biomedical sensors, practitioners who may not be doctors are permanently stationed. Basically these practitioners locally perform medical examinations and capture patients' vital signs under the supervision of doctors at hospitals. During the examination, two-way conversations may be conducted between a doctor and patient or the doctor and a practitioner by using a videoconferencing system. Depending on the system specification, captured vital data are transferred to the doctor either in real-time or after the examination as a compressed form (store-and-forward method).¹ According to these data, medical diagnoses are reported to the patients in a later visit to the telemedicine facility.

Telemedicine is usually conducted at remote facilities supplementally as the second or third visit.

12.1.1 Cost-Effective Portable Telemedicine Kit

Cost-effective portable telemedicine kits were introduced to developing countries. A portable telemedicine kit consists of a digital stethoscope, ECG recorder, medical imaging system, and blood pressure and temperature measurement devices.¹ As electronics in the medical field have progressed, it is possible for biomedical sensors to become lighter and smaller. At the same time, biomedical sensors become more sophisticated and lower priced. From these benefits, the MIT Media Laboratory experimentally developed and deployed a prototype of low-cost portable telemedicine kits in developing countries as part of the Little Intelligent Communities (LINCOS) project, which collaborated with the Costa Rica Foundation

for Sustainable Development and other educational institutions.¹ Usually developing countries suffer a shortage of doctors as well as hospitals. Therefore, patients in these countries basically suffer the physical and monetary burdens of traveling around the country to see doctors. However, from their economic conditions, these countries may not easily agree to increase the number of hospitals. Hence, instead of adding a few new hospitals, it is a rationale that they rather choose to deploy as many telemedicine facilities, which generally cost much less than hospitals, as they can. For example, a prototype of a portable telemedicine kit in the LINCOS project only cost around \$8,000 although it consisted of a digital stethoscope, an ECG recorder, a medical imaging system, and blood pressure and temperature measurement devices.¹ Moreover, in Costa Rica telemedicine centers were made of commercial ISO shipping containers, which were very low cost but equipped with security and structural integrity.

12.1.2 Wireless Telemedicine Models

Another type of telemedicine system utilizes wireless network connectivity so that a wireless personal area network (WPAN) is created around a patient between biomedical sensors and a personal server that runs on a PDA, shown in Figure 12.1.² Besides the WPAN, in this configuration personal servers can establish a wireless local area network (WLAN) to the nearest wireless access point that is deployed around the building to connect to the Internet. The objective of wireless telemedicine is to enhance unobtrusiveness of devices so that patients feel little of the medical equipment during examinations. Therefore, in general, wireless telemedicine is suited for long-term monitoring that usually takes an entire day, e.g., intensive cardiac monitoring.

In the wireless telemedicine configuration, some architecture allows biomedical sensors to communicate with each other so that they can transmit data to other sensors. Some biomedical sensors may be located further away from a personal

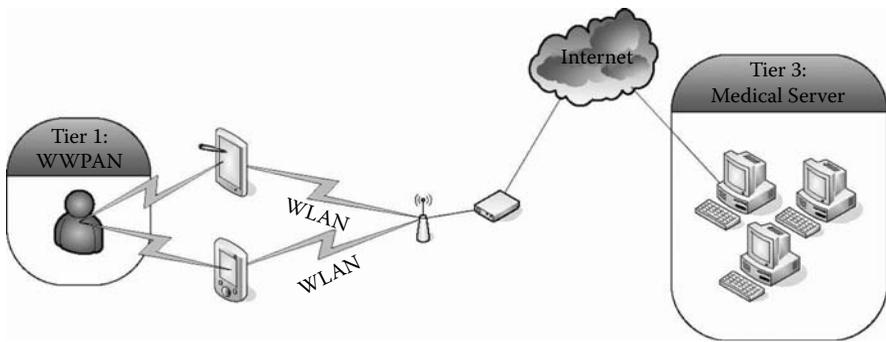


Figure 12.1 An example of a wireless telemedicine system with three-tier architecture: A wearable wireless personal network (WWPAN), personal server, and medical server.

server, and the distance does not allow one-hop transmission of data to the personal server. In that case, these biomedical sensors are required to initiate multi-hop data transmission that relays the data to the personal server by using other intermediate sensors.³ An example of wireless telemedicine architecture utilizing multi-hopping to relay a patient's vital signs to personal servers is presented in Welsh³: (1) medical places vital sensors on disaster victim, (2) medical issues queries for patient vital signs, and (3) patient sensors send data using multicast routing.

When adopting the multi-hop strategy, wireless sensor networks can have the benefit of preserving energy consumption as well as expanding the distance of the data transmission.⁴ Examples of this telemedicine architecture is CodeBlue from Harvard University, which makes use of the multi-hop strategy and the Adaptive Demand-Driven Multicast Routing (ADMR) protocol, which can establish a route from publishers to subscribers.^{5,6}

Regardless of a lot of benefits from the wireless sensor networks, wireless telemedicine still suffers several shortcomings in the domain of the size and weight of biomedical sensors to the network security.⁶ Regarding network security, because wireless sensor networks largely rely on radio broadcasting, this poses the potential risks of disclosing the patient care record (PCR) to a third party, which results in the lack of confidentiality.⁴ However, for current wireless sensor networks, it is still challenging to adopt traditional encrypting techniques to these tiny wireless biomedical sensors due to their limited energy and computational power.⁶

The objectives of this chapter are to explore the current network security issues in the telemedicine field or possible attacks toward the systems, and to survey techniques to solve these issues. The chapter is organized as follows. Section 12.2 introduces biomedical sensor network hardware that has been employed in the wireless telemedicine research by integrating a miniature mote inside to communicate with other sensors. In Section 12.3, possible scenarios of wireless telemedicine architectures are introduced, and within these scenarios, the kinds of attacks that may be launched against them is discussed in Section 12.4. In Section 12.5, security policies that address what kinds of data must be secured by the application are discussed. At last, a couple of security solutions for these attacks are introduced in Section 12.6.

12.2 Biomedical Sensor Network Hardware

This section presents three types of biomedical sensor network hardware that are recently involved in wireless telemedicine. In addition to these three sensors, miniature motes or smart dusts are explained in this section.

12.2.1 Motes

"Mote" is derived from "remote," and it represents miniature wireless transceivers. Regardless of the size and weight, motes have a simple microcontroller and memory,

a low-power radio transmitter, and a couple of sensors so that they can sense, process, and transmit various data within the range. For example, the UC Berkeley and Intel research group developed the MICA mote, a low-power, miniature wireless transceiver.^{3,5,7,8} This palm-sized mote can embed an 8-bit microcontroller, 40 kbps radio, and local storage with low battery consumption, which keeps it working with two AA batteries for 5 to 6 days while in the save mode, over 20 years in nonactive state. In addition, the MICA mote employs a 433- or 916-MHz radio whose maximum bandwidth is 76.8 kbps with a 100-m transmission range.

Likewise, the Harvard University research group developed a tiny prototype mote named Pluto.⁵ Although compared to the MICA mote the Pluto mote sacrifices the expandability and battery durability, it realizes a lighter weight and smaller size than the MICA mote. Thus the Pluto mote is housed in 57 × 36 × 16-mm OEM plastic enclosure, and weighs only 30.5 g, while a tiny rechargeable lithium polymer battery on motes can last around five hours.

Miniature motes generally integrate an on-board bio-amplifier, which can be connected to an electromyograph (EMG) or an electrocardiograph (ECG) sensor, and amplify patients' vital signs so that local practitioners can observe them electrically on the display of the personal server.⁹ In addition to EMG and ECG sensors, because some applications integrate accelerometers into a mote, practitioners can monitor human movement from all three dimensional axes, and the feature can be used for computer-assisted physical rehabilitation.

12.2.2 Pulse Oximeter

The pulse oximeter measures the blood oxygen saturation (SpO₂), heart rate (HR), and plethysmogram waveform in a noninvasive way. Basically a patient is only required to insert his or her index finger or ear lobe into a plastic housing (a plastic box).⁶ Inside of the housing, it projects the infrared and near-infrared light on the index finger, and it observes how much these lights are absorbed by hemoglobin in the blood vessels. Measurement of the amount of light absorption can be translated to the level of SpO₂. Besides the SpO₂ level, a light absorption pattern caused from the blood vessel expansion and contraction can be interpreted to HR.⁶ Thus, patients suddenly having these parameters change must be taken care of immediately.

The Harvard University research group made up a mote-based wireless pulse oximeter from a commercial pulse oximetry sensor module.^{6,7} In designing a wireless pulse oximeter, the Mica2 or MicaZ mote (manufactured by Crossbow Technology, Inc.) was used to integrate a pulse oximetry interface board, and the standard finger housing was connected to the mote by a serial port. Vital signs captured by the finger housing are securely transferred to the mote and then time-stamped, encrypted with a unique node identifier, and transmitted to receiving motes. Because receiving motes are usually integrated in PDAs, practitioners can immediately obtain patients' conditions from their own PDAs.

12.2.3 ECG Sensor

Currently most of the telemedicine systems involve electrocardiographs (ECGs). ECGs are used to detect patient's undesirable heart rhythms, blood and oxygen supplies, as well as the excessively tensed heart muscle.¹ Since the heart muscle is motivated by electricity very similar to an electrical mechanical pump, this activity can be captured by attached electrodes put on the tips of leads that are connected to an ECG recorder.¹⁰ Commonly used ECGs involve 12 to 15 leads with electrodes.⁶ Usually electrodes are fixed on a patient's chest, both arms, and the right leg, and sense the cardio-rhythms and electrical impulses in a short time period.¹⁰

However, attaching a number of electrodes deprives patients of freedom to move and they often feel uncomfortable during the examinations. Therefore, this type of cardiac examination is basically carried out in no more than 30 minutes. However, a short time data sampling may not be enough for physicians to detect cardiac anomalies where problems only emerge once or twice a day.⁶ Moreover, because the contractions of other muscles also produce electricity, 12- to 15-lead ECGs are so sensitive that the results often confuse physicians.¹¹ To avoid these difficulties, continuous ECG telemetry is additionally employed as an intensive cardiac monitoring.⁶ Unlike the short-time electrocardiography, the continuous ECG telemetry utilizes only 2 to 3 electrodes. Thus patients endure fewer physical burdens than with the short-time electrocardiography. Generally these cohesive electrodes are fixed on the right infraclavicular, the left infraclavicular, and possibly the left chest, just below the mammilla, to record three angles of the heart's electrical activity.¹¹

Like wireless pulse oximeters, mote-based ECGs also exist. Apparently the wireless ECGs enhance the portability of the sensors and enable ambulatory cardiac health care in homes or work environments. Although the first generation of the wireless ECG sensors could be used only within hospitals, the next generation could allow the sensors out of hospitals.¹¹ These ECG sensors periodically sent vital signs to a hospital through the Internet. Furthermore, to establish real-time data collection, NASA and Case Western Reserve University's Heart & Vascular Center developed the Arrhythmia Monitoring System (AMS) that enabled real-time ECG data collection from mobile patients through digital, packet-switching telephony services in metropolitan areas.¹¹

12.2.4 Motion Sensors

Primary objectives of motion analysis systems are to record muscular activities and limb movements. In other words, physical rehabilitations from brain damages or Parkinson's disease are aided by these systems.⁶ Brain damages, such as internal brain bleeding or a critical shortage of the brain blood, may result in temporal or terminal functional termination of the partial brain, and these disorders eventually result in the partial disability of the body movement. Parkinson's disease is also a physical impediment which results in unintended hand shaking or a tremor

throughout the body. Parkinson's disease may be initiated with brain injuries or infections. A n e x a m p l e o f m o t e - b a s e d m o t i o n a n a l y s i s s y s t e m s i s p r e s e n t e d i n Welsh³: Telos mote by UC Berkeley and Moteiv Inc. integrates accelerometers, a gyroscope and EMG circuitry, and motion analysis systems measure relations among each body segment attaching sensors, angular velocity, and electrical field during a body movement.

Unintended physical movements from these diseases should be accurately monitored on a regular basis so that doctors can properly and timely prescribe medicines for the diseases.

Usually the motion analysis systems accompany a variety of sensors that are fixed on particular body segments and measure relations among each segment, angular velocity, and electrical field appearance during the movement.^{6,12-14} More precisely, the motion analysis systems typically accompany three types of sensors: accelerometers, gyroscopes, and surface electrodes for electromyography (EMG).

Like other physiological sensors, unobtrusiveness is ideal for long-term motion analysis. In the use of a mote-based motion analysis system, sensors are placed on the upper arm, lower arm, and torso.⁶ A very small module incorporated with accelerometers, a gyroscope, and EMG unit amplifies electrical signals to be captured, while it reduces irrelevant noises from the targeted frequencies. These amplified signals are next sent to motes integrating respective sensors, and they transmit the data to user devices, such as PDAs. Because multiple modules are working together on a patient's body, these data must be properly time-stamped for the synchronization.

12.3 Medical Wireless Sensor Networks Deployment Scenarios

In the previous section, mote-based biomedical sensors are explained for the use of wireless telemedicine. In this section, deployment scenarios are briefly discussed. For the sake of generality, when referring to biomedical sensors, either one or all kinds of the aforementioned sensors can be used in the system if any particular sensors are mentioned.

12.3.1 *RFID-Based Wireless Telemedicine Systems*

The first scenario consists of a number of mote-based biomedical sensors that are put on a patient's body in a hospital; a number of personal servers that could run on PDA, tablet, or laptop PC; and the mote interface board (MIB), which is a gateway between the personal area network and the Internet.⁴ Biomedical sensors periodically collect and send a patient's vital signs to the personal servers. When requested, a sample waveform or other vital data are displayed on the practitioner's PDA. For example, Figure 12.2 shows a sample trace from an ECG sensor.

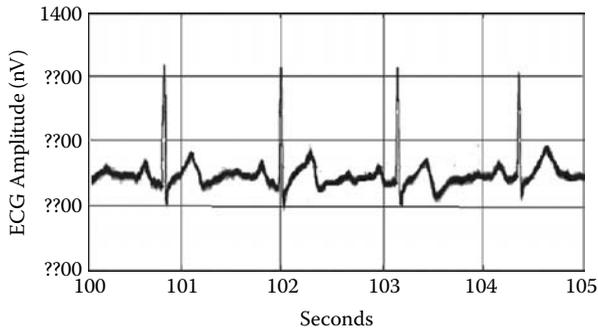


Figure 12.2 A sample trace from an ECG sensor.

In addition, when a sensor detects an anomaly of the patient's physiological condition, it immediately calls any personal servers even though they are located remotely from the sensor.

Because biomedical sensors may be deployed away from the personal servers and any personal server cannot catch the transmitting data directly from the sensors, in this model sensor nodes utilize multi-hop wireless data transmission in order to relay the data to the personal server hop-by-hop.⁴ In addition, this multi-hop wireless transmission also benefits the energy consumption of the sensors because the sensor nodes do not always require performing a long-distance data transmission direct to the personal server.

Moreover, each mote integrates an ultra-high frequency (UHF) RFID reader in itself.⁴ So the motes can read the radio signals from a tag attached to a medicine bottle. This configuration is usually used for controlling the amount of medicine that an elderly patient should intake.

In the scenario in Hu, Kumar, and Xiao,⁴ both the personal servers and the MIBs or the hospital network gateways install security software so that any communication between them is encrypted.⁴

12.3.2 Health Care on 802.15.4 Beacon-Enabled Clusters

In the second scenario, the architecture of the health care wireless sensor networks consists of a number of biomedical sensors, personal area network (PAN) coordinators, room access points, and a central database.¹⁵ Sensors are fixed on the patient's body segments and organized into several sets of sensors, namely clusters, by PAN coordinators. Communications between PAN coordinators and biomedical sensors are established by the IEEE 802.15.4 standard. Therefore, some security suites in the 802.15.4 can be employed for these wireless network communications. Usually each PAN coordinator resides on each patient bed and plays a role of patient security processor (PSP) as well.

Every biomedical sensor participating in a network is supposed to be authenticated and trusted before data transmission is initiated. Therefore, upon communication, encrypted data are first sent to a respective PAN coordinator.¹⁵ The PAN coordinator next sends the data to a central database via a room access point. In general, the room access point is wiring connected to the central database, which is remotely located in the hospital. Because authenticated medical personnel use a PDA or tablet PC running a personal server, they can fetch or update patient's care records (PCRs) from the database by using these devices.

This architecture employs the IEEE 802.15.4 wireless network standard with star topology between a PAN coordinator and biomedical sensors.¹⁵ Although the 802.15.4 standard provides useful security services such as access control lists, data encryption using pre-stored key, message integrity code, and message freshness protection, it is not enough to provide all the security expected in this scenario (e.g., procedures for key management and device authentication). Therefore, the rest of the security has to be compensated by the application layer, such as security application programming interfaces (APIs) by the Zigbee Alliance.

12.3.3 Central Trusted Security Servers

The aforementioned medical wireless sensor networks often involve the Central Trusted Security Server (CTSS) that, like the central database, usually resides in a physically trusted location.¹⁵ Basically the CTSS is responsible for organizing a security protocol between clinician's and patient's devices, and calculating encryption keys for the devices in each cluster. In addition, the CTSS is responsible for enforcing security policies of the medical information systems on the use of both biomedical sensors and personal servers, so any device must not be used for abetment of a breach of confidence; if they are used by malicious adversaries, any patient's privacy must not be vulnerable to them.¹⁵ In storing medical records on the central database, the CTSS performs encrypting processes and puts time-stamps on all patient records, so without proper decryption, patient care records on the database are useless to anyone. At the patient level, patient security processors (PSPs) or personal servers (PSs) share private symmetric keys with all the sensor devices of their own clusters, and these private symmetric keys are provided by the CTSS.

12.4 Attacks on Medical Wireless Sensor Networks

In the wireless sensor network point of view, wireless telemedicine is also inevitably exposed to risks of disclosing patient privacy or patient care records (PCR) to

a third party.^{4,15,16} Therefore, it requires some encryption techniques in order to hide its transmitting data. However, currently from its limited resources, complex encryptions are not applicable to this technology, but lightweight encryption is still challenging in wireless sensor networks.⁴ Before discussing these challenges, we present possible attacks that threaten medical wireless sensor network scenarios. Then, the securities of these attacks are addressed.

12.4.1 General Attacks

From the perspective of the general wireless sensor network, it is possible for medical wireless sensor networks to suffer general attacks, as presented in the literature. Mišić, Amini, and Khan¹⁵ summarize these attacks as follows:

1. *Sybil attack*: Malicious nodes create one or more fake identities so that they can increase the probability to be chosen as route nodes. This attack also degrades fault-tolerance functionality of the sensor networks because only one route acts as if there are many routes by using multiple identities.¹⁷ More precisely, even though packets choose another way to avoid a failure, it might be the same route in which nodes are also fabricated by the same malicious node.
2. *Sinkhole attack*: Malicious nodes attract the neighboring nodes by advertising high quality of connectivity that may be made up so that all the traffic around them is attempting to pass through these malicious nodes.¹⁷
3. *Black hole/gray hole attack*: After establishing routing paths, malicious nodes intentionally drop all the traffic (black hole attack), or only selected traffic (gray hole attack) that is received by them.
4. *Wormhole attack*: Two or more laptop-class attackers establish network tunnels by using out-of-bound channels that usually involve low latency and high bandwidth.¹⁷ Thus these fake network topologies attract neighboring nodes to select them as part of routing paths.
5. *Sleep attack*: Battery life is one of the primary concerns in wireless sensor networks. In a sleep attack, a malicious node keeps interacting with the neighboring nodes to force them to be in the active mode and consume the battery life more quickly than usual.¹⁸ Because these interactions seem to be legitimate, it is difficult to detect these malicious nodes.
6. *Fairness attack*: Malicious nodes continuously send some kind of requests to a coordinator whereby the other nodes suffer receiving unnecessary results from the coordinator, which consumes the total network battery life of the cluster.
7. *Denial of service (DoS) attack*: Malicious nodes continuously retransmit packets or false data requests whereby the performance of the victim slows down, the services for the other nodes temporally stop, and the victim consequently consumes battery life rapidly.

12.4.2 Attacks on Biomedical Sensors

In medical wireless sensor networks, biomedical sensors basically reside on the patient's side for a long time, but in general, these biomedical sensors are quite vulnerable to physical attacks. Therefore, the exposure of the sensors to the external world endangers them to have the IDs and master secret key stolen by adversaries.¹⁵ As a result, the adversaries will be able to steal a new shared secret key that is updated in the next key distribution. In addition, due to preservation of the packet header space, some application determines to use 16-bit-long node IDs instead of 64-bit IDs. In that case, it is comparatively easy for malicious adversaries to make up the fake identities so as to perform Sybil attacks from these identities.¹⁵

When medical wireless sensor networks employ carrier sense multiple access with collision avoidance (CSMA/CA) with star topology, DoS attacks or fairness attacks are possibly launched,¹⁵ whereby malicious nodes continuously send legitimate packets to a coordinator in order to dominate the bandwidth and keep the network busy to prevent the other nodes from communicating with the coordinator. As a result, the availability of the network is largely degraded. A solution of this attack is that coordinators record occurrences of communication with each node, and deny communication from the excessively accessing nodes.

In some applications, when a coordinator initiates a key exchange process, it also notifies every node in the cluster about what nodes will change their shared keys, but this key distribution mechanism is vulnerable to external adversaries.¹⁵ Thus if malicious nodes occasionally receive the IDs of cluster members that are about to change their shared keys, they may send key distribution requests to the coordinator by using these IDs before the legitimate nodes initiate the requests. As a result, before the key exchanges, the legitimate nodes are required to receive the keys, which results in some algorithm problem. Moreover, only a simple collision of the requests may disturb the completion of the key exchange process of legitimate nodes. This attack may be simply defended by authentication of the key exchange requests.¹⁵

When an application adopts a sleep mode to conserve energy of biomedical sensors, it is possible for malicious nodes to pretend to be legitimate nodes while true nodes are sleeping.¹⁵ In that case, it is even more difficult to detect these malicious nodes. Therefore, the coordinators are required to keep track of how long each node is totally in the active mode.

12.4.3 Attacks on PAN Coordinators

Like biomedical sensors, due to the nature of radio broadcasting, the PAN coordinators are also vulnerable to external attack. In the IEEE 802.15.4 standard, because there are only 16 radio frequency channels it is comparatively easy to snoop on the channel that is currently used for the wireless network.¹⁵ Thus, upon finding

out a particular channel, reading messages from beacon frame is also comparatively easy for malicious adversaries. Messages may contain node IDs and the number of nodes. Therefore, adversaries may be able to compute the sleeping time distribution, and as a result, they launch the aforementioned attacks to other nodes by using the information.

The key exchange periods can be sneaked by external adversaries by stealthily sensing the coordinator's notifications.¹⁵ Before initiating key exchange processes, the PAN coordinators broadcast messages to every node of what nodes are required to change their shared secret keys. These messages are easily eavesdropped by malicious nodes. Therefore, by recording these periods, adversaries predict the next key exchange session, and launch some attacks to other nodes.

To avoid these two attacks, messages in beacon frame must be carefully encrypted so that even if they are eavesdropped by malicious nodes, they cannot understand what the messages mean.¹⁵

In addition, location discovery algorithms of the PAN coordinators are also disturbed by external attackers.¹⁵ When determining the location, a PAN coordinator calculates its position in relation to the neighboring PAN coordinators that have GPS sensors. However, because this algorithm largely relies on the strength of signals of the neighboring PAN coordinators, if attackers have higher-power transmitters, these signals are easily disturbed; consequently, this PAN coordinator may report a wrong location to the base station. Moreover, another malicious node may report its position as if it were located at the original position of this PAN coordinator, which now has a wrong position report while the correct PAN coordinator is eliminated from the network.

Moreover, a wormhole-like attack can be launched by two corrupted PAN coordinators, and it is simply performed by one corrupted coordinator always sending its received packets to the other coordinator, creating a wormhole tunnel.¹⁵

Furthermore, the PAN coordinators reside relatively close to biomedical sensors or patients due to their limitation of radio transmission. However, because the PAN coordinators are usually vulnerable to the physical attacks, this situation endangers the coordinators to be physically destroyed or have information of the network or shared secrets as well as patient care records (PCRs) that are currently transmitted stolen by the malicious adversaries.¹⁵

12.5 Security Policies

In the previous section, we learned about possible attacks on wireless telemedicine in terms of wireless sensor networks, on biomedical sensors, and on PAN coordinators. Before establishing secure wireless sensor network models, this section investigates security policies in clinical information systems. In the literature, security policies basically deal with what must and must not be done with aggregated patient care records (PCRs), and they need to be defined in

order to protect confidentiality and integrity of the PCRs in medical information systems.^{16,19}

12.5.1 Threats to Clinical Confidentiality and Integrity

First, this section deals with matters that threaten clinical confidentiality and integrity. Therefore, this section explains the kinds of threats posed to patients' privacy, who targets patient privacy for what purpose, as well as the reasons why clinical confidentiality and integrity are protected.

According to Anderson,¹⁹ in terms of centralized information systems, many organizations' experience recently showed that things to protect information privacy or confidentiality from are treacheries of corrupted insiders. It is because in some cases people inside of organizations are not restricted to access the organizational information with any authentication. For example, in the United Kingdom, a teller in most of the big banks is free to access any customer's account.²⁰ Thus adversaries may bribe the tellers to sell them customers' account information for only £100 or so. Likewise, PCRs are sold for no more than £150 in the United Kingdom.^{20,21}

In order to seek information system efficiency, data need to be gathered up from several separated locations and stored in a central location. However, this centralized database system endangers more PCRs of being disclosed eventually.^{19,22}

Regarding confidentiality, Anderson¹⁹ reports worse experiences in the United States. For example, a list of all patients diagnosed with cancer was stolen by a banker serving on a state health commission; later, this information was sold to the patients' creditors.²³ Another experience shows that these data are used by 500 major U.S. companies, more than half of which refer to PCS to determine whether or not to hire job candidates.²⁴

On the other hand, in terms of lack of integrity, database attacks are used for falsifying clinical information. For example, in medical malpractice, clinicians may conceal relevant information from the patients or their relatives by altering the PCRs,²⁵ and it is also possible to alter prescriptions.²⁶ In addition, in Spain, where health cards are also used as bank cards,²⁷ adversaries may launch attacks to alter the information of the cards so that they can eventually get some monetary benefits from them.¹⁹ The cards may also be used for personal identification, which causes other problems.²⁸

12.5.2 Designing Security Policies

To protect from risks of losing confidentiality and integrity, security policies must be designed carefully. Regarding a security policy model, in order to keep confidentiality of PCRs, Anderson¹⁹ summarizes nine security principles that are based on those regulated by the General Medical Council^{29,30} and the British Medical Association,³¹ and also employed by military and banking systems in England.¹⁹ These

principles mainly concern the following issues: access control list; record opening; control of access control list; consent and notification; persistence; attribution; information flow; aggregation control; and trusted computing base.

1. *Access control list:* The first principle suggests that medical information systems should accompany an access control list to each datum of PCRs in order to restrict their access only to authenticated people who are on the lists.^{16,19} In particular, these authenticated people include patients themselves, the responsible (principal) clinician, and one or more referring clinicians, and the data access includes observations and modifications of data. Data of PCRs may have different levels of confidentiality such that restriction of accessing data relies on how important the data are to the patients.¹⁹ Therefore, according to how important the data, subjects accessing the data can vary in such a way that highly sensitive data must strongly restrict access to only few responsible personnel. In practice, however, determining levels of confidentiality of PCRs is so difficult that only patients themselves can make an appropriate decision. For example, for people in AIDS campaigns, HIV status may be considered open to the public; for Jehovah's Witnesses, blood transfusion may be scandalous and considered to be hushed up.³² Because access control of data largely depends on each patient, any changes from the default lists should be made by the respective patients' own decision.¹⁹
2. *Record opening:* While the first principle defines that every datum should have its own access control list, the second principle restricts the data access only to the personnel who are on the list.¹⁹ Of course, there are various records having different sensitivity, and access control lists should reflect these levels of sensitivity. For example, any clinician in the practice is allowed to access a general PCR of a patient while only a general practitioner can access highly sensitive PCRs.¹⁹ Therefore, in each of these PCRs, there must be an access control list in order to distinguish personnel accessing the data. Thus, by default, the second principle restricts access of PCRs only by a clinician and the respective patient, and possibly the referring clinicians.
3. *Control of access control list:* In the third principle, it is suggested that only a responsible clinician can maintain the access control list, so only he or she can add or delete other health care professionals on the access control lists.^{16,19} To achieve this, exactly one clinician on an access control list must be chosen as a responsible clinician (an administrator) of the list.
4. *Consent and notification:* Every execution of opening PCRs or a addition of people to access control lists must be reported to and, if necessary, permitted by the patients each time.^{16,19} This notification includes the name of personnel who affect the PCRs. Therefore, without the patient's permission, even a responsible clinician cannot affect a PCR as well as add other health care professionals to an access control list. In addition, this principle may protect patients from fraud.¹⁹ For example, some person may pretend to be another

person in order to receive expensive medical treatments. However, every time someone opens a PCR the patient must be informed of it. Therefore, the patient can be notified when someone else tries to access their data through a back door, and this can help detect the fraud.

5. *Persistence:* The fifth principle regulates the persistence of PCRs. Therefore, this principle promises eight years of preservation of most primary PCRs. In addition, PCRs of patients who have cancer must be preserved until the patients die while genetic diseases are longer than these.¹⁹ Thus, during these periods, any PCRs must be preserved and not be deleted. However, there may be a problem when patients demand the deletion of the PCRs by withdrawing their consent. Thus, a possible solution of this problem is to transfer the PCRs to another clinician whom the patient wants rather than delete them. Moreover, simple errors of data must be also preserved by appending the right values instead of overwriting.
6. *Attribution:* Accountability of data access to the respective patients must be established so that how, when, and which authenticated person accessed a PCR can be audited every time.^{16,19} With this functionality, both read and write accesses are recorded, and particularly auditing reading data is employed to prevent breaches of confidence. Moreover, for deletion, when and who deletes data are also monitored by the system, which helps trace intentional deletion of data. However, even if adversaries remove data, the system must allow all the materials deleted to be recovered from the attribution.
7. *Information flow:* Principles even deals with data concatenations from the perspective of access control lists. From this principle, when there are two records A and B for the same patient, if the members of A's access control list entirely includes B's members, then record A may be appended to record B.¹⁹ In this multilevel security mechanism, then only the personnel in the intersection of the access control lists can take over the appended records.
8. *Aggregation control:* As previously mentioned, aggregating PCRs into one centralized place increases risks of disclosing many data at the same time. Therefore, some methods of controlling aggregated data are required. For example, clinicians at a safe haven may inevitably come out on millions of access control lists. Therefore, these people must be notified to patients when they are added to the access control lists of PCRs of those patients.¹⁹ In addition to the notification, access control lists should be preserved outside of their practices, and someone responsible should keep an eye on them. Also a typical method that some institutions are already employing is declaration of penalties, such as dismissal, resulting from unjustified access. However, enforcement is difficult and results are occasional.
9. *Trusted computing base:* To pursue effectiveness, systems will be hierarchically structured, and the subsystems must be obligated to the aforementioned principles.¹⁹ In addition, practical efficiency of the system will be measured

quantitatively by independent professionals. Regarding the measurement, Anderson¹⁹ recommended that the level of measurement should be based on how many people are involved in the system and may endanger PCRs.

12.6 Security Algorithms

On the basis of security policies described in the previous section, in this section several security algorithms for medical wireless sensor networks are introduced. These security algorithms include shared key distribution, secure hash algorithms, and a digital signature scheme. In addition, some algorithms presented in this section are designed particularly for the architecture utilizing multi-hopping. Meanwhile, another algorithm directly applies security application interfaces (APIs) from the current wireless network standards, e.g., IEEE 802.15.4.

12.6.1 Hierarchical Security Architecture

A security infrastructure employing a three-tiered tree model is presented.¹⁶ Its lowest tier is a WPAN around patients. The WPAN consists of biomedical sensors and a personal server or patient security processors (PSP) managing a cluster of sensors creating star topologies. A couple of WPANs are also coordinated by an access point in a patient room, and access points are connected to a hospital network, which includes the central database and Central Trusted Security Server (CTSS).¹⁶ Because medical wireless sensor networks and information systems typically form a hierarchical structure, it had better employ different shared keys and different encryption protocols to each tier, and apply them according to the importance and sensitivity of the data as well as the resource limitation, such as the memory size or computational power.¹⁶

At the PAN level that is in the bottom tier of the hierarchy around patients, because of the resource limitation Mišić and Mišić¹⁶ suggest that a security protocol should employ a symmetric key encryption or private key encryption, such as Triple Data Encryption Standard (3-DES) or Advanced Encryption Standard (AES) rather than more complicated asymmetric key encryption or public key encryption. Compared to asymmetric key encryption, symmetric key encryption is more efficient in terms of usage of physical resources, and consequently lightweighted. On the other hand, because asymmetric key encryption basically implies much more computations and high payloads, limited resources of the wireless sensor networks suffer many disadvantages resulting in the systems not working well.

On the other hand, encryption of higher level such as security between PSPs and room access points may be carried out in the use of the pre-equipped security scheme or Security Architecture for Internet Protocol (IPSec) such as the Encapsulated Security Payload (ESP).^{16,33}

12.6.2 Key Generation Enforcing Security Policies

Concerning the security policies in the previous section, a key generation scheme can be designed. In particular, Mišić and Mišić¹⁶ designed a key generation scheme by taking into account the policies concerning the access control list, record opening, and managing the access control list from the security policies. This key generation scheme can be employed for encrypting operations among the PAN-level devices.

An objective of this key generation is that only certain PAN devices can participate in the key generation process and only the devices that have shared keys can access patient care records (PCRs) of a particular patient's groups. That is, these shared keys are considered as access control lists for particular groups.

Security infrastructure in this model consists of biomedical sensors, a patient security processor managing a group of sensors, clinician's personal servers, and the Central Trusted Security Server (CTSS).¹⁶ According to security policies in the previous section, no PCR can be viewed and also modified by any person who is not on the patient's records access control list.¹⁹ Moreover, if any change is carried out on the PCR or the access control list, the patient must be notified of the changes. On the basis of these policies, Mišić and Mišić¹⁶ designed a key generation protocol that is based on the Diffie–Hellman protocol,³⁴ which enforces the following two rules¹⁶:

1. The key generation must not be carried out unless all the members on an access control list take part in the key generating session.
2. Without participation of the patient, no key generation and no access on the record can be carried out. This key is used for user authentication when accessing a PCR.

In addition, the CTSS described in the previous section is introduced in this model. Again, the CTSS is responsible for managing a security protocol between patient's and clinician's devices and enforcing security policies on all the devices and personnel participating in the system. Thus, the CTSS participates in key generation sessions, and each key generation is carried out by the CTSS interacting with patient's and clinician's devices.¹⁶ The key generation follows the next steps.¹⁶

As described before, this key generation is based on the Diffie–Hellman protocol. For the sake of simplicity, we assume that three people, such as responsible clinical personnel, a referring clinician, and a patient, participate in this key generation, and IDs are represented as ID_1 , ID_2 , and ID_3 . In addition, inverse values Z_1 , Z_2 , and Z_3 , respectively, are used for proofs of participation of key generations by each person, e.g., $ID_i Z_i \bmod (p - 1) = 1$, where p is a large prime number. Moreover, another large prime number g that is a primitive mod p is prepared in this scheme.¹⁶ Basically, this scheme is carried out by a sequence of random number exchanges by a peer-to-peer network.

1. The first member, usually the responsible clinician, generates a large random number x_1 , calculates $g^{x_1} \bmod p$, and sends this result to the second participant. Meanwhile, $g^{Z_1} \bmod p$ is also sent to the second member for the sake of proof of participation.
2. Upon receiving a number that is sent from the first member, the second member, who is the referring clinician, first generates a large random number x_2 , computes $g^{x_2} \bmod p$, and sends the result and a participation proof to the third member.
3. Likewise, the third member, who is the patient, generates a large random number x_3 , computes $g^{x_3} \bmod p$, and sends this back to the first member along with a participation proof.
4. In the next loop, according to a number from the third member, the first member calculates $(g^{x_3} \bmod p)^{x_1} \bmod p = g^{x_3 x_1} \bmod p$, and sends the result to the second member along with a proof calculated in the same manner as the key generation process, resulting in $g^{Z_3 Z_1} \bmod p$. In addition, according to the received numbers, the same computations, which are $(g^{x_1} \bmod p)^{x_2} \bmod p = g^{x_1 x_2} \bmod p$ and $(g^{x_2} \bmod p)^{x_3} \bmod p = g^{x_2 x_3} \bmod p$, are done by the second and third members, and the results are sent to the third member and the first member, respectively, while proofs of participation are sent to each member.
5. At last, according to these received numbers, each member can calculate the result $g^{x_1 x_2 x_3} \bmod p$ and $g^{Z_1 Z_2 Z_3} \bmod p$.

By using this shared key, patient data are encrypted. In each session, all the participants share a unique private key generated in this way and a session expires when either a predefined time period passes or the content of an access control list is modified.¹⁶ After a session expires, a new key generation session is carried out.

Moreover, if the patient wants to set up further restrictions of accessing the PCR by any group member or lock the access, the patient's PSP asks for another large random number x_4 from CTSS as a secret agreement and shares $g^{x_1 x_2 x_3 x_4} \bmod p$ with CTSS.¹⁶ Therefore, no one except the patient can access the data without another key generation process because neither the responsible clinician nor referring clinician can know this secret key.

12.6.3 Hierarchical Security Architecture

At the PAN level, generated keys are used for user authentication of accessing PCRs, and without any shared key no one can perform decryption and consequently view or record vital signs. An example of a security algorithm to be used at this level is the secure hash algorithm (SHA).^{16,35} In this algorithm, a packet is a collection of clinical measurements, a shared key, and a time-stamp as well as the Medium Access Control (MAC) header. Basically time-stamps are put on the data when they are sent from patient's biomedical sensors. According to these parameters, the

packet authentication code for packet i (PAC_i) is calculated by using hash function H in the form of:

$$PA \quad C_i = H(K_H, T_{s,i}, H_i, D_i)$$

where K_H is a shared key, $T_{s,i}$ is a time-stamp, H_i is a MAC header, and D_i is a payload of the measured data. For example, in SHA-1, packets whose length is less than 264 bits are divided into a number of 512-bit blocks and encryption is performed for these blocks returning message digests having the length of 160 bits.^{34,36}

12.6.4 Digital Signature Schemes

Besides the Diffie–Hellman-like key generation scheme, Mišić, Amini, and Khan¹⁵ introduced the Elgamal signature algorithm³⁷ that achieves digital signature to enhance integrity of PCRs and to avoid data manipulations by malicious adversaries in medical wireless sensor networks. To establish a digital signature system among the central database, patient's security processor and clinician's personal servers in the same group, the CTSS and personal servers share another value x that is securely transmitted in a private key encryption by using a session key generated in the aforementioned way, so personal servers in the group as well as the CTSS can generate $y \equiv g^x \pmod{p}$ from x . In addition to computing y , a message sender also computes another value r such that:

$$r \equiv g^k \pmod{p}$$

where k is a random number in the range of 0 to $p - 1$ and also satisfies $\gcd(k, p - 1) = 1$.

According to y and r , the verification of the digital signature of message m , $0 \leq m \leq p - 1$, is calculated such that:

$$g^m \equiv y^r r^s \pmod{p}$$

By substituting y and r , this equation is also written as:

$$g^m \equiv g^{xr} g^{ks} \pmod{p}$$

At this point, if random number k satisfies $\gcd(k, p - 1) = 1$, s can be solved by using the following equation³⁷:

$$m \equiv (xr + ks) \pmod{(p - 1)}$$

Upon calculating s , message m is sent to one or more receivers along with the signature of pair (r, s) . Because every member in the same group and the CTSS have values g, y , and p as well as transmitted message m and pair (r, s) , authentication of the sender is easily computed by substituting the values into $gm \equiv y^r r^s \pmod{p}$.

12.6.5 Zigbee Security Application Programming Interfaces (APIs)

When the patient's PAN utilizes IEEE 802.15.4 standard, applying a security suite for Zigbee is another choice for the packet signing and authentication. In this case, personal servers or PSPs usually become coordinators. The Zigbee Alliance developed security APIs for both symmetric and asymmetric key encryptions for Zigbee devices.¹⁵ Like SHA, the Symmetric-Key Key Establishment (SKKE) protocol is one of keyed-hashing for message authentication techniques. In addition, as a modified AES, the Zigbee Alliance uses 128 bits of the block size.¹⁵ According to the SKKE protocol, the keyed-hash message authentication code (HMAC) with shared key *MacKey*, clinical measurement *MacData*, and one-way hash function H is calculated in the following manner:^{15,36,38}

1. If the length of *MacKey* is shorter than 128 bits, add 0s to *MacKey* until its length becomes 128 bits.
2. Calculate *MacKey* XOR *ipad*, where *ipad* is a constant with 64×36 s.
3. Concatenate *MacData* to a value calculated in 2.
4. Apply hash function H to a value calculated in 3.
5. Calculate *MacKey* XOR *opad*, where *opad* is a constant with 64×5 Cs.
6. Concatenate a value calculated in 4 to a value calculated in 5.
7. Apply hash function H to a value calculated in 6, which is a message digest.

In short, HMAC is calculated by using the following formula:

$$\text{HMAC}(\text{MacData}) = H(\text{MacKey XOR opad}, H(\text{MacKey XOR ipad}, \text{MacData}))$$

In addition, SKKE derives link keys that are used for security transmission between two devices. The link key generation begins between two Zigbee devices, one of which must be the PAN coordinator with exchanging each validation challenge.¹⁵ For example, after exchanging the challenges, two nodes U and V have challenges QEV and QEU , respectively, which are sent by the opponent device. According to these challenges, both devices concatenate each other's IDs and challenges to create a *MacData*:

$$\text{MacData} = (ID_U, ID_V, QEU, QEV)$$

where both devices must have the same value. Then apply the HMAC to this value to generate Z in the manner previously described. Based on Z , two cryptographic hashes are generated with two hexadecimal numbers such as $Hash_1 = H(Z, 0 \times 01)$, $Hash_2 = H(Z, 0 \times 02)$. Between the two values, $Hash_2$ is used for a link key and $Hash_1$ is used for the justification of both devices having the same key.

12.6.6 Security Protocol for Wireless Medical Networks with Multi-Hopping

In the previous models, we assume that personal servers or PSPs are within the coverage of the wireless room access points or gateways of a hospital network so that every data transmission can be achieved with only one hop. Moreover, a ny PSP covers data transmission from patient's biomedical sensors within one hop.¹⁶ However, some medical wireless sensor network architecture may not allow one-hop data transmission between either biomedical sensors and personal server or personal servers and a room access point due to some physical restrictions. In those cases, security protocols must tolerate the multi-hop data transmission. Nevertheless, securing the wireless sensor networks with multi-hopping is still a challenging issue of the field.⁴

At first, this security model focuses on securing data transmission between clinician's PDAs and the hospital network gateways or the mote interface boards (MIBs). To deal with a multi-hopping scheme, a number of nodes or PDAs are reorganized into several clusters, in each of which exactly one node is responsible for the respective cluster. This responsible node is called a cluster head; data communication between cluster heads is called an inter-cluster communication; and between a cluster head and its members is called an intra-cluster communication.⁴ Basically, clusters are organized in such a way that every node in a cluster must be able to transmit its data to the respective cluster head within one hop. As a result, multiple nodes at the clinician level can achieve a hierarchical security where cipher protocols may be different from each other.

12.6.7 Inter-Cluster Session Key Generation

In the inter-cluster session key generation, the MIBs initially generate M session keys, which are computed by a one-way hash function with the next session key in the form of⁴:

$$SK_i = H(SK_{i-1} + 1)$$

where SK_i is the i th session key. After deriving these M keys, the MIBs only multi-cast n th session key SK_n which is in the middle of the key set to every cluster head within the range, where $n \ll M$.

Because every cluster head also has the same hash function H , it can derive a set of n keys $SK_{n-1}, SK_{n-2}, \dots, SK_1, SK_0$ from the n th session key in such a way that:

$$\begin{aligned}
 SK_{n-1} &= H(SK_n) \\
 SK_{n-2} &= H(SK_{n-1}) \\
 &\dots \\
 SK_0 &= H(SK_1)
 \end{aligned}$$

Among a set of $n + 1$ keys $\{SK_n, SK_{n-1}, \dots, SK_1, SK_0\}$, only SK_0 is used for the current session key and the other keys are preserved in a local key buffer of the cluster heads. Because the first key, SK_0 , is initiated, session keys are periodically updated by the MIB, which multicasts a new session key from key repository $\{SK_{n+1}, SK_{n+2}, \dots, SK_{M-1}, SK_M\}$. Because these session keys to be updated also have a relation of $SK_i = H(SK_{i+1})$ with keys that are already saved in the local key buffer of the cluster heads, the key authentication or the MIB authentication is easily conducted by comparing the hashed values to the keys in the local key buffer. At this point, if the authentication succeeds, a session key newly distributed by the MIB is pushed into the local key buffer, while in the key buffer a key having the smallest identification number is shifted out of the buffer and used as a next session key. To the contrary, if the authentication fails, this key is just discarded from the session.

12.6.8 Initialization Vectors (IVs) and SkipJack Algorithm

In addition to facilitating ciphers with multi-hopping, Hu, Kumar, and Xiao⁴ suggest that because there tend to be few variations of patient’s biomedical data, part of the encryption should enhance the variation of the output so that two of the same plaintexts can be encrypted in two different ways. Thus, utilizing the initialization vectors (IVs) within encryption solves this problem. Basically, the IVs are used with the block ciphers. The IVs are random numbers with the same length of the plaintext block to be encrypted.³⁹ With shared key K and hash function H , the IVs are applied in the following manner:

$$\begin{aligned}
 CB_1 &= H(K, TB_1 \text{ XOR IV}), \\
 CB_2 &= H(K, TB_2 \text{ XOR } CB_1), \\
 &\dots \\
 CB_n &= H(K, TB_n \text{ XOR } CB_{n-1})
 \end{aligned}$$

where TB_i is the i th plaintext block, and CB_i is the i th cipher block. One advantage of the IV is that encrypted text blocks can only be readable with the previous cipher block.

Because 3-DES imposes heavy computations on mote-based PDAs and slows software running on them too much, a SkipJack-based symmetric key encryption is employed in this model.

12.7 Conclusions

This chapter presented a security aspect of the current wireless telemedicine technology. Because of the sensitivity of patient care records (PCRs) to everyone, malicious adversaries may target them and steal them so as to make use of them for their benefit. According to introductions of several possible medical wireless network attacks, nine security policies to protect patients from data leakage are introduced. These security policies mainly rely on the access controlling scheme that utilizes the access control list for each datum. Furthermore, the chapter also introduced major possible threats as results of data leakage.

However, in the medical wireless sensor network application, resource limitation is still a bottleneck for efficient security. Therefore, in this chapter, most of the security algorithms discussed are designed to be lightweight with low energy consumption, and consist mainly of key distribution schemes, secure hash algorithms, and digital signature schemes. One of them employs the Diffie–Hellman algorithm, which is designed to help restrict data access only among people who have the access permission, or a shared key. That is, the shared keys are considered as the access control lists for particular PCRs. The Elgamal signature algorithm is based on the Diffie–Hellman algorithm, and is used for the digital signature scheme, so it ensures that data from particular sources are guaranteed to come from the legitimate members of the same group. Also one security architecture employs a hierarchical structure so that it can be applied to the wireless sensor networks with multi-hopping. Furthermore, to expand signal variations of vital signs, the initialization vectors are equipped in the encrypting process.

In a future work, intrusion detection systems could be designed for medical wireless sensor networks.¹⁵ In addition, according to the progress of the wireless network resources, the security architectures should be remodeled to keep the confidentiality and integrity of the PCRs as much as possible.

Acknowledgment

This work is partially supported by the U.S. National Science Foundation (NSF) under grants CNS-0716211 and CNS-0716455.

References

1. A.T. Adler, A Cost-Effective Portable Telemedicine Kit for Use in Developing Countries (master's thesis, Massachusetts Institute of Technology, May 2000).
2. A. Milenković, C. Otto, and E. Jovanov, *Wireless Sensor Networks for Personal Health Monitoring: Issues and an Implementation*, Elsevier B.V., 2006.
3. M. Welsh, CodeBlue: A Wireless Sensor Network for Medical Care and Disaster Response, Matt Welsh—Harvard University, 2005.
4. F. Hu, S. Kumar, and Y. Xiao, Towards a secure, RFID/sensor based tele-cardiology system, *Proceedings of IEEE CCNC'07*.
5. D. Malan, T. Fulford-Jones, M. Welsh, and S. Moulton, CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care, Division of Engineering and Applied Sciences, Harvard University, and School of Medicine, Boston University.
6. V. Shnayder, B.-R. Chen, K. Lorincz, T.R.F. Fulford-Jones, and M. Welsh, Sensor Networks for Medical Care, Division of Engineering and Applied Sciences, Harvard University, 2005.
7. M. Welsh, D. Myung, M. Gaynor, and S. Moulton, Resuscitation Monitoring with a Wireless Sensor Network, Harvard University and 10Blade, Inc.
8. M. Welsh, D. Malan, B. Duncan, T. Fulford-Jones, and S. Moulton, Wireless Sensor Networks for Emergency Medical Care, Matt Welsh—Harvard University, 2004.
9. E. Jovanov, A. Milenković, C. Otto, and P.C. de Groot, A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation, *Journal of NeuroEngineering and Rehabilitation*, March 2005, <http://www.jneuroengrehab.com/content/2/1/6>
10. K.J. Liszka, D.W. York, M.A. Mackin, and M.J. Lichter, Remote monitoring of a heterogeneous sensor network for biomedical research in space, *Proceedings of ICWN 2004*.
11. K.J. Liszka, M.A. Mackin, M.J. Lichter, D.W. York, D. Pillai, and D.S. Rosenbaum, Keeping a beat on the heart, *IEEE CS and IEEE ComSoc*, 2004.
12. P. Bonato, P. Mork, D. Sherrill, and R. Weggaard, Data mining of motor patterns recorded with wearable technology, *IEEE Engineering in Medicine and Biology*. 22(3), 2003.
13. M. Mathie, A. Coster, N. Lovell, and B. Celler, Accelerometry: Providing an integrated, practical method for long-term, ambulatory monitoring of human movement, *Physiological Measurement*, 25(2), 2004.
14. J. Bussmann, J. Tulen, E. van Herel, and H. Stam, Quantification of physical activities by means of an ambulatory accelerometer: A validation study, *Psychophysiology*, 35(5), 1998.
15. J. Mišić, F. Amini, and M. Khan, On security attacks in healthcare WSNs implemented on 802.15.4 beacon enabled clusters, *Proceedings of IEEE CCNC 2007*.
16. J. Mišić and V.B. Mišić, Implementation of security policy for clinical information systems over wireless sensor networks, *Ad Hoc Networks*, 5: 134–144, 2007.
17. Y.-C. Wang and Y.-C. Tseng, Attacks and defenses of routing mechanisms in ad hoc and sensor networks, in *Security in Sensor Networks*, Y. Xiao, Ed., Auerbach Publications, Boca Raton, FL, 2007, pp. 3–25.

18. M. Pirretti, S. Zhu, N. Vijaykrishnan, P. McDaniel, M. Kandemir, and R. Brooks, The sleep deprivation attack in sensor networks: Analysis and methods of defense, *International Journal of Distributed Sensor Networks*, 2(3): 267–287, 2006.
19. R.J. Anderson, A Security Policy Model for Clinical Information Systems, *Proceedings of the 1996 IEEE Symposium on Security and Privacy*, 1996, 34–48.
20. N. Luck and J. Burns, “Your secrets for sale,” *The Daily Express*, Feb. 16, 1994, pp. 32–33.
21. L. Rogers and D. Leppard, For sale: Your secret medical records for £150, *Sunday Times*, November 26, 1995, 1–2.
22. Protecting Privacy in Computerized Medical Information, Office of Technology Assessment, U.S. Government Printing Office, 1993.
23. E. Bartlett, RMs need to safeguard computerized patient records to protect hospitals, *Hospital Risk Management*, 9: 129–140, 1993.
24. Is your health history anyone’s business? *McCall’s Magazine*, April 1995, p. 54, reported by M. Bruce on Usenet newsgroup comp.society.privacy, March 22, 1995.
25. K. Alderson, Nurse sacked for altering records after baby’s death, *The Times*, Nov. 29, 1995, p. 6.
26. Nurse jailed for hacking into computerized prescription system, *British Journal of Healthcare Computing and Information Management*, 1(94), 7, 1994.
27. S. Brown, Sanitas launches health credit card, *Cards International*, October 6, 1995, p. 3.
28. Identity Cards: A Consultation Document CM2879 — Response of the Data Protection Registrar, October 1995.
29. Good Medical Practice, General Medical Council, 178–202, Great Portland Street, London.
30. Confidentiality, General Medical Council, 178–202, Great Portland Street, London.
31. A. Sommerville, Medical ethics today: Its practice and philosophy, *BMA’s Handbook of Ethics and Law*, 2nd ed., 1993.
32. A. Griew and R. Currell, A Strategy for Security of the Electronic Patient Record, IHI, University of Wales, Aberystwyth, March 14, 1995, www.ihl.org.
33. O. Elkeelany, M. M. Matalgh, K. P. Sheikh, M. Thaker, G. Chaudhry, D. Medhi, et al., Performance analysis of IPSec protocol: Encryption and authentication, *Proceedings of IEEE International Conference on Communication ICC 2002*, vol. 2, pp. 1164–1168, 2002.
34. W. Diffie and M. E. Hellman, New direction in cryptography, *IEEE Transactions on Information Theory*, 22(6): 644–654, 1976.
35. National Institute of Standards, Digital Signature Standard, Gaithersburg, MD, U.S. Department of Commerce, 1994.
36. H. Krawczyk, M. Bellare, and R. Canetti, HMAC: Keyed-Hashing for Message Authentication, The Internet Society, 1997, <http://www.ipa.go.jp/security/rfc/RFC2104EN.html>
37. T. Elgamal, A public key cryptosystem and a signature scheme based on discrete logarithms, *IEEE Transactions on Information Theory*, 4, IT-31, July 1985, 11–31.
38. M. Bellare, R. Canetti, and H. Krawczyk, Keyed hash function and message authentication, *Proceedings of Crypto’96, LNCS 1109*, pp. 1–15.

39. W. Pitt, Block Ciphers and Initialization Vectors (IVs), The Security Samurai, 2005, <http://marvets.com/blog/archive/2005/06/10/193.aspx>
40. D. Eastlake and P. Jones, U.S. Secure Hash Algorithm 1 (SHA1), The Internet Society, 2001, <http://www.ipa.go.jp/security/rfc/RFC3174EN.html>
41. J.S. Reddy, Zigbee Security, Zigbee Alliance, 2004.
42. P. Baronti, P. Pillai, V. Chook, S. Chessa, A. Gotta, and Y. Fun Hu, Wireless Sensor Networks: A Survey on the State of the Art and the 802.15.4 and Zigbee Standards, <http://dienst.isti.cnr.it/Dienst/Repository/2.0/Body/ercim.cnr.isti/2006-TR-18/pdf>

Chapter 13

Power Management and Security in IEEE 802.15.4 Clusters: How to Balance?

Fereshteh Amini, Moazzam Khan, and Jelena Mišić

CONTENTS

13.1 I ntroduction.....	238
13.2 I EEE 802.15.4	239
13.3 S ecurity in IEEE 802.15.4	240
13.3.1 S ecurity Building Blocks.....	240
13.3.1.1 K eying Models.....	241
13.3.1.2 K eeyed Hash Function for Message Authentication	242
13.3.2 S ymmetric-Key Key Establishment Protocol (SKKE)	242
13.3.2.1 E xchange of Ephemeral Data	242
13.3.2.2 G eneration of Shared Secret	243
13.3.2.3 D erivation of Link Key	244
13.3.2.4 C onfirming Link Key.....	245
13.3.3 U se of SKKE in Our Simulation Model	246
13.3.4 L ink Key Updates	248

13.4 P ower Management	249
13.5 S imulation Model	250
13.5.1 Beacon-Enabled IEEE 802.15.4 Simulation Model	250
13.5.2 Adding Key Exchange Mechanism to the Simulation Model of IEEE 802.15.4 Network	252
13.5.3 Simulation Run and Analysis	253
13.6 Conclusion and Future Work.....	261
References	262

The IEEE 802.15.4 specification is a recent low data rate wireless personal area network standard. While basic security services are provided for, there is a lack of more advanced techniques, which are indispensable in modern personal area network applications. In addition, performance implications of those services are not known. In this chapter, we describe a secure data exchange protocol based on the ZigBee specification and built on top of IEEE 802.15.4 link layer. This protocol includes a key exchange mechanism. We assume that all nodes are applying power management technique based on the constant event sensing reliability required by the coordinator. Power management generates random sleep times by every node, which in average fairly distributes the sensing load among the nodes. Key exchange is initiated by cluster coordinator after some given number of sensing packets has been received by the coordinator. We develop and integrate simulation model of key exchange and power management technique into cluster's reliable sensing function. We evaluate the impact of security function and its periodicity on cluster performance.

13.1 Introduction

The IEEE 802.15.4 specification outlines a class of wireless radios and protocols targeted at low power devices, personal area networks, and sensor devices. IEEE 802.15.4 specification employs a number of well-known security services that can be implemented but at the cost of memory and communication overhead. Currently, not many wireless sensor network overhead statistics are available when security is employed in such networks. Sensor network application developers and network administrators always need these overhead statistics in choosing the security option that best suits the security for a particular threat environment. For evaluating these security overheads on wireless sensor networks, we will simulate IEEE 802.15.4 media access control layer and secure data exchange once the devices exchange link keys with the PAN coordinator. We will measure communication costs that are incurred after employing these security features under different inputs to a wireless sensor network model.

Key update provides an automated mechanism for restricting the amount of data which may be exposed when a link key is compromised. Key update frequency depends on the key update overheads and threat environment under which network is working. Hence controlling the lifetime of keys and determination of how the key update occurs is a technical challenge. Some authors have reported the activity management and network behavior without considering any security parameters.¹ In this work we develop a simulation model for the cluster behavior including periodic key exchange (with variable update threshold), power management, and sensing data application. For activity management, nodes in cluster apply sleep technique in order to deliver only the required number of packets per second (which we will call event sensing reliability) to the coordinator. We obtained simulation results to evaluate the overhead of key exchange in terms of medium behavior, total number of delivered packets, nodes' utilization, and its effect on node's lifetime.

The chapter is organized as follows. We give an overview of IEEE 802.15.4 specification in Section 13.2 and later in Section 13.3 we introduce the security features addressed in IEEE 802.15.4. As IEEE 802.15.4 does not address any keying model, we are relying on keying model from ZigBee specification and discuss this in Section 13.3.2, and Section 13.3.4 explains the key update technique used in our study. In Section 13.4 we explain the approach used for key activity management of the network to provide predetermined reliability. In Section 13.5 we explain the simulation model, how it is implemented and in the same section will present our results. Finally we conclude our work in Section 13.6.

13.2 IEEE 802.15.4

The need for low-cost, low-power, short-range communication is the main reason for introducing IEEE 802.15.4 low rate wireless personal area network (LR-WPAN) standard.² According to this specification, such WPAN consists of devices which are the basic components of these networks. Two or more devices communicating in a common physical channel create a WPAN.

Star topology is one option for communication in LR-WPAN. In this topology devices communicate via a single central controller called PAN coordinator. After deciding on a PAN identifier, PAN coordinator may decide whether a device can join the PAN.

In the current work we concentrate on beacon-enabled based communication. In this form of communication, devices first listen for the network beacon. When the beacon is found, the device synchronizes to the superframe structure. At the appropriate point, the device transmits its data packet, using slotted CSMA-CA, to the coordinator (uplink). The coordinator acknowledges the successful reception of the data by transmitting an acknowledgment frame.

On the other hand, when the PAN coordinator has something to send to a device (downlink), it informs the device by including in the network beacon that

a data message is pending. The device periodically listens to network beacon and, if a message is pending, transmits a request frame to the coordinator using slotted CSMA-CA. The coordinator acknowledges the successful reception of the data request by transmitting an acknowledgment frame. The pending data frame is then sent using slotted CSMA-CA. The device acknowledges the successful reception of the data by transmitting an acknowledgment frame.

13.3 Security in IEEE 802.15.4

IEEE 802.15.4 standard provides physical and link layer solutions for wireless personal area networks. It also provides well-known and well-understood cryptographic techniques^{3,4} by supporting authentication, message integrity, confidentiality, and freshness check for preventing replay attacks. Application of such security mechanisms comes at a cost that includes processing overhead, memory overhead, power consumption, and resulting low bandwidth.⁵

An application implemented using IEEE 802.15.4 has a choice of different security suites that control the type of security protection by setting appropriate control parameters in the link layer security suite stack. A long Message Authentication Code (MAC) size improves the security feature of authentication and it is very difficult for an adversary to break or guess a MAC of longer size.⁶ But this improved security is achieved at the cost of longer packet size. In IEEE 802.15.4 compliant wireless sensor networks, packet size is very crucial to the overall throughput that is required by the application. Applications that support continuous data flow would be affected more than the applications in which data flow is periodic. Applications used for real-time monitoring of some critical environments rely on continuous flow of data and hence by implementing security will affect the overall throughput and lifetime of such network by increasing the packet size. For the current work we will employ the security suite specified in IEEE 802.15.4 that supports both encryption and data integrity with MAC size of 128 bits. The security suite uses Counter with CBC-MAC (CCM)⁷ mode of AES (Advanced Encryption Standard) for encryption and authentication. This cryptographic technique uses Counter by first applying integrity protection both on message header and data payload and later it encrypts the data payload and MAC using AES. At the receiver end the receiver gets the packet and applies decryption using parameters based on sender's address from its access control list.

13.3.1 Security Building Blocks

The IEEE 802.15.4 specification provides basic security mechanisms but these security features cannot work on their own. The level of security in any network

revolves a round of the keys that are shared among devices. Different approaches have been suggested to distribute and manage these keys. Because IEEE 802.15.4 does not suggest any keying mechanism, in this work we will follow the keying mechanism from ZigBee Alliance specifications.⁴ In this section we will first introduce the keying mechanisms and later explain how this is handled in ZigBee specification by taking advantage of the inherent security mechanisms already provided by IEEE 802.15.4.

13.3.1.1 Keying Models

As explained above, the IEEE 802.15.4 addresses good security mechanisms but it still does not address what type of keying mechanism will be used to employ the above techniques.

ZigBee Alliance⁴ is an association of companies working together to enable wireless networked monitoring and control products based on IEEE 802.15.4 standard. After the acceptance of 802.15.4 as IEEE standard, ZigBee Alliance is mainly focused on developing network and application layer issues. ZigBee Alliance is also working on application programming interfaces (API) at network and link layer of IEEE 802.15.4. ZigBee Alliance also introduces secure data transmission in wireless sensor network that are based on IEEE 802.15.4 specification but most of this work is in general theoretical descriptions of security protocol at the network layer. No specific study or results have been published by ZigBee Alliance in regard to which security suite performs better in different application overheads. ZigBee Alliance has also recommended both symmetric and asymmetric key exchange protocols for different networking layers. Asymmetric key exchange protocols that mainly rely on public key cryptography are computationally intensive and their feasibility in wireless sensor networks is only possible with devices that are resource-rich both in computation and power.

Application support sublayer of ZigBee specification provides the mechanism by which a ZigBee device may derive a shared secret key (link key) with another ZigBee device. Key establishment involves two entities, an initiator device and a responder device, and is prefaced by a trust provisioning step. Trust information (e.g., master key) provides a starting point for establishing a link key and can be provisioned in-band or out-band.

ZigBee Alliance uses Symmetric-Key Key Establishment (SKKE) protocol for link key establishment. In SKKE an initiator device establishes a link key with a responder device using a master key. This master key, for example, may be pre-installed during manufacturing, may be installed by a trust center, or may be based on user-entered data (PIN, password). In the current study we assume that all the devices and PAN coordinator have pre-installed master keys and we will focus mainly on link key establishment.

13.3.1.2 Keyed Hash Function for Message Authentication

A hash function is a way of creating a small digital fingerprint of any data. Cryptographic hash function is a one-way operation and there is no practical way to calculate a particular data input that will result in a desired hash value; thus, it is difficult to forge. A practical motivation for constructing hash functions from block ciphers is that if an efficient implementation of block cipher is already available within a system (either in hardware or in software), then using it as the central component for a hash function may provide later functionality at little additional cost. IEEE 802.15.4 protocol supports a well-known block cipher AES and hence ZigBee Alliance specification also relied on AES. ZigBee Alliance suggested the use of Matyas–Meyer–Oseas⁸ as the cryptographic hash function that will be based on AES with a block size of 128 bits.

Mechanisms that provide integrity checks based on a secret key are usually called message authentication codes (MACs). Typically, message authentication codes are used between two parties that share a secret key in order to authenticate information transmitted between these parties. ZigBee Alliance specification suggests the keyed hash message authentication code (HMAC) as specified in the FIPS Pub 198.⁹ A MAC takes a message and a secret key and generates a *MACtag*, such that it is difficult for an attacker to generate a valid (message, tag) pair and to forge messages. The calculation of *MacTag* (i.e., HMAC) of data *MacData* under key *MacKey* will be shown as follows:

$$MacTag = MAC_{MacKey} MacData$$

13.3.2 Symmetric-Key Key Establishment Protocol (SKKE)

Key establishment involves two entities, an initiator device and a responder device, and is prefaced by a trust provisioning step. Trust information (e.g., a master key) provides a starting point for establishing a link key and can be provisioned in-band or out-band. In the following explanation of the protocol we assume unique identifiers for initiator device as *U* and for responder device (PAN coordinator) as *V*. The master key shared among both devices is represented as *Mkey*.

We will divide Symmetric-Key Key Establishment Protocol (SKKE) between initiator and responder in the following major steps.

13.3.2.1 Exchange of Ephemeral Data

Figure 13.1 illustrates the exchange of the ephemeral data where the initiator device *U* will generate the challenge *QEU*. *QEU* is a statistically unique and unpredictable

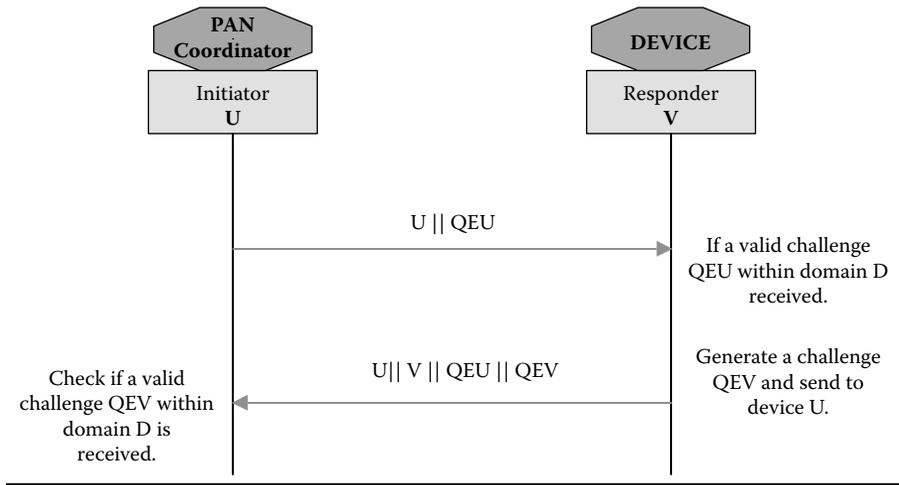


Figure 13.1 Exchange of ephemeral data.

bit string of length *challengeLen* by using either a random or pseudorandom string for a challenge *Domain D*. The challenge domain *D* defines the minimum and maximum length of the challenge.

$$D = (\text{minchallengeLen}, \text{maxchallengeLen})$$

Initiator device *U* will send the challenge *QEU* to responder device, which upon receipt will validate the challenge *QEU* by computing the bit length of bit string challenge *QEU* as *ChallengeLen* and verify that

$$\text{ChallengeLen} \in [\text{minchallengeLen}, \text{maxchallengeLen}]$$

Once the validation is successful the responder device will also generate a challenge *QEV* and send it to initiator device *U*. The initiator will also validate the challenge *QEV* as described above.

13.3.2.2 Generation of Shared Secret

Both parties involved in the protocol will generate a shared secret based on unique identifiers (i.e., distinguished names for each of the parties involved), symmetric master keys, and challenges received and owned by each party (Figure 13.2).

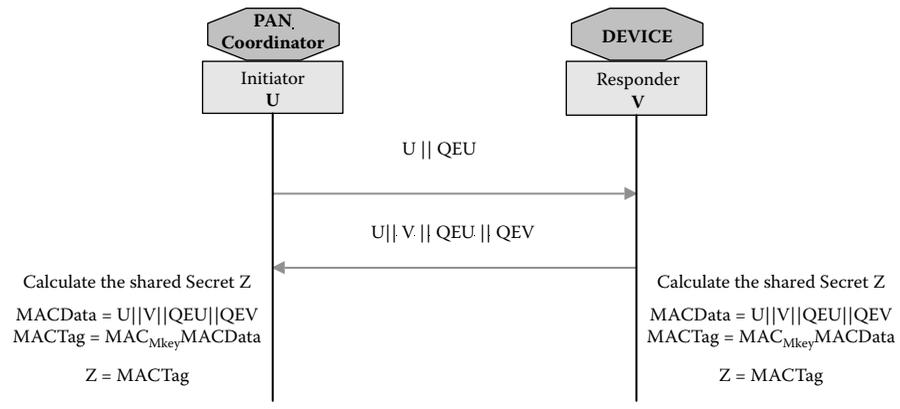


Figure 13.2 Generation of shared secret.

1. Each party will generate a *MACData* by appending their identifiers and respective valid *Challenges* together as follows:

$$MACData = U || V || QEU || QEV$$

2. Each party will calculate the *MACTag* (i.e., keyed hash) for *MACData* using *Mkey* (master key for the device) as the key for keyed hash function as follows:

$$MACTag = MAC_{Mkey} MACData$$

3. Now both parties involved have derived the same secret *Z*. (*Note:* This is just a shared secret, not the link key. This shared secret will be involved in deriving the link key but is not the link key itself.)

$$Z = MACTag$$

13.3.2.3 Derivation of Link Key

Each party involved will generate two cryptographic hashes (this is not the keyed hash) of the shared secret as described in ANSI X9.63-2001¹⁰:

$$Hash_1 = H(Z || 01)$$

$$Hash_2 = H(Z || 02)$$

The hash value $Hash_2$ will be the link key among two devices (Figure 13.3). Now for confirming that both parties have reached the same link key ($KeyData = Hash_2$), we will use value $Hash_1$ as the key for generating keyed hash values for confirming stage of the protocol.

$$MACKey = Hash_1 \quad (1) \tag{3.1}$$

$$KeyData = Hash_2 \quad (1) \tag{3.2}$$

$$K KeyData = Hash_1 \parallel Hash_2 \quad (1) \tag{3.3}$$

13.3.2.4 Confirming Link Key

Until this stage of protocol both parties are generating the same values and now they want to make sure that they reached the same link key values, but they do not want to exchange the actual key at all. For this they will once again rely on keyed hash functions and now both devices will generate different *MACTags* based on different data values, but will use the same key (i.e., *MACKey*) for generating the keyed hashes (*MACTags*).

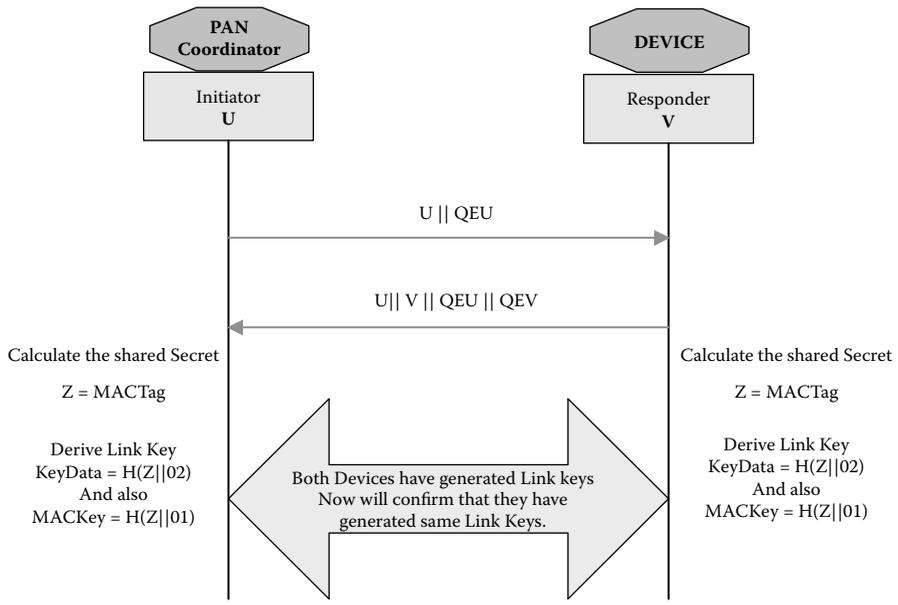


Figure 13.3 Generation of link key.

1. Generation of MACTags

Initiator and responder devices will first generate *MACData* values and based on these values will generate *MACTags*. Initiator device *D* will receive the *MACTag₁* from the responder device *V* and generate *MACTag₂* and send to device *V*.

We explain the generation of both *MACData* values and *MACTags* as follows. First both devices will calculate *MACData* values:

$$MACData_1 = 02_{16} \| V \| U \| QEU \| QEV$$

$$MACData_2 = 03_{16} \| V \| U \| QEU \| QEV$$

From the above *MACData* values both devices will generate the *MACTags* using the key *MACkey* (Equation 13.1) as follows:

$$MacTag_1 = MAC_{MacKey} MacData_1$$

$$MacTag_2 = MAC_{MacKey} MacData_2$$

2. Confirmation of MACTags

Now the initiator device *D* will receive *MacTag₁* from responder and responder device *V* will receive *MACTag₂* from device *D* and both will verify that the received *MACTags* are equal to corresponding calculated *MACTags* by each device. Now if this verification is successful each device knows that the other device has computed the correct link key (Figure 13.4).

13.3.3 Use of SKKE in Our Simulation Model

We have implemented SKKE in four major communication steps as are described in ZigBee specification⁴ (Figure 13.5):

1. SKKE-1: Initiator *U* will send the challenge *QEU* and wait for the challenge *QEV* from responder *V*.
2. SKKE-2: Responder *V* will receive the challenge *QEU* from initiator *U*, calculates its *QEV* and in the same data packet will send the *MacTag₁*.
3. SKKE-3: Initiator will verify the *MacTag₁* and if it is verified successfully, will send its *MacTag₂*. Now the initiator has a link key but will wait for an acknowledgment that its *MacTag₂* has been validated by the responder *V*.

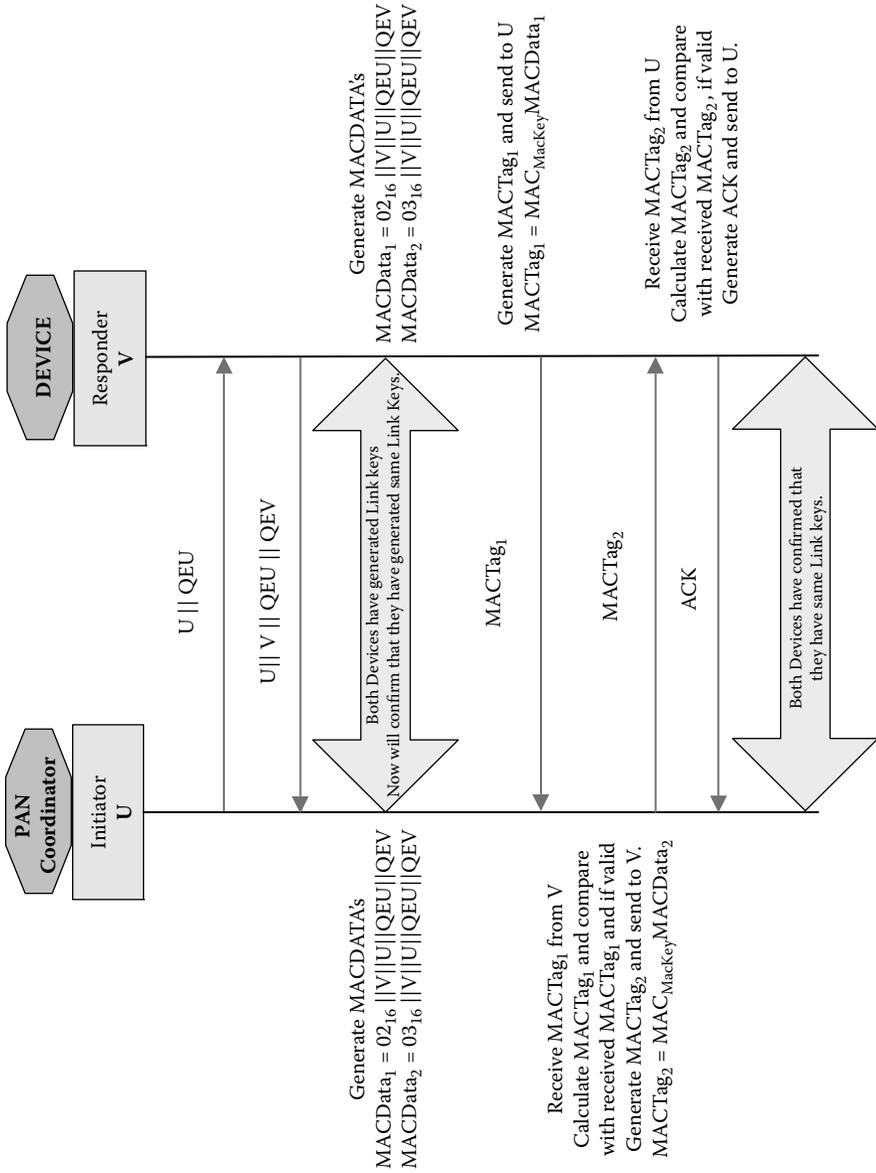


Figure 13.4 Confirmation of link keys.

4. SKKE-4: Responder will receive and validate the $MacTag_2$ from the initiator. If $MacTag_2$ is validated successfully, the responder will send an acknowledgment and now both initiator and responder have link keys. Once initiator receives this SKKE-4 message, key establishment is complete and now regular secure communication can proceed using link key among the initiator and the responder.

13.3.4 Link Key Updates

Key management consists of techniques and procedures supporting the establishment and maintenance of keying relations between authorized parties. Key management is simplest when all keys are fixed for all time. The time period over which these keys are valid for use is limited because use of the same key may result in giving enough information relating to a specific key for cryptanalysis and also may expose network traffic in case of compromise of a single key.

Depending on the severity of the threat environment, it is possible that a node or link key is somehow compromised by an adversary and can send false data to the

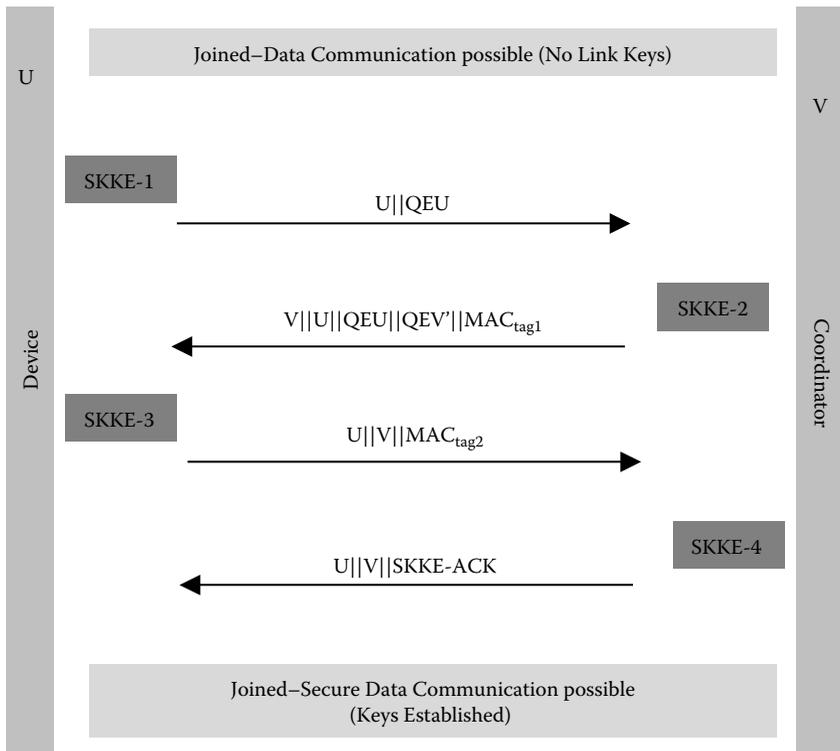


Figure 13.5 SKKE Protocol.

PAN coordinator. Key update provides an automated mechanism for restricting the amount of data that may be exposed when a link key is compromised. Sound security policies regarding transparent key updates is a fundamental component of sound security practices. But key updates protocol depends on the key update overheads and threaten environment under which network is working. Hence controlling the lifetime of keys and determination of how the key update occurs is a challenging task in any network. Approaches for key updates in general wireless networks mainly target networks that have group key structures and have high communication bandwidth.^{11,12} For resource-scarce IEEE 802.15.4 networks these key updates will affect the performance adversely.

In this work we assume that a PAN coordinator maintains a counter for each node that keeps track of the number of packets exchanged under the same key (figure 13.5). When the threshold value of the counter is reached for any device, the PAN coordinator will initiate the key exchange with all the devices in the cluster. During the key exchange, all devices will temporarily stop the data transmission and resume it when they acknowledge the new key. The alternative approach is to use the single counter for all the devices. However, this approach may open a security hole for a denial-of-service attack by a single corrupted device.

13.4 Power Management

Power management consists of adjusting the frequency and ratio of active and inactive periods of sensor nodes.^{13,14} For IEEE 802.15.4 nodes it can be implemented in two ways. In the first one, supported by the standard,⁴ the interval between the two beacons is divided into active and inactive parts, and the sensors can switch to low-power mode during the inactive period. Activity management for individual nodes can be accomplished through scheduling of their active and inactive periods.

Let us consider a sensing application in which redundant sensors are used to achieve the desired value of event sensing reliability (number of packets per second needed for reliable event detection).¹⁴ We assume that individual nodes sleep for a random time interval, the duration of which is a geometrically distributed random variable regulated with probability P_{sleep} . When a node wakes up, it waits for the beacon from the coordinator before it attempts to transmit the packet. We have used Bernoulli scheduling for the packet scheduling during the active period of the node. In this approach, at the end of each packet transmission the node checks its uplink buffer. If it is empty, the node immediately goes to sleep; if there are packets to send, the node transmits the next packet from the buffer with the probability P_{active} , or goes to sleep with the probability $1 - P_{active}$. Therefore, two control parameters are needed: one, P_{sleep} , regulates the duration of the inactive period; the other, P_{active} , regulates the duration of the active period. When individual nodes begin to cease functioning, either because of battery exhaustion or for other reasons, the

remaining nodes will have to extend their activity to achieve satisfactory reliability, and the importance of the Bernoulli mechanism will increase.

Depending on whether we split the computational load of activity management, we can have two approaches of centralized and distributed controls. By choosing the latter approach to distribute the computational load more evenly, we assume that the network coordinator is aware of the number of sensor nodes (which have to be explicitly admitted to the network²) and their packet arrival rates (which may be obtained as simple long-term averages, as packet headers contain the source node address). The coordinator first determines node utilization based on the number of live nodes and then calculates the individual reliability r per node (by dividing the required collective reliability R by the number of live nodes n) and sends this information within the beacon frame. Over time some sensors die, and the coordinator has to broadcast updated values of individual reliability, which grow whenever one of the sensors die. Note that the sleep time is geometrically distributed, and the mean sleep time is $t_{bof f} / (1 - P_{sleep}) = 1/r$, where $t_{bof f} = 0.32ms$ corresponds to the duration of one backoff period. Therefore, each sensor node starts with $P_{sleep} = 1 - rt_{bof f}$ and $P_{active} = 0$. It then monitors the utilization of its radio transmitter/receiver subsystem, using a monitoring window of specified size. Utilization is simply calculated as the count of backoff periods in which the node was active during the recent window divided by the total size of the window.

We have developed a Markov chain model for node behavior, which includes all phases of SKKE protocol and subsequent sleep and transmission phases. We assume that the PAN coordinator maintains a separate counter for the number of transmissions by each node. When the counter value reaches threshold n_k , key update protocol is triggered. Updated keys are used to generate message authentication code. The high-level Markov chain, which includes key update sleep periods followed by the transmissions, is presented in Figure 13.6.

13.5 Simulation Model

We have simulated the key exchange mechanism and sleep mechanism for the IEEE 802.15.4 network using Artifex,¹⁵ a general development platform for discrete event simulations. For the remainder of this section we first give a quick introduction of beacon-enabled simulation model of 802.15.4^{1,16} and later explain the simulated key exchange and update process and power management in our current work.

13.5.1 Beacon-Enabled IEEE 802.15.4 Simulation Model

The network communication model of this simulation is based on star topology. The model is built on three primary objects: PAN coordinator, device, and medium.

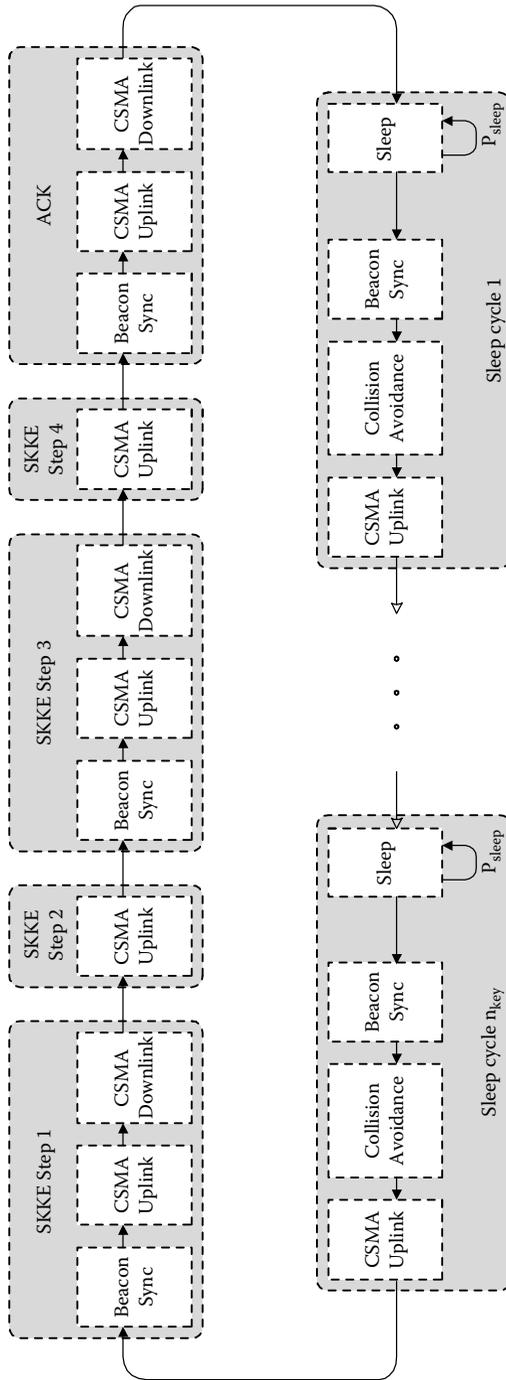


Figure 13.6 Markov chain for the node behavior under threshold triggered key exchange.

The device and PAN coordinator objects are interconnected via medium object in our simulation model.

Two different token types are defined that play the role of packet and backoff. Packets can be any beacon, MAC request, data, and acknowledgment (ack) types. The communication is initiated when the PAN coordinator first sends beacon to medium (beacons are sent after every $48t$ where t is duration of one backoff period). After receiving the beacon the medium starts a clock and sends pulse to all devices every t time.

Data packets are generated by device object following exponential distribution and are destined to a randomly chosen device. The packet is then sent to the medium and a copy of it is kept for retransmission if needed. Data packets are then received by the medium. If the number of received packets in the medium is greater than 1, collision occurs. If there are no collisions, data packets are sent successfully to the PAN coordinator and the medium status is set to busy.

PAN coordinator is the next stop for data packets and is responsible for sending ack type packets to the corresponding device after a specified delay. As with every packet, ack will be received first by the medium and then sent to the corresponding device. When PAN coordinator is sending data to a device it keeps a finite buffer for each device in the PAN. If the buffer of the device which the data packet is destined for is full, the packet will be discarded. In the case that there is still room in that device's buffer, the coordinator adds the destination ID of a packet to the pending devices list and advertises the ID in the beacon. The device will notice that there is a packet waiting for it and will initiate a MAC request packet to be sent to the coordinator. The PAN coordinator after receiving the request will perform round robin scheduling algorithm and choose the device to send the packet from its corresponding downlink buffer.

13.5.2 Adding Key Exchange Mechanism to the Simulation Model of IEEE 802.15.4 Network

In this section we describe the communication between the ordinary nodes and PAN coordinator, which occurs as a result from the link key exchange. We assume that devices are attached to the cluster and the formation of the piconet is finalized. Also, we assume the master keys are established, so that there is no threat of eavesdropping during exchange of master keys. The next step is generating link keys between each device and PAN coordinator. For the exchange of link keys, we will follow SKKE protocol as described in Section 13.3.3.

The process of key generation starts by PAN coordinator's advertisement for the first phase of key generation packets. Depending on which stage of generation we are in, the corresponding SKKE type of data packet (ranging from 1 to 4) will be processed (e.g., the first data packet has the type of SKKE-1 and so on). According to the standard specification, at most seven devices can be advertised in each

beacon. Therefore the PAN coordinator will advertise seven devices in each beacon. According to the standard, each device listens to each beacon and if its ID is being advertised the device will send a request packet. Request packet is transmitted in CSMA-CA mode and can collide with other packets. If it is received successfully by the PAN coordinator it will be acknowledged and downlink packet transmission carrying the SKKE protocol data will follow in the downlink transmission.

In our model, key exchange packets have nonpreemptive priority over data packets. If the node has started backoff process for data packet and it hears its ID in the beacon it will finish the current packet transmission before sending the request packet. However, if data packet arrives to the device's buffer while the key exchange is going on, its transmission will be postponed until device receives the new link key. PAN coordinator will first check key for the destination device from its access control list and no packet will be sent to the specific destination until the corresponding link key is already exchanged between PAN coordinator and the node. From this point on regular secure data packets will be immediately sent to the destination.

13.5.3 Simulation Run and Analysis

We have implemented the physical, data link, and security layers of an IEEE 802.15.4 cluster operating in beacon-enabled, slotted CSMA-CA mode. The packet size without security overheads includes all physical layer and medium access control layer headers, and it is set to 30 bytes, i.e., to three backoff periods. When packet signature (message authentication code) of 16 bytes is added to the total, packet size had to be rounded to five backoff periods (the largest packet size could be set to 13 backoff periods).

The cluster under consideration contains 14 devices, each having buffer capacity for three packets. Packet arrival per device followed the Poisson process with average rate of 90.5 packets per minute. When the coordinator announces key exchange in the beacon, all nodes have to stop uplink data transmissions temporarily until they receive new key initialization values from the coordinator in the downlink packets. Due to complex downlink data-link transmission algorithm we expect that key exchanges will adversely affect the regular sensing traffic.

We considered the impact of the increase of packet size due to addition of message authentication code, increased processing time needed for encryption in AES with CBC-MAC, and key exchange between the nodes over various packet arrival rates and cluster sizes. Figure 13.7 presents throughput, access probability (probability of no packet collision), and blocking probability at the node's buffer when all security overhead is included. Results were taken for varying number of nodes and varying packet arrival rate per node. Figure 13.8 presents the same parameters (except the key exchange cost because it does not exist) when no security measures are deployed in the network. We observe that without security measures, blocking probability is equal to zero, i.e., the network works without losses.

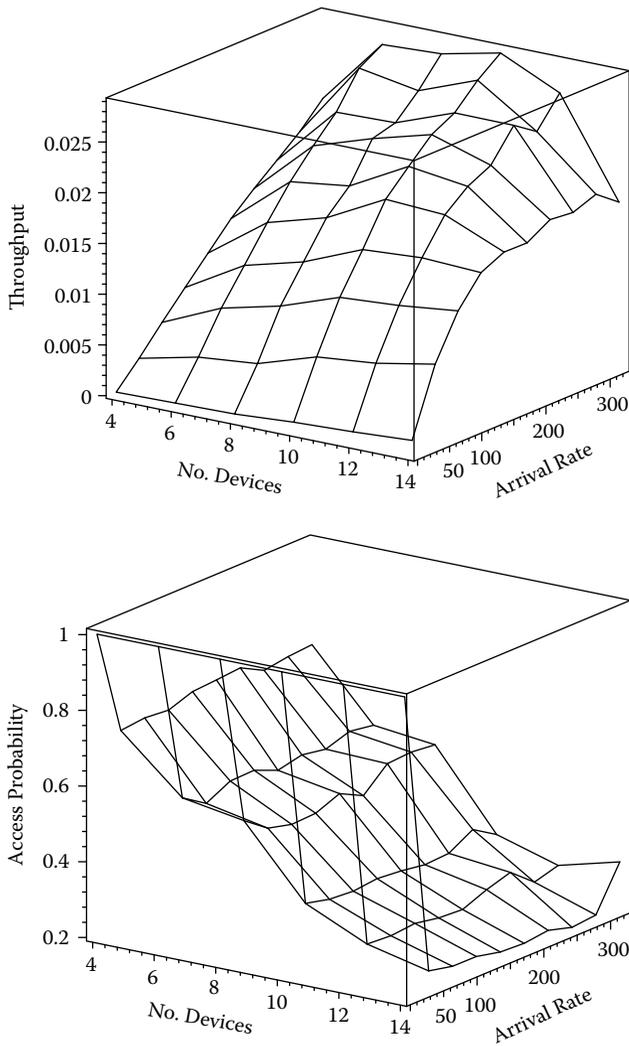


Figure 13.7 Throughput, access probability, and blocking probability as the function of simulation time (backoffs) for the case when security is employed and all devices stop their communications to update their keys.

The experiment to measure the cost of key update in the cluster contains seven devices only, and therefore it was possible to advertise the keys for devices in a single beacon. All devices temporarily stopped their data transmission during the key exchange. The behavior of the cluster over time is presented in Figure 13.9. Figure 13.9(a) shows number of backoff periods spent in key exchange. We notice that average cost of key exchange is slightly below 2000 backoff periods, which

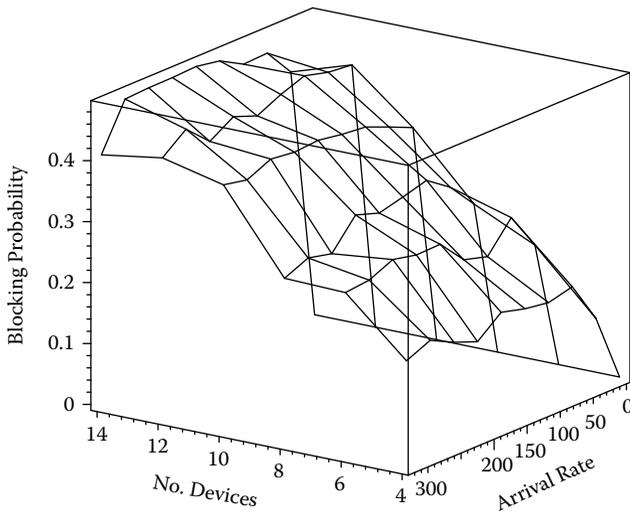


Figure 13.7 (Continued)

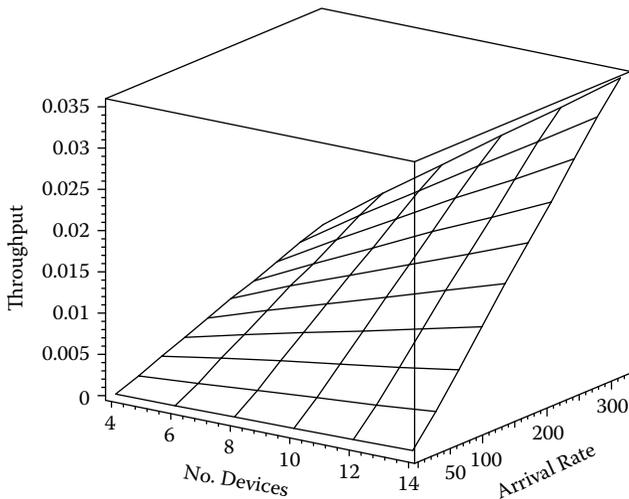


Figure 13.8 Throughput, access probability, and blocking probability as the function of simulation time (backoffs) when no security technique is employed.

gives 250 to 270 backoff periods per device. Knowing that the key exchange involves a total of two downlink (uplink request + downlink data) transmissions and three uplink transmissions, we conclude that one CSMA-CA access takes approximately 40 backoff periods. Given the backoff window sizes of (8, 16, 32) we conclude that transmission commences in average after third backoff attempt,

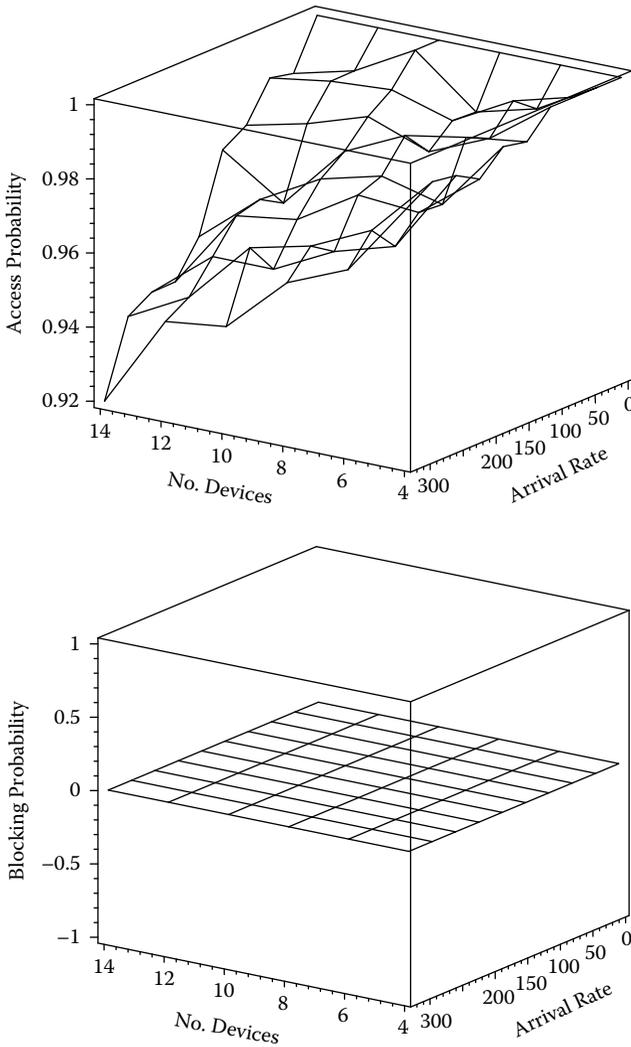


Figure 13.8 (Continued)

which indicates moderate to high activity over the medium. The blocking probability at individual sensor node buffer over the snapshot periods is shown in Figure 13.9(b). Due to large periods when device transmission is prevented during key exchange (well over 1500 backoff periods), the blocking probability skyrockets to values between 0.7 and 1. When the key exchange is finished, normal data communications resume. As a result, the blocking probability drops

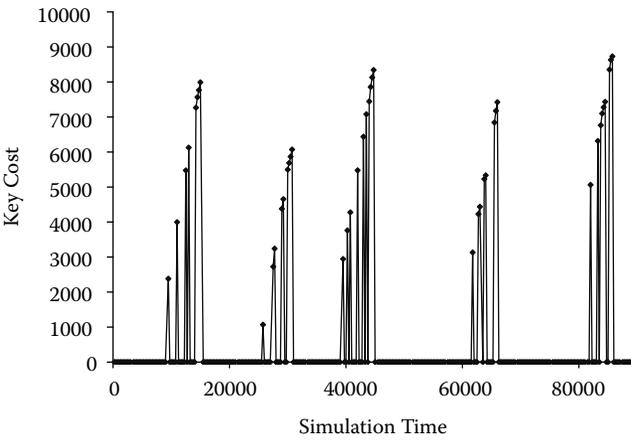
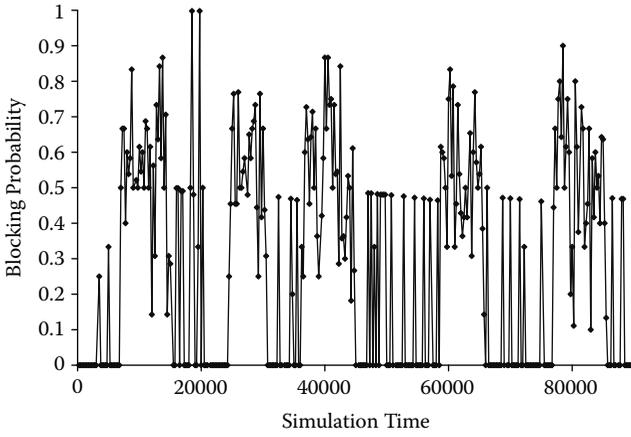


Figure 13.9 (a) Key cost, (b) blocking probability, and (c) throughput as the function of simulation time (backoffs) when devices stop their communications for key updates

abruptly to values around 0.3 and slowly declines further as the backlogged packets clear.

Figure 13.9(c) shows the throughput values measured during snapshot intervals of 250 backoff periods. The throughput of data packets is shown in white, while the throughput of key-exchange packets is shown in black. According to the throughput results reported in Mišić, Shafi, and Mišić,¹ the observed network regime without

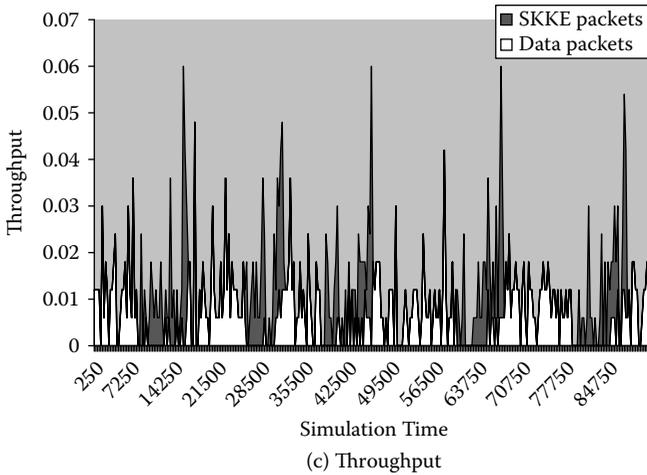


Figure 13.9 (Continued)

key exchange is slightly below the saturation condition (in saturation condition, all data transmissions end up in collisions).

We have also implemented distributed activity management in our simulator, assuming that the battery for each node has a fixed capacity. Battery capacity, which is expressed in backoff periods, is decremented by one for each backoff period in which the radio subsystem is active.

We have varied the key exchange threshold (n_k) between 40 and 100 packets while the requested event sensing reliability was kept at $R=10$ packets per second. Cluster size (n) was varied between 5 and 30 nodes. We have assumed that the network operates in the ISM band at 2.45 GHz, with raw data rate of 250 kbps. The packet size was fixed at 12 backoff periods, and the device buffers have a fixed size of three packets. The packet size includes message authentication code and all physical layer and medium access control protocol sublayer headers, and is expressed as the multiple of the backoff period.⁴ We also assume that the physical layer header has 6 bytes, and that the medium access control sublayer header and frame check sequence fields have a total of 9 bytes.

Figure 13.10(b) shows the total number of successfully transmitted packets (including key and data information) transmitted per second for requested data reliability of $R=10$ packets per second (which is shown on Figure 13.10(a)). We note that the total number of packets hyperbolically grows when the key exchange threshold decreases linearly in Figure 13.10(b). This is intuitive because the frequency of key updates is R/n_k per second and number of overhead packets with key information per second is equal to $8R/n_k$. We note that key exchange overhead becomes negligible only for $n_k \geq 90$. Probability that packet will not suffer from

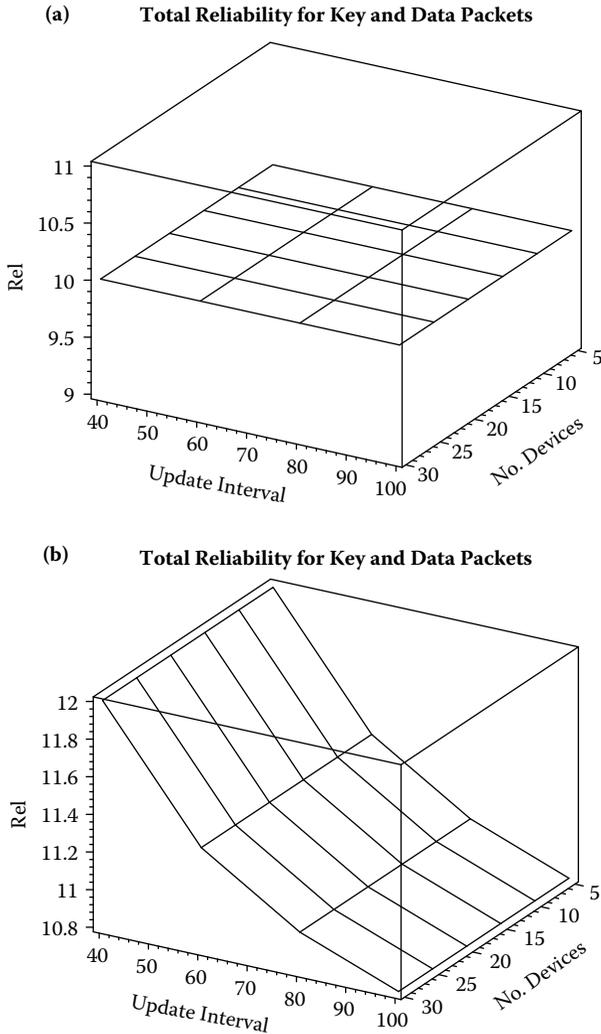


Figure 13.10 Event sensing reliability for (a) and (b) data and key + data, (c) inactive period, (d) total utilization and utilization for key packets, (e) average number of active devices, and (f) sleep probability for a node.

collision or noise error sharply drops when threshold for key exchange drops below 40 packets. Both the reliability overhead and success probability depend only on the requested event sensing reliability except for very small key update threshold. Sleep period, on the other hand, depends mostly on the number of alive nodes and impact of key exchange overhead is barely noticeable.

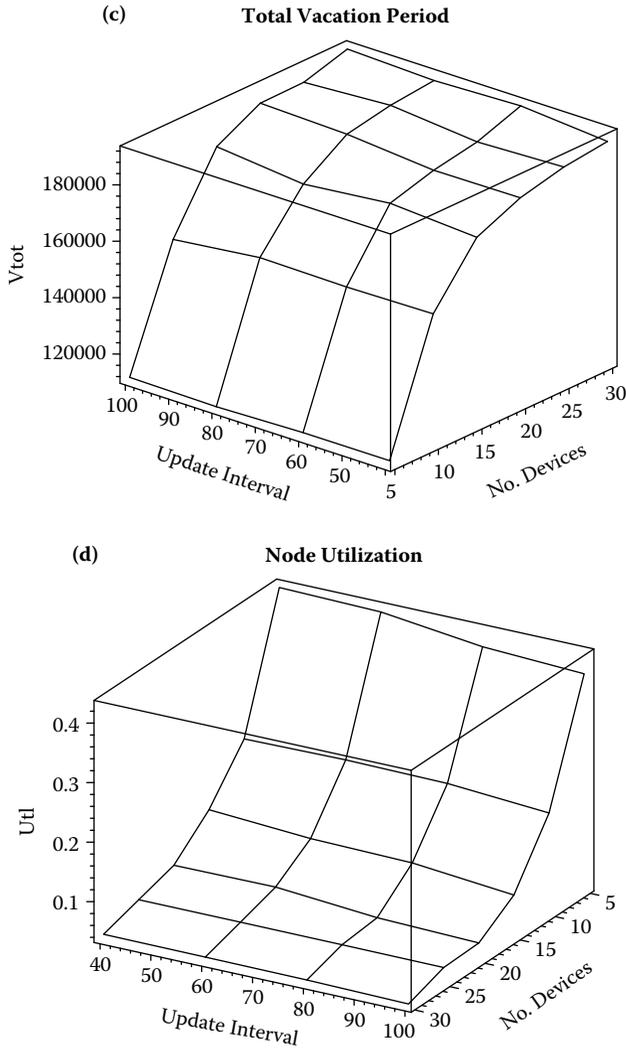


Figure 13.10 (Continued)

Total node utilization shown in Figure 13.10(d) depends mostly on the number of alive nodes, but it also increases with increase of the number of key exchanges per second and exact impact of the key exchange overhead is shown in Figure 13.10(e). Finally, sleep probability for each node is shown in Figure 13.10(f). Sleep probability dominantly changes with n , while the changes with n_k are much milder.

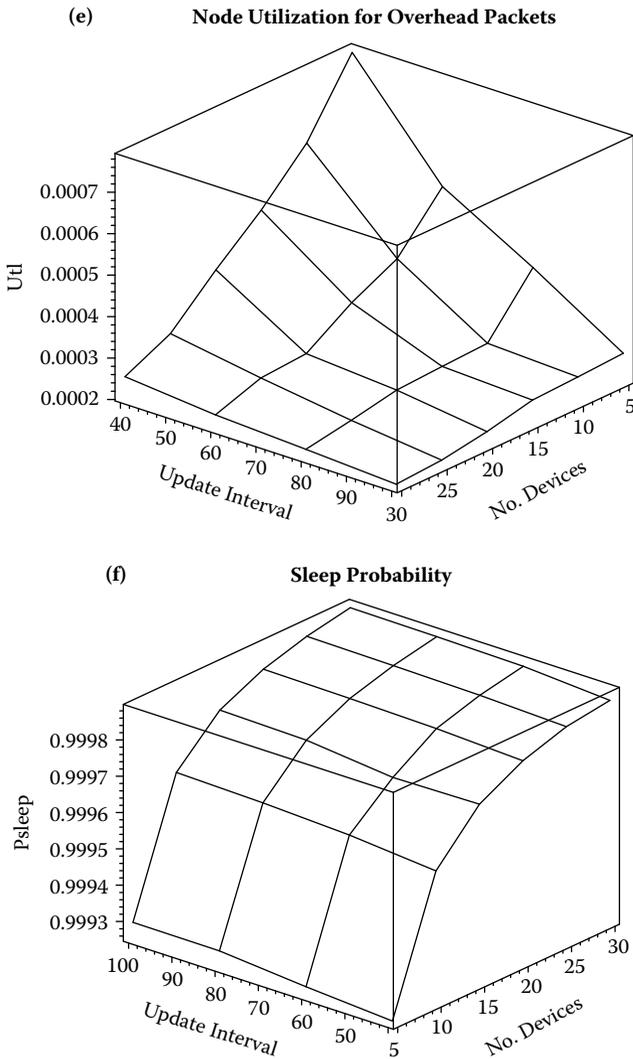


Figure 13.10 (Continued)

13.6 Conclusion and Future Work

We have simulated and studied key exchange process integrated with reliable sensing and power management in beacon-enabled IEEE 802.15.4 cluster and the results confirm our expectations. Data encryption is provided by exchanging link keys between each device and PAN coordinator. The signature payload plays a big role in performance of the network. We have developed a model of key exchange integrated

into the sensing function of a beacon-enabled IEEE 802.15.4 cluster. Our results show important impact of the ratio of the event sensing reliability and key update threshold on the cluster's energy consumption. We have evaluated the impact of the threshold for key update on the cluster's descriptors. The results can give useful hints for the choice of frequency of key updates for required event sensing reliability.

For the future work we will measure more realistically the performance of secure IEEE 802.15.4 personal area network. Combination of other key exchange protocols, activity management, and key management techniques will be compared to get measures that will be used to enhance the lifetime of wireless personal area network along with defense against expected threats from the environment.

References

1. Mišić, J., Shafi, S., Mišić, V.B. Performance of a beacon-enabled IEEE 802.15.4 cluster with downlink and uplink traffic. *IEEE Transactions on Parallel and Distributed Systems*, 17 1–16 (2006).
2. Standard for part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (WPAN). IEEE Std 802.15.4, IEEE (2003).
3. Stallings, W. *Cryptography and Network Security: Principles and Practice*. Prentice Hall, Upper Saddle River, NJ (2003).
4. ZigBee specification (ZigBee document 053474r06, version 1.0). ZigBee Alliance (2004).
5. Sastry, N., Wagner, D. Security considerations for IEEE 802.15.4 networks. In: *WiSe '04: Proceedings of the 2004 ACM Workshop on Wireless Security*. 32–42 (2004).
6. Bellare, M., Kilian, J., Rogaway, P. The security of the cipher block chaining message authentication code. *Computer and System Sciences* 61: 362–399 (2000).
7. Whiting, D., Housley, R., Ferguson, N. Counter with cbc-mac (CCM). <http://www.rfc-archive.org/getrfc.php?rfc=3610> (2003).
8. Menezes, A., Oorschot, P.V., Vanstone, S. *Handbook of Applied Cryptography*. CRC Press, Boca Raton, FL (1997).
9. FIPS Pub 198, The Keyed-Hash Message Authentication Code (HMAC). Federal Information Processing Standards Publication 198, U.S. Department of Commerce/N.I.S.T. (2002).
10. ANSI X9.63-2001, Public Key Cryptography for the Financial Services Industry—Key Agreement and Key Transport Using Elliptic Curve Cryptography. American Bankers Association (2001).
11. Wang, W., Bhargava, B. Key distribution and update for secure inter-group multicast communication. In: *SASN '05: Proceedings of the 3rd ACM Workshop on Security of Ad Hoc and Sensor Networks*, ACM Press, New York, 43–52 (2005).
12. Zhang, X.B., Lam, S.S., Lee, D.Y., Yang, Y.R. Protocol design for scalable and reliable group rekeying. *IEEE/ACM Transactions on Networking*, 11: 908–922 (2003).
13. Stojmenović, I., Ed. *Handbook of Sensor Networks: Algorithms and Architectures*. John Wiley & Sons, New York, (2005).

14. Sankarasubramaniam, Y., Akan, Ö.B., Akyildiz, I.F. E SRT: event-to-sink reliable transport in wireless sensor networks. In: *Proceedings of the 4th ACM MobiHoc*, Annapolis, MD 177–188 (2003).
15. Inc., R.D. Artifex v.4.4.2 (2003).
16. Shafi, S. Performance of a beacon enabled IEEE 802.15.4-compliant network. Master's thesis, Department of Computer Science, University of Manitoba, Winnipeg, Canada (2005).

NETWORKING SUPPORT

5

Chapter 14

**Fourth Generation
Heterogeneous Wireless
Access Networks for
eHealth Services:
Architecture and Radio
Resource Management**

Dusit Niyato, Ekram Hossain, and Jeffrey Diamond

CONTENTS

14.1 Introduction.....269
14.2 Fourth Generation Heterogeneous Wireless Access Networks270
 14.2.1 Wireless Access Technologies270
 14.2.2 4G Heterogeneous Wireless Networks.....272
 14.2.3 Research Issues in Heterogeneous Wireless Access
 Networks273

14.3	4G Wireless Systems and eHealth Services.....	275
14.3.1	Different Types of eHealth Services	275
14.3.1.1	Follow-Up Service.....	275
14.3.1.2	Telehomecare Service.....	276
14.3.1.3	Perihospital Service.....	276
14.3.1.4	Mobile Health Care Service.....	276
14.3.1.5	Intrahospital Monitoring System	276
14.3.1.6	Medical Information Management Service.....	277
14.3.2	Application of 4G Heterogeneous Wireless Access Systems for eHealth Services	277
14.3.2.1	Limitations of Traditional Wireless eHealth Systems.....	277
14.3.2.2	Usage Scenarios of eHealth Services in a 4G Heterogeneous Wireless Access Network.....	278
14.4	A General 4G Wireless eHealth Network Architecture and the Related Research Issues.....	280
14.4.1	4G Wireless eHealth Network Architecture.....	280
14.4.1.1	eHealth Agent	280
14.4.1.2	Intelligent Software-Defined Radio	282
14.4.2	Research Issues	284
14.5	Bandwidth Allocation and Admission Control for eHealth Services in a Heterogeneous Wireless Environment: A Game- Theoretic Approach.....	285
14.5.1	System Model.....	286
14.5.2	Bandwidth Allocation and Admission Control	287
14.5.2.1	Bankruptcy Game	287
14.5.2.2	Bandwidth Allocation	289
14.5.2.3	Admission Control	290
14.5.3	Numerical Results.....	290
14.5.3.1	The Core and the Shapley Value	290
14.5.3.2	Performances of Bandwidth Allocation and Admission Control Algorithms	291
14.6	Conclusions	292
	References	293

The evolving fourth generation (4G) wireless networks are envisioned to provide high-speed wireless communication services with quality-of-service (QoS) support and seamless mobility in a heterogeneous wireless access environment. One of the potential applications of heterogeneous 4G wireless systems is to provide telemedicine/eHealth services in the health care sector. In this chapter we provide a survey of the different wireless access technologies for 4G wireless systems, identify

the communication requirements for the different health care services, and describe the different eHealth usage scenarios in such a network. The radio resource management issues (e.g., bandwidth allocation, admission control, and load balancing) in a 4G wireless eHealth network architecture are discussed. To this end, a game-theoretic approach to bandwidth allocation and admission control is presented for a heterogeneous wireless access network to provide eHealth services. This bandwidth allocation approach is based on the cooperative game theory in which multiple types of wireless networks cooperatively offer bandwidth to a new connection and thus achieve load balancing among the different wireless networks. The numerical results show that this bandwidth allocation scheme can improve radio resource utilization as well as connection blocking probability performance for eHealth services.

14.1 Introduction

The fourth generation (4G) wireless communications systems are envisioned to integrate multiple wireless access technologies to provide high-speed multimedia communications services to mobile users in a seamless manner. In such a heterogeneous wireless access environment, a mobile will be able to connect to multiple wireless networks simultaneously to achieve high data rate through load balancing, global mobility through seamless handoff, and quality-of-service (QoS) support through tight integration of network services with applications in the higher layers. In this way, wireless services can be provided anytime, anywhere, regardless of the type of available wireless access to a mobile user.

To achieve the above objectives, the available protocols and technologies (from the physical to the application layers) must be revisited; at the same time new techniques need to be developed to provide an efficient and a flexible heterogeneous wireless access architecture. In this chapter, we provide a survey of the wireless communications technologies in a 4G heterogeneous wireless access system and outline the related research issues. The specific requirements of such a network to provide different types of eHealth services are discussed and different usage scenarios are described. A general wireless eHealth network architecture based on heterogeneous wireless access is presented and the related research issues on radio resource management (e.g., resource reservation, bandwidth allocation, admission control) are outlined. To this end, a game-theoretic approach to bandwidth allocation through load balancing in the different access networks, namely, cellular, wireless local area, and wireless metropolitan area networks, is presented for the different types of eHealth services. Performance evaluation results show that the proposed bandwidth allocation method through load balancing can improve system performances in terms of resource utilization and connection blocking probability.

The rest of the chapter is organized as follows. A survey of the different wireless access technologies to be integrated in a 4G system is provided in Section 14.2. Also, issues related to radio resource management and QoS provisioning in such a heterogeneous wireless access environment are discussed. Section 14.3 describes the different types of eHealth services and their usage scenarios in a 4G wireless system. Section 14.4 presents the high-level architecture for a 4G wireless-based eHealth network and identifies the related radio resource management issues. The game-theoretic approach for bandwidth allocation for eHealth services and the performance evaluation results are presented in Section 14.5. Section 14.6 states the conclusions.

14.2 Fourth Generation Heterogeneous Wireless Access Networks

14.2.1 Wireless Access Technologies

- *Wireless metropolitan area network (WMAN)*: WMAN based on IEEE 802.16/WiMAX¹ and 802.20 (MobileFi) for fixed and mobile broadband wireless access (BWA) will provide high-speed wireless connectivity in large coverage areas to support multimedia and other Internet applications. IEEE 802.16 operates either at 10 to 66 GHz or 2 to 11 GHz band (IEEE 802.16a) and it supports three different air interfaces: single-carrier time-division multiple access (TDMA), orthogonal frequency division multiplexing (OFDM)/TDMA, and orthogonal frequency division multiple access (OFDMA). In OFDM, high-speed transmission is achieved by transmitting data using several subcarriers at the same time. OFDM technique is also robust to intersymbol interference and frequency-selective fading. Adaptive modulation and coding (AMC) is used to enhance transmission rate according to channel quality. IEEE 802.16d (802.16-2004) and IEEE 802.16e, which evolved from 802.16a, use advanced physical layer techniques to support non-line-of-sight communication. IEEE 802.16e is specifically designed to support user mobility. The IEEE 802.16g standard aims to support mobility at higher layers (transport and application) and across the backhaul network for multinet network operation.

IEEE 802.16 supports two operation modes, namely, point-to-multipoint and multihop (mesh). In point-to-multipoint mode (referred to as last-mile BWA), the base stations (BSs) control all communications to and from the subscriber stations (SSs). On the other hand, in the mesh mode, the SSs cooperate with each other to relay traffic to an Internet gateway through multihop communication. Three major types of services, namely, unsolicited grant service (UGS), polling service (PS), and best-effort (BE) service, are supported within the IEEE 802.16 QoS framework. IEEE 802.16-based technology can be used for providing broadband wireless telemedicine/eHealth services.²

IEEE 802.20/MobileFi-based systems are expected to provide mobile BWA with data transmission speed over 10 Mbps using frequency spectrum below 3.5 GHz for users with high mobility (< 250 km/hour). At the physical layer, OFDM-based wireless transmission will be used. IEEE 802.20 will be optimized for IP services for fixed and mobile users.³ In addition, to extend the transmission range of the base station a mesh network can be formed by multiple mobile stations.

- *Cellular network*: By adopting advanced radio transmission technologies, the second generation (2G) and the third generation (3G) cellular networks have now evolved to provide high-speed wireless communication for multimedia services. The high-speed downlink packet access (HSDPA) based on wideband-CDMA (code division multiple access), which was proposed by the 3G Partnership Project (3GPP), can support data rate up to 14 Mbps by incorporating AMC as well as hybrid automatic request (HARQ) in the standard. Wireless telemedicine systems based on cellular technology have been developed.⁴
- *Wireless local area network (WLAN)*: Due to low cost and ease of deployment, IEEE 802.11 (WiFi) has become a popular technology for WLAN. IEEE 802.11 supports contention-free and contention-based medium access control (MAC) protocols, i.e., point coordination function (PCF) and distributed coordination function (DCF), respectively. IEEE 802.11b, a, and g standards use 2.4, 5.8, and 2.4 GHz bands to support maximum data rate of 11, 54, and 54 Mbps, respectively. WLANs have been used to provide eHealth services such as home care⁵ and biosignal monitoring in a hospital environment.⁶
- *Wireless personal area network (WPAN)*: IEEE 802.15 was proposed for WPAN for short-range wireless communications (10 to 50 meters). IEEE 802.15.1 is based on Bluetooth v1.1 foundation specifications and this standard supports data rate up to 1 Mbps. IEEE 802.15.3 was designed to support high data rate (20 Mbps or more) in the 2.4 GHz band. In the MAC layer, IEEE 802.15.3 supports both contention-free (e.g., TDMA-based) and contention-based (i.e., carrier sense multiple access/collision avoidance [CSMA/CA]) access schemes. The newer version IEEE 802.15.3a intends to support data rate up to 110 Mbps for multimedia traffic. IEEE 802.15.4 (ZigBee) was designed for low data rate communication (i.e., 250 kbps) with CSMA/CA-based channel access. This is suitable for implementing wireless sensor networks. Bluetooth-based wireless systems have been used for transmission of biosignal sensor data.⁷

The major features of these wireless access technologies are summarized in Table 14.1.

When evolving towards a 4G technology, some of the above systems will adopt advanced physical layer techniques such as multiple input/multiple output (MIMO) and ultra-wideband transmission (UWB).

Table 14.1 Major Features of Wireless Access Technologies

<i>Standard</i>	<i>Bandwidth (Mbps)</i>	<i>Frequency</i>	<i>Spectrum Type</i>	<i>Spectrum Size (MHz)</i>
802.16e	70	2.3, 2.5, 3.5, 3.7, 5.8 GHz	Licensed	10
W-CDMA HSDPA	14.4	850 MHz; 1.9, 1.9/2.1, 1.7/2.1 GHz	Licensed	10
UMTS-TDD	16	450, 850 MHz; 1.9, 2, 2.5, 3.5 GHz	Licensed	5
802.11a	54	5.25, 5.6, 5.8 GHz	Unlicensed	20
802.11b	11	2.4 GHz	Unlicensed	20
802.11g	54	2.4 GHz	Unlicensed	20
802.11n	200	2.4 GHz	Unlicensed	20 or 40

By transmitting and receiving data over multiple antennas, a MIMO system can improve the transmission rate and the error performance significantly. MIMO takes advantage of space diversity due to independent multipath to create multiple subchannels to improve the transmission performance. In the IEEE 802.16 standard, MIMO implementation in the radio transceiver is optional. MIMO will be integrated with 3GPP's HSDPA to double the transmission rate. IEEE 802.11n is being designed specifically for WLAN with MIMO links.

In UWB networks, data transmission takes place over a very large bandwidth using low transmission power density. UWB systems can be implemented in two different approaches, i.e., pulse-based and multiband-orthogonal frequency-division multiplexing (MB-OFDM). In the pulsed-based approach, a signal pulse is transmitted in a short period of time. On the other hand, in MB-OFDM, an OFDM technique is combined with frequency hopping. This approach will be adopted by the IEEE 802.15.3a WPAN standard.

14.2.2 4G Heterogeneous Wireless Networks

The next generation (4G) wireless networks will have the following features:⁸

- *Heterogeneous environment:* Heterogeneity of the wireless access environment will be a major feature of the evolving 4G wireless networks. This will enable high-speed wireless communications under seamless mobility. In a heterogeneous environment, multiple service providers will collaboratively offer wireless access services to mobiles with multi-radio interfaces.

- *Multiple types of services:* For different wireless applications, 4G networks must support different QoS guarantees. For example, wireless telemedicine/eHealth applications may require transmission of huge volume of non-real-time diagnostic data (e.g., images) and real-time video data simultaneously. QoS support should be provided in different mobility scenarios (e.g., in a mobile ambulance environment as well as in a static in-hospital environment).
- *Adaptive resource allocation:* An adaptive and efficient radio resource management framework is required to meet the diverse requirements of wireless applications. A simple objective of such a framework is to maximize the service provider's revenue (and/or maximize the satisfaction of mobile users) under constraints on available radio resource and QoS performance requirements.

An example of a 4G heterogeneous wireless access scenario is shown in Figure 14.1 in which there are three wireless technologies, namely, IEEE 802.16-based WMAN, CDMA cellular, and IEEE 802.11 WLAN (which could be operated by the same or different service providers). A mobile can connect to multiple wireless access networks simultaneously.

14.2.3 Research Issues in Heterogeneous Wireless Access Networks

- *Load balancing and network selection:* While the aim of network selection is to choose the best network to connect to (based on the network status and the user's requirement), that of load balancing is to achieve fair resource allocation and congestion avoidance in the access networks (e.g., the lesser the congestion, the larger the amount of bandwidth allocated to a connection from an access network). In Song and Jamalipour,⁹ a network selection mechanism was proposed for a 4G heterogeneous wireless network consisting

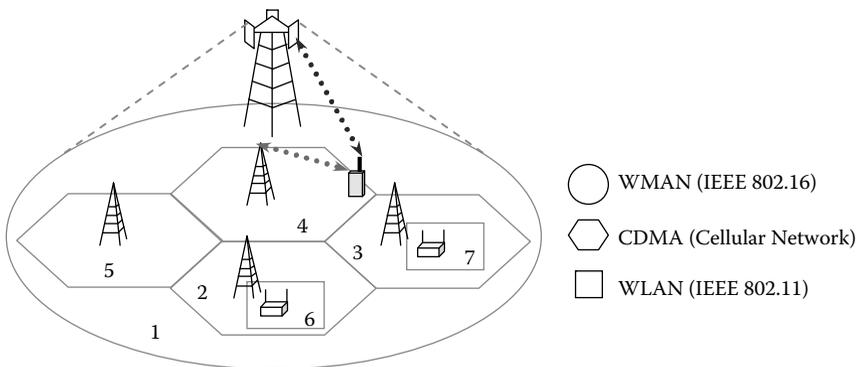


Figure 14.1 Service area under consideration in a heterogeneous wireless access environment.

- of UMTS and WLAN. This scheme contains two major parts to evaluate and to rank the networks based on the preferences of the mobiles. For the first part, an analytic hierarchy process was applied to compute the relative weights of evaluation criteria sets based on the application requirements. For the second part, a grey relational analysis was used to rank the network alternatives. Several QoS parameters (i.e., throughput, timeliness, reliability, security, and cost) were taken into account in this network selection scheme.
- *Resource allocation and admission control:* Because the QoS requirements for mobile users need to be satisfied in a seamless manner, the resource allocation and admission control methods in a 4G heterogeneous wireless access network should consider the states of the different networks.⁸ Due to mobility, a new connection may eventually need to be handed off vertically to a new network. Therefore, resource allocation can be based on cooperation among the different access networks so that network utility can be maximized. Also, both packet-level (e.g., throughput and packet delay) and connection-level (e.g., new connection blocking and handoff connection dropping probabilities) performances must be considered while developing a radio resource management policy.
 - *Vertical handoff and mobility management:* In a heterogeneous wireless access environment, fast and efficient vertical handoff mechanisms (e.g., for handoff between WLANs and cellular networks) are required to achieve seamless mobility and QoS guarantee. A vertical handoff mechanism needs to consider not only the radio link and the physical layer parameters but also the network and the transport level parameters. Therefore, handoff detection policy and decision metrics used in traditional wireless systems may not be applicable to 4G networks. Also, handoff protocols need to be redesigned. In McNair and Zhu,¹⁰ a framework for vertical handoff was presented where the handoff decision metrics include service type, data rate requirement, network condition, and cost of handoff. A dynamic optimization model was proposed to provide a QoS guarantee to the mobile users while maximizing the network utilization.
 - *Protocols for multihop communications and mesh networking:* Multihop communications will be used in 4G networks (e.g., both IEEE 802.11 and 802.16 support mesh operation mode) to enhance network performance and service availability.¹¹ IEEE 802.16-based mesh networks can serve as backhauls for cellular networks and WLAN hotspots. Efficient resource allocation and QoS support mechanisms need to be developed for multihop mesh networking.
 - *QoS provisioning:* A unified QoS provisioning framework would be required for 4G heterogeneous wireless access networks. In Gao, Wu, and Miki,¹² the state-of-the-art QoS techniques and the standardization activities were summarized. Then, a ubiquitous QoS framework for heterogeneous wireless access environment was presented, which integrates a three-plane network infrastructure in interacting with a terminal-based cross-layer adaptation framework. In this framework, different classes of vertical handoff were used

for different types of users with different QoS requirements. This three-plane infrastructure was comprised of management function, QoS control, and packet transfer. While management function and QoS control are used to govern the QoS support parameters, packet transfer utilizes those parameters to achieve QoS guarantee to the users.

- *Application of cognitive radio techniques:* Cognitive radio is being considered as a useful technique to improve radio spectrum utilization and communication reliability. By using intelligent software-defined radio, a cognitive radio transceiver is able to adjust its radio bandwidth usage adaptively. Cognitive radio techniques can be used in a 4G heterogeneous wireless access network. In particular, an intelligent software radio at the mobile can estimate and learn the system behavior to obtain interference and congestion information in different access networks. Then, based on this information, a decision can be made to choose the best network to connect to.

14.3 4G Wireless Systems and eHealth Services

14.3.1 Different Types of eHealth Services

To provide improved health care services, advanced telecommunication technologies have been adopted widely for eHealth services such as remote patient monitoring and trauma care of injured patients. 4G heterogeneous wireless access networks can be used to provide improved eHealth services in both indoor (e.g., in-hospital) and outdoor (e.g., mobile ambulance) environments. In general, eHealth services include follow-up service, home care service, pre-hospital service, and medical information management service used by physicians and medical staff to process patient information.

14.3.1.1 Follow-Up Service

Follow-up services are provided to a patient in stable condition to monitor the biosignal data periodically so that if any unusual condition is detected in the patient's monitored biosignal data, proper treatment can be applied promptly. Follow-up services (e.g., for monitoring ECG signals) were designed based on different wireless technologies.^{13–17} For example, the *airmed-cardio* system proposed in Salvador et al.,⁴ developed based on GSM (Global System for Mobile Communication) and wireless application protocol (WAP), is an example of such a system. This was designed to provide out-of-hospital follow-up services to cardiac patients. It was found that such follow-up services can reduce the frequency of face-to-face meetings between patients and health care staffs and also can shorten the time required for a adaptation of treatment and improve dose control for an individual patient.

14.3.1.2 *Telehomecare Service*

Telehomecare is generally used for rehabilitation of the patients to minimize the number of visits for therapists, and thereby it reduces the risk involved in moving the patients. In Guillen et al.,¹⁸ a multimedia telehomecare system was presented, which enabled online interaction between a patient and the physician as well as online remote monitoring of electrocardiogram (ECG) signal, heart sound, and blood pressure signals. Similar systems were also proposed in Braecklein et al. and Kao et al.^{19,20}

14.3.1.3 *Prehospital Service*

Prehospital services provide efficient trauma care to injured patients to reduce mortality and morbidity from trauma due to accidents. In this case, paramedical staff can provide initial assessment, resuscitation, and treatment while the patient is transported to the hospital. In Chu and Ganz,²¹ a prototype of a mobile teletrauma system using a 3G wireless network was developed. This system was deployed on a tablet personal computer in an ambulance and the patient information (i.e., video, medical image, and ECG signal) can be transmitted simultaneously through a cellular network. Also, emergency patients can be benefited through the online communications to and from ambulances before arriving at hospital. Telediagnosis, long-distance support, and teleconsultation by expert physicians from a mobile health care system can increase the survival rate of patients significantly.²²

14.3.1.4 *Mobile Health Care Service*

The quality of health care can be improved through early diagnosis and treatment by using mobile eHealth service. This also reduces the cost of patient transportation. In Takizawa et al.,²³ an eHealth system using a computed tomography (CT) van was developed to provide health care services to patients in remote areas. The CT images are transferred via satellites to the health care center for diagnosis.

14.3.1.5 *Intrahospital Monitoring System*

In a hospital, the vital biosignal data of a patient can be monitored and transmitted to the corresponding medical staff through wireless communication. In Lin et al.,⁶ a portable physiological monitoring system based on wireless PDA was presented. In this system, sensor devices were connected to the PDA to record biosignal (e.g., heart rate and three-lead electrocardiography). The PDA transmitted the biosignal data to a management unit in real-time using a WLAN connection.

14.3.1.6 Medical Information Management Service

Medical information management services provide control and management services to the physicians and health care staffs. These services can be provided in both offline and online manners. For offline operation, pre-recorded or historical data stored in personal medical information servers are retrieved through a hospital's intranet to be processed. For example, after biosignal data or radiology images are collected, they are analyzed and the results of this analysis are used for scheduling and planning treatment as well as health care services to the patients. On the other hand, for online operation, an injured patient can be observed and diagnosed through prehospital service in real-time so that preliminary treatment can be provided instantly.

In a clinical grade health care network²⁴ (i.e., which provides a complete information system for clinicians) medical data including medical history records, patient treatments, test results, diagnoses for current admission, patient monitoring data, and laboratory tests for ongoing treatment need to be communicated among clinics and hospitals/health care centers. Also, applications such as multimedia collaboration among doctors, pharmacists, and drug dosage management staff need to be supported.

14.3.2 Application of 4G Heterogeneous Wireless Access Systems for eHealth Services

14.3.2.1 Limitations of Traditional Wireless eHealth Systems

Traditionally, eHealth services have been provided using wired communication channels such as telephone and cable modems. The major shortcoming of these technologies is the lack of mobility. By using 4G wireless systems, the quality of the eHealth services can be improved with higher availability. Because the traditional wireless eHealth services mostly rely on a single wireless technology, the usability, manageability, and performance of these systems are limited.

- *Usability:* The limitation of a single wireless technology for eHealth services results due to limited coverage area. For example, if the patients need to be transported out of the hospital, follow-up systems based on WLAN cannot be used. In this case, systems with multi-interface wireless technologies are needed to enhance the usability of the wireless service.
- *Manageability:* eHealth services/systems developed using different wireless technologies are often difficult to manage and to operate in an efficient way. For example, inter- and intrahospital follow-up systems based on cellular networks and WLANs require different system operators and also the radio equipment may not interoperate.

- *Performance:* The major limitation of a wireless eHealth system using a single wireless technology arises due to the fact that each network has limited capacity and most of the time the network must serve several types of users (e.g., follow-up patient needs to share wireless channels with voice users in a cellular network). In this case, a heterogeneous wireless access environment is desired to share the total load efficiently.

14.3.2.2 Usage Scenarios of eHealth Services in a 4G Heterogeneous Wireless Access Network

Figure 14.2 shows how the different eHealth services can be provided through a heterogeneous wireless access network. Four different usage scenarios of eHealth services in a 4G wireless network are shown in Figure 14.3.

- *Follow-up service:* Because follow-up patients can move freely from one place to another (e.g., from indoors to outdoors and vice versa), flexibility of the eHealth services can be improved in a heterogeneous access network by using vertical handoff among the different networks. To achieve this capability, the radio equipment can be employed with multiple wireless interfaces (e.g., cellular network, WMAN, and WLAN). When the patient is at an outdoor location, cellular network and WMAN are preferable. However,

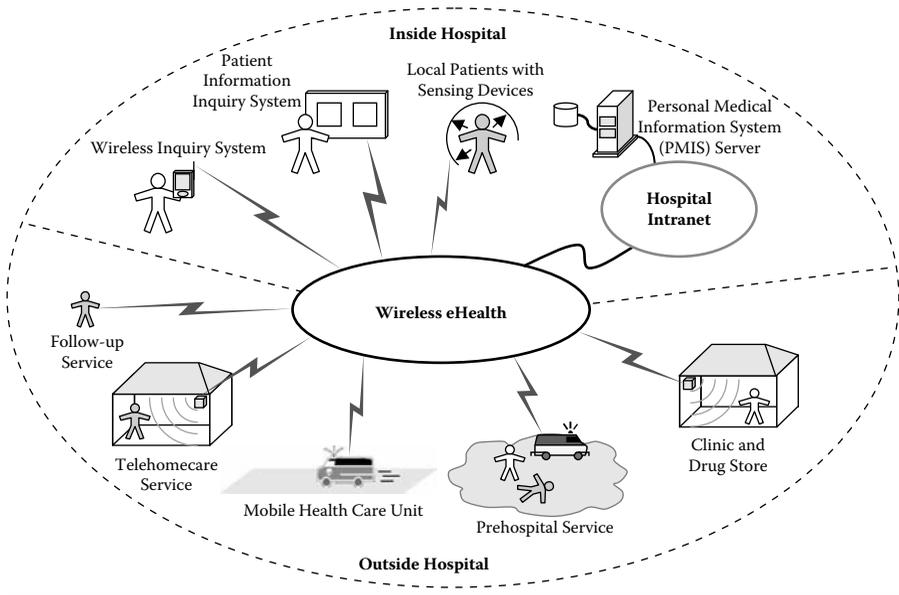


Figure 14.2 4G wireless system and eHealth services.

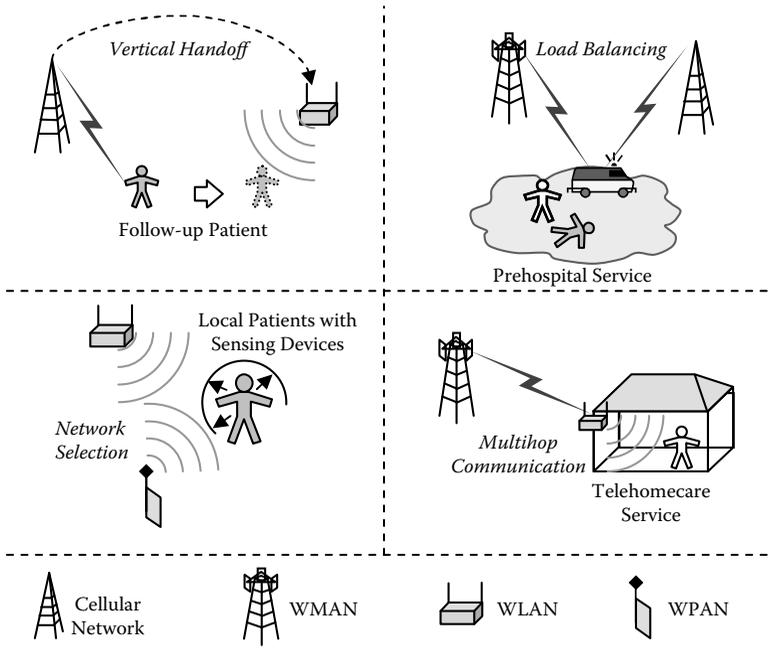


Figure 14.3 Usage scenarios for eHealth services based on 4G wireless systems.

if the patient is inside the hospital or in a place where wireless access to a hotspot is available, follow-up service can be handed off vertically to the corresponding WLAN whose bandwidth is much larger than that of cellular network and WMAN.

- *Intrahospital service:* Inside a hospital, a biosignal monitoring system can be developed based on a wireless personal area network (WPAN), for example, a ZigBee-based wireless sensor network, which can interoperate with a WLAN or a cellular wireless network. In this scenario, intelligent and effective network selection policy is required.
- *Prehospital service:* High-speed wireless transmission with QoS support would be required for video and voice communications between injured patients and physicians during prehospital service. While an ambulance is moving, eHealth traffic can be transferred through a WMAN or cellular network. In this case, radio resource reservation is needed to achieve zero blocking probability for emergency connections originating from mobile ambulances.
- *Telehomecare and mobile eHealth service:* Telehomecare and mobile eHealth services can be provided through an integrated WLAN/cellular/WMAN network architecture. For example, biosignal monitoring/videoconferencing equipment can connect to WLAN access points (APs), which can in turn connect to a health care center through WMAN BSs.

14.4 A General 4G Wireless eHealth Network Architecture and the Related Research Issues

14.4.1 4G Wireless eHealth Network Architecture

A general 4G wireless eHealth network architecture for eHealth services is shown in Figure 14.4. This architecture basically integrates the hospital intranet with the service provider's infrastructure.

- Hospital intranet is used as the main infrastructure inside the hospital to support health care applications. This intranet is connected with the Internet and also with the enterprise wireless network such as WLAN hotspots and WPANs (in-hospital sensor network for patient monitoring) inside the hospital.
- Service provider's infrastructure for different types of networks (e.g., cellular network, WMAN, and public WLAN hotspot) is responsible for transferring data traffic to and from the hospital/health care center.

14.4.1.1 eHealth Agent

In this architecture, a software eHealth agent is located in the service provider's infrastructure for resource allocation, admission control, and mobility management in a 4G wireless system. This eHealth agent maintains the QoS requirement and application-specific parameters for eHealth services. Although the basic structure of this agent can be the same for different types of health care services, some

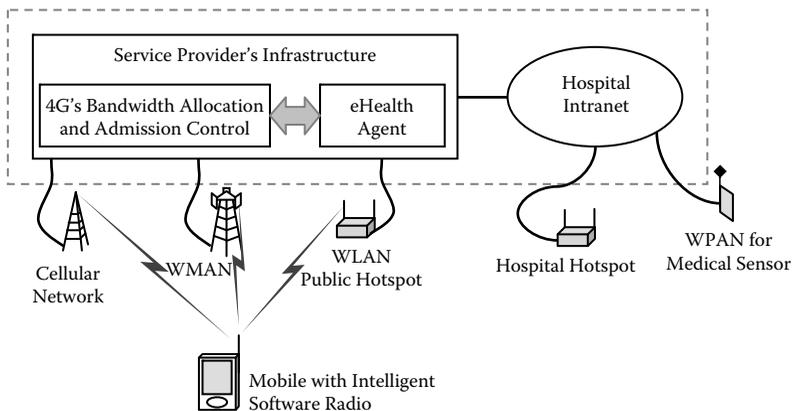


Figure 14.4 4G wireless eHealth network.

details may need to be customized based on the usage scenarios (e.g., follow-up or prehospital service). The functions of the eHealth agent are as follows:

- *Wireless connection management:* When a n eHealth connection is initiated, the service provider (e.g., cellular BS, WLAN AP, WMAN BS) consults with the eHealth agent to retrieve information on the connection's communication requirements. Related information (e.g., QoS requirements and security credential) are exchanged between the agent and the service provider for bandwidth allocation and admission control. Also, the agent will facilitate handoff management (e.g., similar to the home agent in IPv6).
- *Gather data from multiple wireless networks:* The main function of the eHealth agent is to collect data from multiple wireless networks (Figure 14.5). Simultaneous transmission over multiple networks will enhance the application's throughput. This collected data can be communicated to the hospital intranet. For this function, the eHealth agent will also be required to perform security functions such as authentication and authorization.

Note that the concept of agent was used in the context of eHealth service.^{25–30} In Hayes-Roth and Larsson,²⁵ an agent was used in a patient monitoring application in a surgical intensive care unit. It was designed to support collaboration among medical staffs by sharing knowledge. In Koutkias, Chouvarda, and Maglaveras,²⁸ an intelligent eHealth agent was designed to perform real-time patient health monitoring in a home care service. By using biomedical measurement, this agent is able to identify and classify the patient's health problem. Appropriate recommendations and suggestions can be then made to the patient and the medical staff.

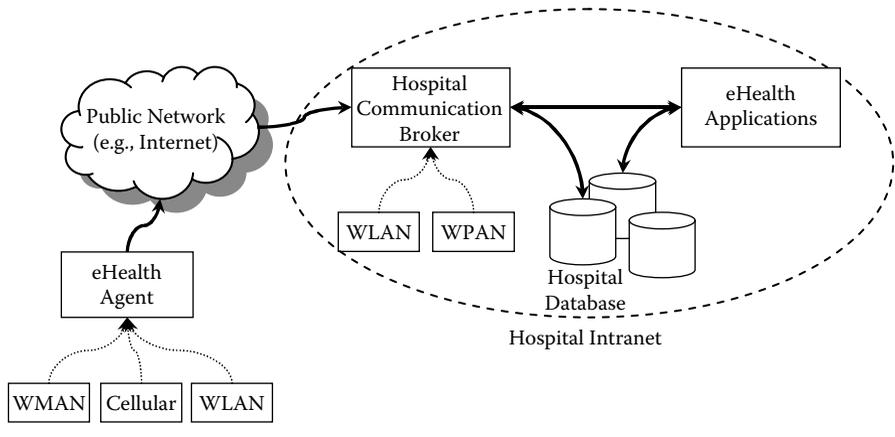


Figure 14.5 eHealth agent and its interaction with an intelligent software radio modem.

14.4.1.2 Intelligent Software-Defined Radio

At the mobile side, an intelligent software radio is required to communicate with the eHealth agent. The function of this software radio can be summarized as follows:

- *Network selection:* With multi-radio interface, a mobile in 4G networks will be able to connect to multiple wireless access networks simultaneously. Therefore, the fundamental function of intelligent software radio is to select an access network among a set of available networks. The network selection in an eHealth environment should be based on the QoS requirement and transmission quality (e.g., signal-to-noise ratio, transmission rate, collision rate) of each network. Also, price of the network access can be taken into account for the selection.
- *Determine transmission parameters:* Based on the selected network, transmission must be performed to satisfy the QoS requirement based on the granted radio resource. The transmission parameters are different depending on the physical and MAC specifications of each network. In a WLAN, the transmission parameters are the channel access probability and bandwidth reservation if it is supported (i.e., this needs special MAC protocol because IEEE 802.11 standard does not support explicit bandwidth allocation). In a CDMA cellular network, transmission rate and transmit power are the transmission parameters. In IEEE 802.16, delay and throughput requirements are key performance parameters for real-time and non-real-time polling services, respectively.
- *QoS management:* At a mobile unit, it is possible to service multiple traffic flows with different QoS requirements for different eHealth applications. QoS mapping would be one of the main functions of an intelligent software radio to schedule eHealth traffic to be transmitted to suitable network interface (Figure 14.6).

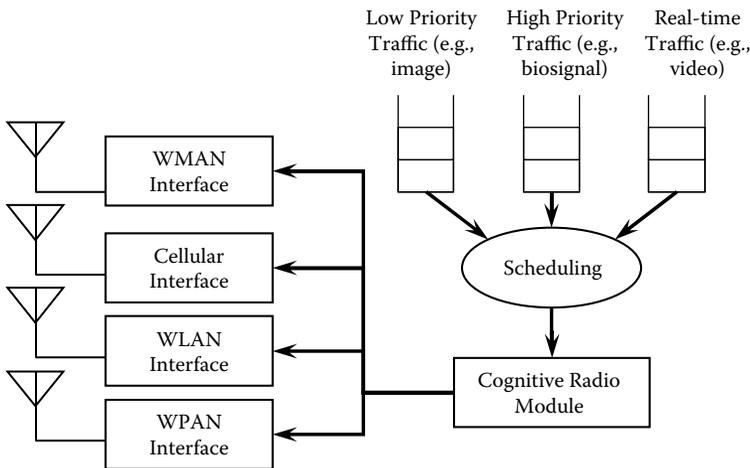


Figure 14.6 Components of an intelligent software radio modem in a mobile unit.

For example, low priority traffic can be transmitted through a WLAN while real-time traffic (e.g., video) can be transmitted through a WMAN by using the QoS support mechanism in IEEE 802.16.

- *Mobility/handoff management:* To provide seamless mobility, an intelligent software radio must include the functionalities to make a handoff decision. Also, it has to determine which network it will be handed over to.

This cognitive radio³¹ module in the intelligent software radio modem (Figure 14.7) is able to observe the radio environment. The observed data is used to gain knowledge of the environment through a learning process. From this knowledge, the mobile is able to plan and optimize wireless access. If any decision has to be made, this will be based on the optimization result. Then, an appropriate action is taken.

In a 4G network-based eHealth service environment, the mobile needs to observe all available wireless services (i.e., WMAN, cellular, WLAN, and WPAN) and also their characteristics (e.g., channel quality, congestion level, and price). After gathering the information, the mobile optimizes the possible plan for transmission in each network according to the QoS requirement of the application. Then, the decision is made so that the desired objective is achieved (e.g., minimize network congestion for follow-up service or minimize transmission delay for the real-time video from ambulances). Based on the optimized decision, network selection is performed and the optimized amount of bandwidth is requested.

One important issue in designing intelligent software radio is the interference management. For example, in an intrahospital environment, different locations (e.g., surgical rooms, patient bed area) have different acceptable levels of electromagnetic interference. Because some medical equipment (e.g., microelectronic sensor with radio communication module) is sensitive to electromagnetic interference, intelligent software radio must be aware of such constraints and adapt the transmission accordingly (i.e., frequency band and transmit power). In this case,

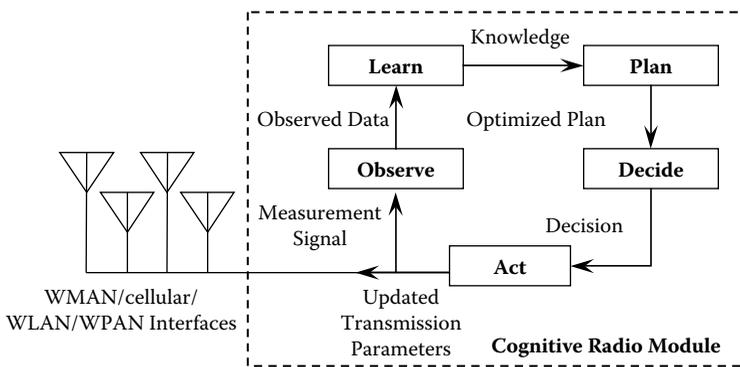


Figure 14.7 Cognitive radio module in an intelligent software radio modem.

an intelligent software radio can be “location-aware.” This location information can be used for transmission adaptation. For example, wireless transmission in a surgical room, which consists of many pieces of medical equipment (e.g., wireless operating room controllers, wireless monitors, high-frequency surgical devices, and diathermy), is more constrained than that in an intermediate care unit in which basic biosignal monitoring systems (e.g., ECG) are installed. Also, to avoid interference to the communication among the medical equipment, spectrum sensing should be performed efficiently in an intelligent software radio. Many spectrum sensing techniques have been proposed in the literature.^{32–35}

14.4.2 Research Issues

Several research issues related to this wireless eHealth architecture can be outlined as follows:

- *Resource reservation and admission control:* Because different eHealth services have different priorities and also eHealth traffic shares the wireless resources with normal data traffic, resource reservation and admission control methods need to be designed accordingly. Emergency connections (corresponding to prehospital service) from ambulances carrying injured patients must be prioritized over all other connections and should experience zero blocking probability. However, connections corresponding to mobile health care and follow-up services have less priority but still are more important than normal traffic.
- *Load balancing:* This is important for transmission of rate-sensitive eHealth traffic corresponding to prehospital and mobile health care services, which require voice and video communications as well as file transfer. Because these connections may require bandwidth from the different access networks simultaneously, efficient load balancing schemes would be required. A game-theoretic method for load balancing in a heterogeneous wireless access system will be presented later in this chapter.
- *Intelligent software radio:* Software radio at the mobiles and APs/BSs can learn and intelligently adapt to the network conditions to improve network performance. This adaptation should be performed in a cross-layer fashion.
- *Multihop communications:* Different from the homogeneous case, asymmetric wireless link characteristics (e.g., interference, capacity, and congestion in different frequency bands) in a heterogeneous wireless access environment must be taken into account while designing packet routing, scheduling, and radio resource management protocols for multihop communications. For example, in telehomecare service, traffic from multiple follow-up patients can be aggregated through a WLAN AP and transmitted by a WMAN to the hospital in a multihop fashion.

14.5 Bandwidth Allocation and Admission Control for eHealth Services in a Heterogeneous Wireless Environment: A Game-Theoretic Approach

As shown in Figure 14.3, one of the major issues in a 4G wireless eHealth network is load balancing. Load balancing among multiple wireless access networks can improve the QoS performance of the service significantly by utilizing simultaneous transmission over multiple wireless interfaces.

In this section, we propose a bandwidth allocation and admission control method for eHealth services in a 4G heterogeneous wireless access network in which a mobile can connect to multiple networks simultaneously. This method is based on load balancing across the access networks. For load balancing and satisfying users' QoS requirements, multiple networks need to cooperate while offering bandwidth to a connection. Therefore, we formulate this situation as a bankruptcy game, which is a special type of N-person cooperative game.

In a scenario where the different access networks are operated by different operators, the solution of a cooperative game model can satisfy all the operators (i.e., players in the game). In particular, a coalition among the networks can be formed and the amount of bandwidth to be offered to the new connection by the different networks can be determined. The solution is fair and can achieve an effective load balancing among the different access networks.

Through a cooperative game-theory framework, the stability of the bandwidth allocation scheme is analyzed by using the concept of core. The amount of allocated bandwidth to a connection in each network is determined from the Shapley value, which is one of the solutions in an N-person cooperative game. Then, an admission control algorithm is proposed to ensure that the QoS requirements of the ongoing and the incoming connections can be met.

Two major techniques are used in radio resource allocation for wireless networks, i.e., classical optimization and game theory. In classical optimization, a single objective is defined for the entire system that is to be optimized. Also, many constraints can be incorporated in the formulation. Standard techniques in classical optimization (e.g., linear programming, linear integer programming, convex optimization, Hungarian method, etc.) can be applied to obtain a solution that is systemwise optimal. That is, a single objective of the system is maximized or minimized. On the other hand, in a system consisting of multiple competing entities, a single objective cannot be defined due to the self interest of each of these entities (e.g., each user wants to maximize its QoS performance). In such a system, a game-theoretic solution (i.e., equilibrium) can ensure satisfaction for each of the entities involved.

Game theory was applied to the problem of resource allocation in wireless networks.^{36–40} In Yaiche, Mazumdar, and Rosenberg,³⁶ a bargaining game was formulated for bandwidth allocation for elastic services (i.e., QoS performance of traffic

can be adapted according to radio resource availability) in wired networks. The Nash bargaining solution from cooperative game theory, which provides an efficient and fair solution, was considered. Also, a distributed algorithm was proposed to achieve the solution. In Lin et al.,³⁷ the problem of admission control in a CDMA network was formulated as a noncooperative game considering churning of user from one service provider to another. The Nash equilibrium was considered as the solution, which satisfies both the user and the service provider. In Xia,³⁸ game theory was used to address the admission control problem in large-scale media delivery systems. A cooperative game was formulated between proxy and media servers to allocate disk bandwidth. A fair solution was obtained from this game formulation.

None of these works considered the bandwidth allocation problem in a heterogeneous wireless networking scenario in which a single user can simultaneously use services offered by multiple wireless networks.

14.5.1 System Model

We consider the service area shown in Figure 14.1 in which each sub-area is covered by one or more wireless access networks, namely, IEEE 802.11-based WLAN, IEEE 802.16-based WMAN, and CDMA-based cellular networks. In particular, in sub-area 1, only WMAN services are available. In sub-areas 2, 3, 4, and 5, services from cellular networks and WMAN are available. In sub-areas 6 and 7, a mobile can connect to all of the wireless access networks. A perfect power control is assumed to ensure uniform available transmission rate across the coverage area. A mobile uses multiple radio transceivers so that it can connect to multiple radio access networks simultaneously.

We consider emergency prehospital service, mobile health care service, and follow-up service. Traffic corresponding to emergency prehospital connections consists of ECG, voice and video traffic, and traffic corresponding to mobile health care services consists of ECG, voice, video, and file transfer for radiology images. For follow-up service, we consider only ECG traffic. We assume that the connections for emergency prehospital services have the highest priority and those for mobile health care and follow-up services have higher priority than normal voice and data connections. We assume that the connection holding times for prehospital, mobile health care, follow-up services, and normal data services are exponentially distributed with means of 20, 30, 45, 15, and 5 minutes with bandwidth requirements of 500 kbps, 700 kbps, 24 kbps, 200 kbps, and 16 kbps, respectively. Connection arrival processes are assumed to be Poisson.

- *WLAN*: We consider IEEE 802.11 WLAN radio interface with early backoff announcement (EBA),⁴¹ which is an enhanced version of DCF. This medium access control (MAC) protocol is compatible with the CSMA/CA protocol in the IEEE 802.11 standard. By incorporating the backoff information into

the MAC header, mobiles can completely avoid collisions. The data transmission rate is 11 Mbps and the maximum saturation throughput in a WLAN achieved through EBA is 6.2 Mbps.⁴¹

- *Cellular network*: We consider a wideband CDMA cellular wireless access system.⁴² In order to satisfy a particular bit-error-rate requirement, the corresponding E_b/I_e must be guaranteed. We assume that the transmission bandwidth is 5 MHz, and $E_b/I_e=8.17$ dB so that the bit-error-rate is less than 10^{-4} . The total transmission rate in a CDMA cell is 2 Mbps.
- *WMAN*: We consider an IEEE 802.16-based WMAN radio interface, which supports data rate in the range of 10 to 50 Mbps depending on the bandwidth of operation as well as the modulation and the coding schemes used. WirelessMAN-OFDM with TDMA/TDD (i.e., OFDM/TDMA-TDD)-based transmission is assumed using 50 sub channels, each of which has a bandwidth of 320 KHz. The total bandwidth required (including the guard bands) is 20 MHz. The frame size is assumed to be 2 ms. In each subchannel AMC is applied with seven transmission modes. In the MAC layer, the Ss use contention-free (polling) mode to request bandwidth by using BW-request PDUs.

14.5.2 Bandwidth Allocation and Admission Control

14.5.2.1 Bankruptcy Game

Bankruptcy game is a special type of cooperative game which is generally used to divide a limited amount of resources among multiple players (i.e., more than two) in a fair manner. A description of this game is as follows. When a company owing money to N creditors becomes bankrupt, the creditors will divide the bankrupted company's property among themselves. In general, the sum of the claims from all creditors is larger than the value of this property (or the equivalent amount of money). This conflicting situation can be formulated as an N -person game where the creditors of the game are seeking for the equilibrium point to divide the money. A detailed analysis of this bankruptcy game was presented in O'Neill.⁴³

A standard bankruptcy game can be expressed by a finite set of players (i.e., creditors) \mathbb{A} , a real positive number M which denotes the amount of the bankrupted company's money, and a nonnegative value d_i for claim from player i . To satisfy every player, the solution of the bankruptcy game must have the following two properties. First, the money must be completely distributed, and second each player has to obtain a nonnegative amount of money not exceeding this claim. Let x_i denote the solution (i.e., amount of money) for player i .

In a cooperative game, a coalition is established by the players to gain better benefit than playing a game without cooperation. The coalition form is used to represent such a game, and the payoff of coalition is expressed by the characteristic

function. A coalition \mathbb{S} is defined as a subset of \mathbb{A} . In this case, \emptyset and \mathbb{A} denote an empty coalition and a grand coalition, respectively. The coalition form of an N -person game is defined by the pair (\mathbb{A}, v) , where v is a characteristic function of the game. For the bankruptcy game considered here, the characteristic function can be defined as follows:⁴³

$$v(\mathbb{S}) = \max \left(0, M - \sum_{j \in \mathbb{S}} d_j \right) \tag{14.1}$$

for all possible coalition \mathbb{S} . In particular, the characteristic function of a coalition gives the minimum total amount of money that the players in that coalition will receive.

The core is generally used to obtain the stability region for the solution of an N -person cooperative game. The core is established based on the assumption that a player will not agree to receive money less than that the player could obtain without coalition. This core of the game can be determined based on the imputations which are the payoff vectors

$$\mathbb{P} = \left\{ \mathbf{x} = [x_1, \dots, x_n] \mid \sum_{i \in \mathbb{A}} x_i = v(\mathbb{A}), \text{ and } x_i \geq v(\{i\}), \forall i \in \mathbb{A} \right\} \tag{14.2}$$

where $\sum_{i \in \mathbb{A}} x_i = v(\mathbb{A})$ and $x_i \geq v(\{i\}), \forall i \in \mathbb{A}$ denote a group and an individual rational, respectively. In particular, group rational is the highest total payoff that can be achieved by forming a coalition among all players and individual rational is the payoff for which each individual agrees to join the coalition.

An imputation \mathbf{x} is unstable with any coalition \mathbb{S} if $v(\mathbb{S}) > \sum_{i \in \mathbb{S}} x_i$. Specifically, if the imputation is unstable, there is at least one player who is unsatisfied due to the coalition (i.e., that player receives an amount of money which is less than what he could have obtained without joining the coalition). In other words, that coalition is infeasible, and therefore, cannot be formed. The stable imputation must satisfy all possible coalitions. Therefore, the core is defined as the set of stable imputations, and hence, it represents the area corresponding to a set of stable solutions for all players and can be expressed mathematically as follows:

$$\mathbb{C} = \left\{ \mathbf{x} = [x_1, \dots, x_n] \mid \mathbf{x} \in \mathbb{P} \text{ and } \sum_{i \in \mathbb{S}} x_i \geq v(\mathbb{S}), \forall \mathbb{S} \subset \mathbb{A} \right\}. \tag{14.3}$$

A method is required to obtain a specific solution (inside the core) in this N -person cooperative game. For this, we use the Shapley value because its computational complexity is small and it can provide relatively fair solutions compared with other methods. The Shapley value can be obtained as follows:

$$\phi_i(\mathbf{v}) = \sum_{\mathbb{S} \subset \mathbb{A}, i \in \mathbb{A}} \frac{(|\mathbb{S}|-1)!(n-|\mathbb{S}|)!}{n!} (\mathbf{v}(\mathbb{S}) - \mathbf{v}(\mathbb{S} - \{i\})) \quad (14.4)$$

where $|\mathbb{S}|$ indicates the number of elements in the set \mathbb{S} and $x_i = \phi_i(\mathbf{v})$.

14.5.2.2 Bandwidth Allocation

We utilize the bankruptcy game formulation for load balancing and bandwidth allocation for a new connection that is serviced by multiple networks simultaneously. Here, a mobile/connection is analogous to the bankrupted company and the requested bandwidth from the different networks is equivalent to the money that has to be offered to the different networks that are the players in this game. The n networks cooperate to offer bandwidth to a new connection so that the bandwidth requirement of the new connection is met and all of the networks are satisfied with the solution (i.e., stable). The total number of players here is 3 and the set of players is defined as $\mathbb{A} = \{wl, ce, wm\}$ for WLAN, cellular network, and WMAN, respectively.

When a new connection requests for bandwidth M , a central controller (e.g., eHealth agent) determines the amount of offered bandwidth to this connection from each network. This offered bandwidth is a function of the subscription class for that connection/mobile and the available bandwidth in each network. In particular, the offered bandwidth from each network i can be obtained as follows:

$$d_i = \begin{cases} \tilde{b}_{k,i}, & \tilde{b}_{k,i} < (B_i^{(a)})^r \\ (B_i^{(a)})^r + N \left(B_i^{(a)} - (B_i^{(a)})^r \right), & \tilde{b}_{k,i} \geq (B_i^{(a)})^r \end{cases} \quad (14.5)$$

where $\tilde{b}_{k,i}$ is the predefined offered bandwidth by network i to a new connection (or the corresponding mobile) with subscription class k , $B_i^{(a)}$ is the available bandwidth in network i , $b_k^{(req)}$ is the amount of requested bandwidth by a new connection in class k , N is a uniform random number between zero and one, and r is a control parameter that will be referred to as the bandwidth-shaping parameter (i.e., $0 < r \leq 1$). Then, the Shapley value gives the amount of allocated bandwidth in each network i , i.e., $x_i = \phi_i(\mathbf{v})$, $\forall i \in \mathbb{A}$.

14.5.2.3 Admission Control

When a mobile initiates a new connection, the information on the required bandwidth is sent to the central controller, which computes the offered bandwidth by each network (i.e., the Shapley value). The new connection is accepted if $\sum_{i \in \mathbb{A}_k} x_i \geq b_k^{(req)}$ and $x_i \in \mathbb{C}, \forall i \in \mathbb{A}$ (i.e., the Shapley value is in the core, namely, the solution is stable), and rejected otherwise. However, to prioritize eHealth connections over normal voice and data connections, a threshold-based admission control is used (e.g., threshold T_k for connection type k). In particular, if the summation of total allocated bandwidth to the ongoing connections plus the bandwidth requirement for the new connection when divided by the network capacity is less than this threshold, a new connection is accepted, and rejected otherwise.

14.5.3 Numerical Results

14.5.3.1 The Core and the Shapley Value

We vary the amount of requested bandwidth and the resulting allocation in each of the networks as shown in Figure 14.8. As expected, the amount of allocated bandwidth increases as the requested bandwidth increases. In Figure 14.8, the allocation can be divided into four intervals according to the amount of requested bandwidth (i.e., $[0, 150]$, $[150, 400]$, $[400, 600]$, and $[600, \infty]$). In the first interval, because the entire amount of requested bandwidth can be offered by each network, the bandwidth allocation in every network is equal. Therefore, the fair way to allocate bandwidth is to allocate an equal amount from each network. For the second interval,

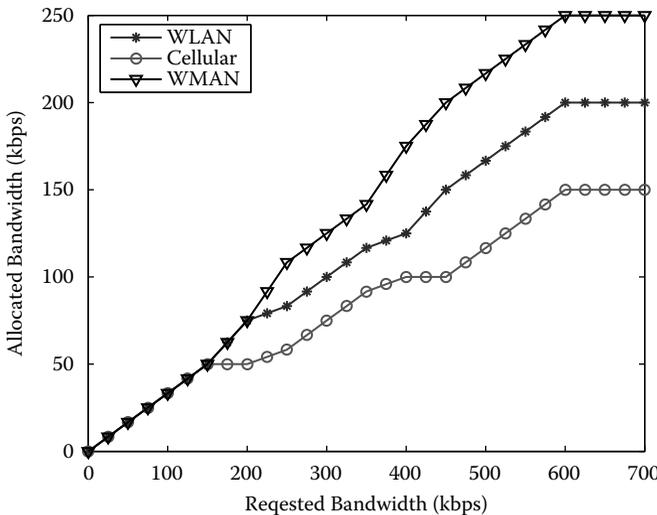


Figure 14.8 Example of bandwidth allocation.

because the requested bandwidth becomes larger than the offered bandwidth from one of the networks, the bandwidth allocation in each network becomes different. In the third interval, the differences among the allocated bandwidths in different networks become larger because the requested bandwidth is larger than the offered bandwidth in two of the networks. If the requested bandwidth becomes increasingly higher and becomes larger than the offered bandwidth in all of the networks, the allocated bandwidth becomes constant.

14.5.3.2 Performances of Bandwidth Allocation and Admission Control Algorithms

Next, for the proposed bandwidth allocation and admission control method, we evaluate connection blocking probability and network utilization under different connection arrival rates (Figure 14.9[a] and Figure 14.9[b]). Here, the connection blocking probability is defined as a probability that an incoming connection is blocked due to an admission control decision. This can be determined by the ratio of blocked connections to the total number of connections arriving at a particular service area. Network utilization determines the proportional of allocated bandwidth to the total available bandwidth. It is averaged over the total simulation time. The thresholds corresponding to connections for emergency prehospital, mobile health care, follow-up, and normal voice and data services are set to 1.0, 0.95, 0.9, 0.80, and 0.85, respectively, for all the networks. For comparison purposes, we also consider the traditional bandwidth allocation without load balancing. In this

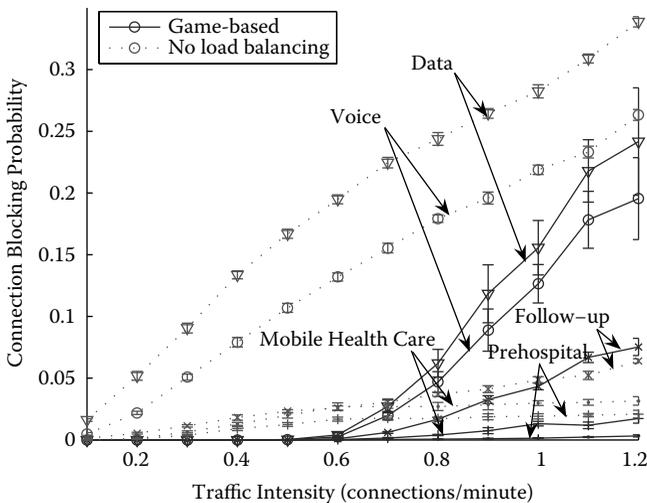


Figure 14.9 Variations in (a) connection blocking probability for each connection type and (b) network utilization.

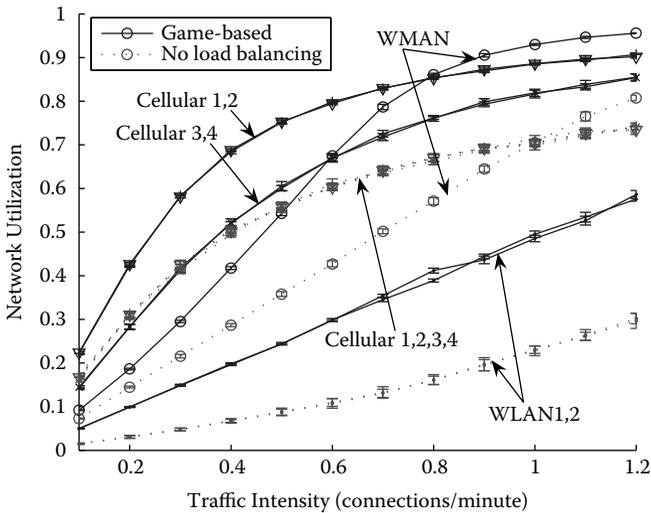


Figure 14.9 (Continued)

traditional scheme, a connection will be accepted by the network with the largest available bandwidth. In this case, an incoming connection will be blocked if none of the available networks have enough bandwidth.

From Figure 14.9, connection blocking probabilities increase as traffic intensity increases. However, because radio resource reservation is used, connection blocking probabilities for eHealth connections are much lower than those for normal data and voice connections. The connection blocking probabilities for the proposed bandwidth allocation and admission control scheme are compared with those of traditional bandwidth allocation in which load balancing is not used. The proposed game theory-based load balancing can achieve much lower connection blocking probability. Also, the network utilization is much higher.

14.6 Conclusions

We have reviewed the basic features and the research issues in 4G heterogeneous wireless access networks and described potential applications of such a network to provide different eHealth services. The limitations of the traditional wireless eHealth systems have been pointed out. A general architecture for a 4G wireless eHealth network has been presented and the related research issues have been outlined. To this end, for such a heterogeneous wireless access network we have presented a bandwidth allocation and admission control scheme. This scheme is based on cooperative game theory in which different types of networks make a coalition to offer bandwidth to the different types of eHealth connections. The

numerical results have shown that this bandwidth allocation scheme can improve the network performances in terms of resource utilization and connection blocking probability.

References

1. IEEE 802.16 Standard — Local and Metropolitan Area Networks — Part 16, IEEE Std 802.16a-2003.
2. D. Niyato, E. Hossain, and J. Diamond, IEEE 802.16/WiMAX-based broadband wireless access and its application for telemedicine/e-health services, *IEEE Wireless Communications*, 14, 1, 72–83, 2007.
3. W. Bolton, Y. Xiao, and M. Guizani, IEEE 802.20: Mobile broadband wireless access, *IEEE Wireless Communications*, 14, 1, 84–95, 2007.
4. C.H. Salvador et al., Airmed-cardio: A GSM and Internet services-based system for out-of-hospital follow-up of cardiac patients, *IEEE Transactions on Information Technology in Biomedicine*, 9, 1, 73–85, 2005.
5. B. Fong, A.C.M. Fong, and G.Y. Hong, Interoperability in a wireless home networking system for healthcare monitoring, in *Proceedings of IEEE ICCE'07*, 1–2, Jan. 2007.
6. Y.H. Lin, I.C. Jan, P.C.-I. Ko, Y.Y. Chen, J.M. Wong, and G.J. Jan, A wireless PDA-based physiological monitoring system for patient transport, *IEEE Transactions on Information Technology in Biomedicine*, 8, 4, 439–447, 2004.
7. M.F.A. Rasid and B. Woodward, Bluetooth telemedicine processor for multichannel biomedical signal transmission via mobile cellular networks, *IEEE Transactions on Information Technology in Biomedicine*, 9, 1, 35–43, 2005.
8. D. Niyato and E. Hossain, Call admission control for QoS provisioning in 4G wireless networks: Issues and approaches, *IEEE Network*, 19, 5, 5–11, 2005.
9. Q. Song and A. Jamalipour, Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques, *IEEE Wireless Communications*, 12, 3, 42–48, 2005.
10. J. McNair and F. Zhu, Vertical handoffs in fourth-generation multinet network environments, *IEEE Wireless Communications*, 11, 3, 8–15, 2004.
11. R. Bruno, M. Conti, and E. Gregori, Mesh networks: Commodity multihop ad hoc networks, *IEEE Communications Magazine*, 43, 3, 123–131, 2005.
12. X. Gao, G. Wu, and T. Miki, End-to-end QoS provisioning in mobile heterogeneous networks, *IEEE Wireless Communications*, 11, 3, 24–34, 2004.
13. B. Bernard, K. Paul, O. Juan, T. Betty, and S. Fontaine, Telemedicine: A solution to the follow-up of rural trauma patients, *Journal of the American College of Surgeons*, 192, 4, 447–452, 2001.
14. R. Fensli, E. Gunnarson, and O. Hejlesen, A wireless ECG system for continuous event recording and communication to a clinical alarm station, in *Proceedings of IEEE EMBC'04*, 1: 2208–2211, 2004.
15. R. Boussejot, U. Grieger, D. Kreisler, L. Schmitz, H. Koch, S. Beckmann, C. Bethge, H. von Nettelhorst, and K. Stangl, Telemetric ECG diagnosis follow-up, in *Proceedings of IEEE Computers in Cardiology*, 121–124, Sept. 2004.

16. M.F. Murad, R. Ahmad, S. Naeem, Q. Ali, A. Ehsan, T. Sohail, and A. Zafar, Follow-up of earthquake victims in a remote hospital using telemedicine, in *Proceedings of IEEE HEALTHCOM'06*, 228–231, Aug. 2006.
17. V.C. Protopappas, D.A. Baga, D.I. Fotiadis, A.C. Likas, A.A. Papachristos, and K.N. Malizos, An ultrasound wearable system for the monitoring and acceleration of fracture healing in long bones, *IEEE Transactions on Biomedical Engineering*, 52, 9, pp. 1597–1608, 2005.
18. S. Guillen, M.T. Arredondo, V. Traver, J.M. Garcia, and C. Fernandez, Multimedia telehomecare system using standard TV set, *IEEE Transactions on Information Technology in Biomedicine*, 49, 12, 1431–1437, 2002.
19. M. Braecklein, I. Tchoudovski, C. Moor, K. Egorouchkina, L.P. Pang, and A. Bolz, Wireless telecardiological monitoring system for the homecare area, in *Proceedings of IEEE EMBS'05*, 3793–3795, 2005.
20. W.-C. Kao, W.-H. Chen, C.-K. Yu, C.-M. Hong, and S.-Y. Lin, Portable real-time homecare system design with digital camera platform, *IEEE Transactions on Consumer Electronics*, 51, 4, 1035–1041, 2005.
21. Y. Chu and A. Ganz, A mobile teletrauma system using 3G networks, *IEEE Transactions on Information Technology in Biomedicine*, 8, 4, 456–462, 2004.
22. S. Pavlopoulos, E. Kyriacou, A. Berler, S. Dembeyiotis, and D. Koutsouris, A novel emergency telemedicine system based on wireless communication technology-AMBULANCE, *IEEE Transactions on Information Technology in Biomedicine*, 2, 4, 261–267, 1998.
23. M. Takizawa, S. Sone, K. Hanamura, and K. Asakura, Telemedicine system using computed tomography van of high-speed telecommunication vehicle, *IEEE Transactions on Information Technology in Biomedicine*, 5, 1, 2–9, 2001.
24. A.F. Graves, B. Wallace, S. Periyalwar, and C. Riccardi, Clinical grade — A foundation for healthcare communications networks, in *Proceedings of the International Workshop on Design of Reliable Communication Networks (DRCN'05)*, Oct. 2005.
25. B. Hayes-Roth and J.E. Larsson, A domain-specific software architecture for a class of intelligent patient monitoring agents, *Journal of Theoretical and Experimental Artificial Intelligence*, 8, 2, 149–171, 1996.
26. V.D. Mea, Agents acting and moving in healthcare scenario — a paradigm for telemedical collaboration, *IEEE Transactions on Information Technology in Biomedicine*, 5, 1, 10–13, 2001.
27. M. Li and R.S.M. Istepanian, 3G network oriented mobile agents for intelligent diabetes management: A conceptual model, in *Proceedings of IEEE EMBS'03*, 31–34, Apr. 2003.
28. V.G. Koutkias, I. Chouvarda, and N. Maglaveras, A multiagent system enhancing home-care health services for chronic disease management, *IEEE Transactions on Information Technology in Biomedicine*, 9, 4, 528–537, Dec. 2005.
29. B.-M. Han, S.-J. Song, K.M. Lee, K.-S. Jang, and D.-R. Shin, Multi-agent system based efficient healthcare service, in *Proceedings of IEEE ICACT'06*, 1, Feb. 2006.
30. D.L. Hudson and M.E. Cohen, Intelligent agent model for remote support of rural healthcare for the elderly, in *Proceedings of IEEE EMBS'06*, 6332–6335, Aug. 2006.
31. S. Haykin, Cognitive radio: Brain-empowered wireless communications, *IEEE Journal on Selected Areas in Communications*, 23, 2, 201–220, 2005.

32. D. Cabric, A. Tkachenko, and R.W. Brodersen, Spectrum sensing measurements of pilot, energy, and collaborative detection, in *Proceedings of IEEE MILCOM'06*, 1–7, Oct. 2006.
33. G. Ganesan and Y. Li, Cooperative spectrum sensing in cognitive radio networks, in *Proceedings of IEEE DySPAN'05*, 137–143, Nov. 2005.
34. M.P. Wylie-Green, Dynamic spectrum sensing by multiband OFDM radio for interference mitigation, in *Proceedings of IEEE DySPAN'05*, 619–625, Nov. 2005.
35. A. Ghahmazi, and E.S. Sousa, Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing, *IEEE Communications Letters*, 11, 1, 34–36, 2007.
36. H. Yaiche, R.R. Mazumdar, and C. Rosenberg, A game theoretic framework for bandwidth allocation and pricing in broadband networks, *IEEE/ACM Transactions on Networking*, 8, 5, 667–678, 2000.
37. H. Lin, M. Chatterjee, S.K. Das, and K. Basu, ARC: An integrated admission and rate control framework for competitive wireless CDMA data networks using noncooperative games, *IEEE Transactions on Mobile Computing*, 4, 3, 243–258, 2005.
38. Z. Xia, W. Hao, I.-L. Yen, and P. Li, A distributed admission control model for QoS assurance in large-scale media delivery systems, *IEEE Transactions on Parallel and Distributed Systems*, 16, 12, 1143–1153, 2005.
39. T. Alpcan, T. Basar, and S. Dey, A power control game based on outage probabilities for multicell wireless data networks, *IEEE Transactions on Wireless Communications*, 5, 4, 890–899, 2006.
40. J. Musacchio and J. Walrand, WiFi access point pricing as a dynamic game, *IEEE/ACM Transactions on Networking*, 14, 2, 289–301, 2006.
41. J. Choi, J. Yoo, C. Kim, and S. Choi, EBA: An enhancement of the IEEE 802.11 DCF via distributed reservation, *IEEE Transactions on Mobile Computing*, 4, 4, 378–390, 2005.
42. L. Xu, X. Shen, and J. W. Mark, Fair resource allocation with guaranteed statistical QoS for multimedia traffic in wideband CDMA cellular network, *IEEE Transactions on Mobile Computing*, 4, 2, 166–177, 2005.
43. B. O'Neill, A problem of rights arbitration from the Talmud, *Mathematical Social Sciences* 2, 345–371, 1982.

Chapter 15

3G/WLAN Cross-Layer Design for Ultrasound Video Transmission in a Robotic Tele-Ultrasonography System

Maria G. Martini, Robert S.H. Istepanian, Matteo Mazzotti, and Nada Philip

CONTENTS

- 15.1 Introduction298
- 15.2 Mobile Robotic Tele-Echography and the OTELO System 300
 - 15.2.1 3G/WLAN Functional Modalities of the OTELO System301
 - 15.2.2 802.11e WLAN Link Considerations in Expert (Hospital) Station..... 304
- 15.3 Robust Multi-Layer Controller Structure for Enhanced Medical Video Streaming..... 304
- 15.4 Results and Discussion 308

15.5 Conclusions	315
Acknowledgments	315
References	316

The rapid evolution of wireless technology is allowing mobile health care technologies to become a reality. These include the transmission of high-quality medical video sequences over unreliable wireless links. However, one of the key challenges in this area consists in the contrasting requirements of almost lossless compression and low available bandwidth, especially in ultrasound and radiology telemedical consultation services. Joint source and channel coding and cross-layer design approaches have proven suitable for wireless video transmission over IP networks, but such approaches have not been specifically considered for wireless medical video applications. In this chapter we present a cross-layer design approach for enhanced wireless ultrasonography video streaming from a remote robotic scanning system in a combined 3G/WLAN environment. The chapter outlines the robotic ultrasonography system, then describes the joint source and channel coding methodology designed specifically for enhanced-quality ultrasound streaming. Some preliminary test results are presented and future work in this area is also addressed.

15.1 Introduction

Wireless medical imaging and teleconsultation represent emerging areas within the mobile health care domain. Current and emerging developments in wireless communications integrated with developments in pervasive and wearable technologies will have a radical impact on future health care delivery systems, especially in wireless and mobile ultrasonography teleconsultation. M-health can be defined as mobile computing, medical sensors, and communications technologies for health care.^{1,2} This emerging concept represents the evolution of e-health systems from traditional desktop telemedicine platforms to wireless and mobile configurations.

The main wireless technologies suitable for enhanced wireless telemedicine systems are 3G (W-CDMA, CDMA 2000, TD-CDMA) and WLAN, not forgetting emerging technologies such as WiMAX, personal area networks, and ad hoc/sensor networks (exploiting short-range communications technologies such as Bluetooth and Zigbee). Fourth generation (4G) systems will integrate existing technologies in the common framework of IP-based systems.

In this chapter, we present an advanced mobile health care application represented by a mobile robotic tele-echography system that requires a demanding medical data and videostreaming traffic for high-quality diagnosis in a heterogeneous network topology that combines 3G and WLAN environments.

We describe our approach based on a network-aware joint source channel coding and decoding (JSCC/D)³ method for bandwidth demanding medical video streaming systems (referred to as “medical JSCC/D” or “m-JSCC/D”). The detailed JSCC approach is partly described elsewhere.⁴

Medical video streaming represents one of the highest demanding telemedical applications in wireless environments. Medical video compression techniques for telemedical applications have requirements of high fidelity, in order to avoid the loss of information that could help an acceptable diagnosis quality. In order to keep diagnostic accuracy, lossless compression techniques are often considered when medical video sequences are involved. In any case, when transmission is over band-limited, error-prone channels, a compromise solution should be considered between compression fidelity and protection and resilience to channel errors and packet loss, to provide a medically acceptable diagnosis quality. From the medical imaging perspective, it has been observed that when lossy compression is limited to ratios from 1:5 to 1:29, compression can be achieved with no loss in diagnostic accuracy.⁵ Furthermore, even if the final diagnosis should be done using an image that has been reversibly compressed, irreversible compression still plays a critical role when quick access to data stored in a remote location is needed. For these reasons, lossy compression techniques, such as H.264, have been considered for medical images and in particular ultrasound medical video compression.⁶ These critical requirements demand a careful and robust balance between diagnostic requirements and fragility issues of current wireless networks.

In this work, we particularly focus on compression based on video coding standards suitable for an ultrasound telemedical robotic system,⁷ such as MPEG-4⁸ and H.264,⁹ where the available error resilience tools may help in keeping an acceptable quality after transmission. The system we propose represents a new end-to-end optimization of an advanced m-health system over an integrated IP 3G/WLAN environment.

The approach presented in this chapter is based on the joint source and channel coding paradigm and provides an evolution from the communications theory perspective, apparently in contrast with the well-known “separation theorem” derived from Shannon’s theory.¹⁰ It has been shown in fact that separation is not optimal for wireless audio and video transmission, particularly when transmitting data with real-time constraints or operating on sources whose encoded data bit error sensitivity varies significantly.

Recently, JSCC/D techniques that include a co-ordination between source and channel encoders were investigated,¹¹ e.g., for transmission of audio data,¹² images,¹³ and video.¹⁴ In some of these works the transmission is adapted to the source characteristics (unequal error protection [UEP]), either at channel coding level or through source adaptive modulation (see, e.g., Hagenauer, Seshadri, and Sundberg,¹² and Dardari et al.¹⁵). JSCC/D techniques may also require the use of rate/distortion curves or models of the source in order to perform the optimal compromise between source compression and channel protection.¹³

IP-based transmission systems require a further analysis step, because we have to consider that in order to allow joint design of the source encoder, at application layer, and of the channel coding strategy, at physical layer, information on the two blocks should flow in the protocol stack to allow the source encoder to profit from information on the channel conditions and the channel encoder to exploit information from the source. Information about the network condition can be exploited as well for system optimization, e.g., by compressing the source more when the network is overloaded. This approach, developed in recent years, is known as “cross-layer design”^{16,17} and consists of jointly designing the different layers of the protocol stack, by exploiting at each layer global information from the system.

In the remainder of this chapter, we present the cross-layer design framework together with the application of this approach on a wireless robotic tele-ultrasonography system for acceptable-quality medical video streaming requirements.

Some preliminary performance analysis results and discussion for future work in this area are also provided.

15.2 Mobile Robotic Tele-Echography and the OTELO System

Tele-ultrasound systems for remote diagnosis have been proposed in recent years^{18–26} given the need to allow tele-consultation when access of the medical specialist to the sonographer is not possible. Ultrasonography is an examination strongly relying on specialized skills. The clinician performs the diagnosis both from static measurements performed on images and from dynamical behavior analysis of the organs. Good hand–eye coordination is thus required and this aspect has to be considered in tele-consultation. The possibility to allow the expert to drive the probe and perform real-time evaluation of the relevant video is thus of high importance.

The LOGINAT project was developed in France since 1993 for inter-hospital perinatal care. Due to the limited performance of communication networks at that time, only static images and related documents were transmitted in real-time. Similarly, the TeleInViVo European project²¹ developed a portable station to allow echographic exams in isolated regions. A digital tele-ultrasound system for real-time JPEG image transmission over wireline (leased 1.5 Mbps lines) is described in Sublett and Weaver.¹⁸ In all these cases, the expert had to perform the diagnosis on a purely visual basis.

First examples of a robotic tele-ultrasonography system were represented by the SYRTECH system²⁵ and the MIDSTEP European project.²⁰ HIPPOCRATE, a robot for remote ultrasonography, is described in Masuda et al.²³ A further example is Vilchis,²² where a tele-operated robot for remote ultrasonography (TER) is described, with MJPEG/MPEG video transmission over wireline ISDN/LAN systems. Such systems provide a better performance with respect to the previous group, allowing the clinician to perform a more reliable diagnosis based not only on

single images and video, but also on their relation with the probe position when the image is captured. All the described systems only address wireline transmission.

In recent years, an advanced medical robotic system named OTELO (mobile tele-echography using an ultra-light robot) was developed, in the framework of a European IST funded project, consisting of a fully integrated end-to-end mobile tele-echography system for population groups that are not served locally, either temporarily or permanently, by medical ultrasound experts.^{7,24} It comprises a fully portable tele-operated robot allowing a specialist sonographer to perform a real-time robotized tele-echography to remote patients. Figure 15.1 shows the main operational blocks of the system.

This tele-echography system is composed of the following:

1. An expert site, where the medical expert interacts with a dedicated patented pseudo-haptic fictive probe instrumented to control the positioning of the remote robot and emulating an ultrasound probe that medical experts are used to handling, thus providing better ergonomics.
2. The communication media. We developed communication software based upon IP protocol to adapt to different communication (wired and wireless links).
3. A patient site made up of the six degrees of freedom (DoF) lightweight robotic system and its control unit. Further details on this system are described in Istepanian, Jovanov, and Zhang² and Garawi, Istepanian, and Abu-Rgheff.⁷

15.2.1 3G/WLAN Functional Modalities of the OTELO System

OTELO can be considered as a bandwidth-demanding data traffic advanced m-health system, with challenging classes of QoS requirements, as several medical ultrasound images, robotic, and other data have to be transmitted simultaneously. Such requirements are summarized in Table 15.1.

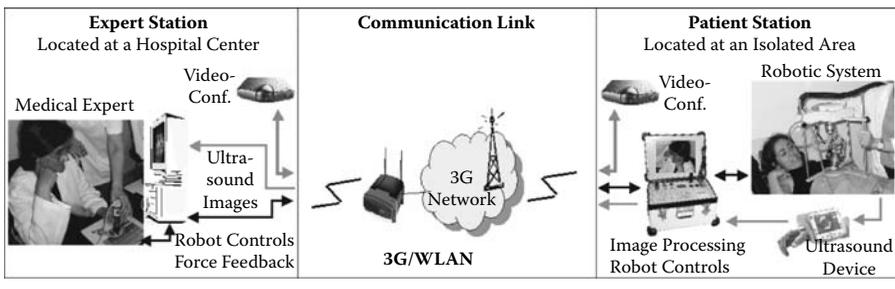


Figure 15.1 The OTELO mobile robotic tele-ultrasonography system.

Table 15.1 OTELO Medical Data Requirements and Corresponding Data Rates

	<i>Ultrasound Video</i>	<i>Ultrasound Still Images</i>	<i>Ambient Video Stream</i>	<i>Voice</i>	<i>Robot Control Data</i>
Flow direction	Simplex: patient to expert	Simplex: patient to expert	Duplex	Duplex	Duplex
Transport protocol	RTP/UDP/IP	TCP/IP	RTP/UDP/IP	RTP/UDP/IP	UDP/IP
Speed requirement	Real-time	Non real-time	Real-time	Real-time	Real-time
Payload data rate requirement	15 frames/210 kbps	1 frame/10 sec uplink	15 to 1 frame/sec symmetrically	16 kbps symmetrically	0.3 kbps symmetrically

Figure 15.2 shows the general 3G/WLAN connectivity of the OTELO system and the interface requirements. In this scenario, we assume that OTELO’s expert station is connected to the OTELO system via the specialist hospital WLAN network.

The detailed medical and non-medical OTELO data traffic are shown in Table 15.1.⁷ As the ultrasound images are mostly transferred from the robot probe to the OTELO expert station, the air interface, *Uu*, between the OTELO patient station and the radio network controller (RNC) bearer, is characterized by asymmetric traffic load. The still ultrasound images, stream ultrasound images, ambient video, sound, and robot control data are sent over the uplink channel, while only robot control, ambient video, and sound need to be downloaded to the patient side (i.e., expert station uploading).

From Table 15.1, it can be seen that for the OTELO system the most bandwidth-demanding traffic is the medical ultrasound (US) data. For this reason, the focus in the following sections is on the transmission of U.S. data in wireless environments. According to the communication link limitations, various scenarios can be identified with respect to the data traffic that should be sent simultaneously so as to enable performance of the medical examination, as explained briefly in the following operational steps of this medical robotic system:

1. When the expert is searching for a specific organ (liver, kidney, etc.), high-quality images may not be required. Simple compression methods or lossy techniques can be applied. The lowest data rate acceptable to medical experts is about 210 kbps with a frame update of 15 frames/sec.²⁴
2. When the organ of interest is found and small displacements of the robot are applied, it may be necessary to consider lossless compression techniques that

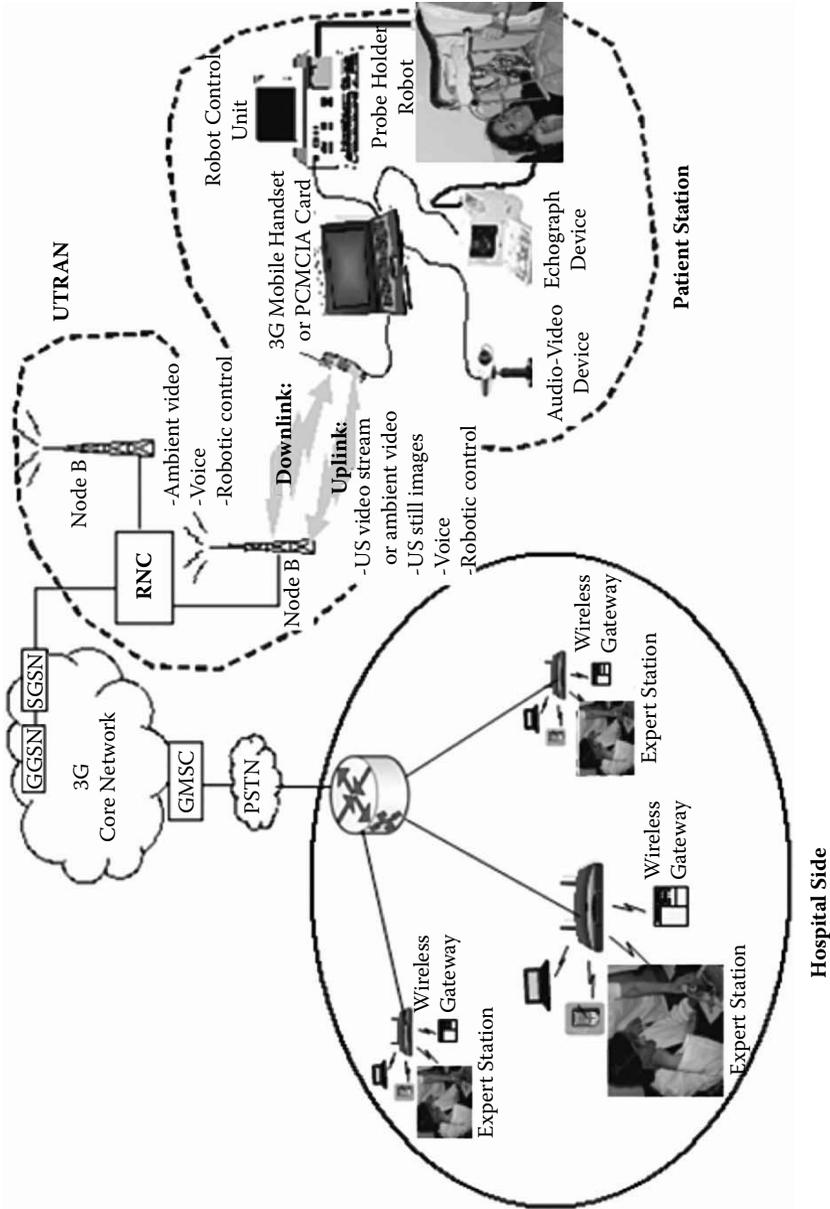


Figure 15.2 3G/WLAN wireless connectivity of OTELO system.

would bring higher image quality to the expert. This lossless compression can be applied on the whole image or on a region of interest (ROI). From the medical perspective and in order to provide real-time virtual interactivity between the remote consultant and the manipulated robot, the best round-trip delay from the expert station between the robot commanded position and the received corresponding image should not exceed 300 ms.⁷

3. For clinical validation purposes, there is the need to have a multi-site specialist wireless connectivity in the hospital and to provide an assured medical diagnosis of the received ultrasound images. Hence in this study we assume an additional multi-specialist WLAN connectivity system to provide such wireless teleconsultation service.

15.2.2 802.11e WLAN Link Considerations in Expert (Hospital) Station

In this work we consider IEEE 802.11e WLAN connectivity, as it provides enhanced QoS features and multimedia support compared to the existing IEEE 802.11b and IEEE 802.11a wireless standards, while maintaining full backward compatibility with them. In the 802.11a standard, an orthogonal frequency division multiplexing (OFDM) encoding scheme is used rather than FHSS or DSSS. 802.11b, often called Wi-Fi, considers complementary code keying (CCK) as the modulation method, which allows higher data speeds and is less susceptible to multipath-propagation interference.

Although WLAN connectivity allows the possibility to use higher bandwidth, we need to consider that data transmitted in the hospital WLAN is possibly received from the UMTS link. In this case, the UMTS link represents the bottleneck because the source bit rate in the WLAN section is limited to the one received from the UMTS link. In any case, more error protection can be provided to the medical US data in the WLAN section given its higher available bandwidth.

15.3 Robust Multi-Layer Controller Structure for Enhanced Medical Video Streaming

In this section we describe the proposed cross-layer design framework for OTELO compressed ultrasound video transmission over 3G/WLAN environments (ROAM).

Figure 15.3, in which a single wireless hop is depicted, illustrates the overall system architecture proposed. The figure shows the transmitter side (patient side) in the upper part of the figure and the receiver side (expert side) in the lower part, including the signalization used for transmitting the medical JSCC/D control

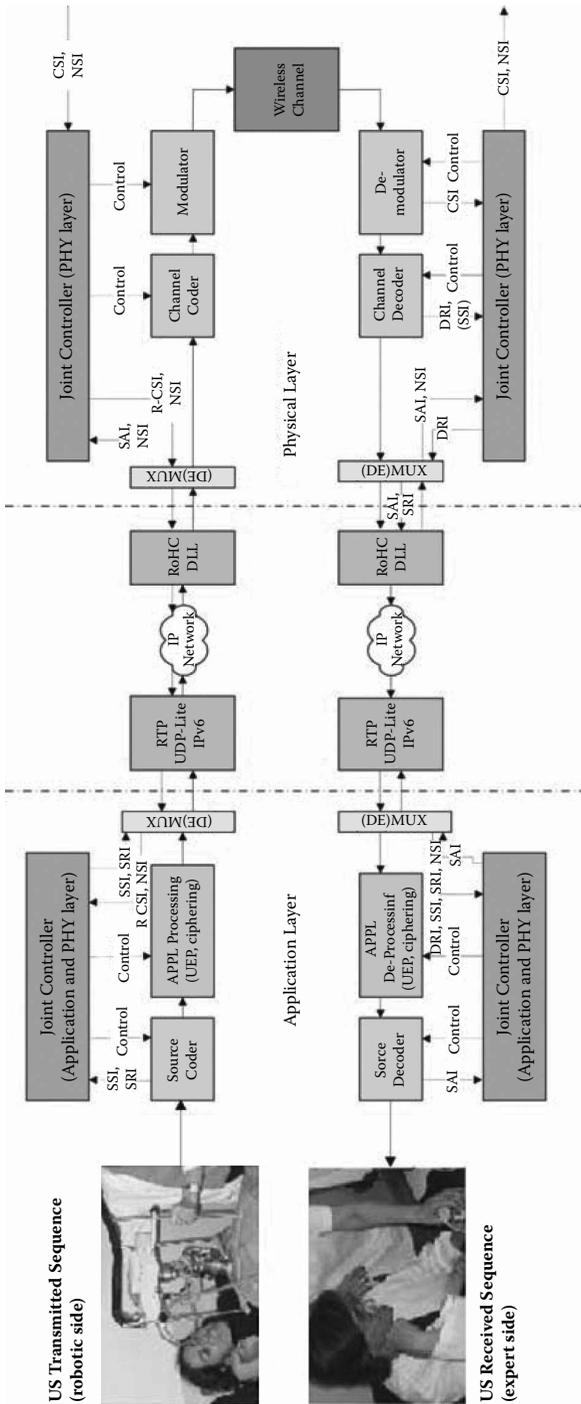


Figure 15.3 ROAM architecture for the OTELO system.

information in the system. We focus, in fact, on the transmission of ultrasound video from patient to specialist. In addition to the traditional tasks performed at the application level (source encoding, application processing such as ciphering), at the network level, including RTP/UDP (possibly UDP-Lite)/IPv6 packetization, impact of IPv6 wired network, and robust header compression (RoHC), medium access (including enhanced mechanisms for 802.11e), and radio access (channel encoding, interleaving, modulation), the architecture includes two controller units at the physical and application layers. Those controllers, depicted in Figure 15.4, are introduced for supervising the different (de)coders, (de)modulation, and (de)compression modules and to adapt said module parameters to changing conditions, through the sharing of information about the source, network, and channel conditions and user requirements.

Table 15.2 shows the different I/O information of each controller.

Clearly, when considering real-time diagnostic systems, this control information needs to be transferred through the network and system layers, in a timely and bandwidth-efficient manner. The impact of the network and protocol layers is often neglected when studying joint source and channel coding and only minimal effort is made in finding solutions for providing efficient inter-layer signaling mechanisms for JSCC/D. The authors have identified different mechanisms that could allow information exchange transparently for the network layers (see, e.g., Martini and Chiani²⁷ and Martini et al.²⁸).

Finally, it should be noted that additional information is requested by the system for the set-up phase, where information on available options (e.g., available channel encoders and available channel coding rates, available modulators, etc.) and a-priori information on the transmitted ultrasound video (e.g., statistical characterization of video sequence) are exchanged, the session is negotiated, and default parameters are set (e.g., authentication key, module's default settings).

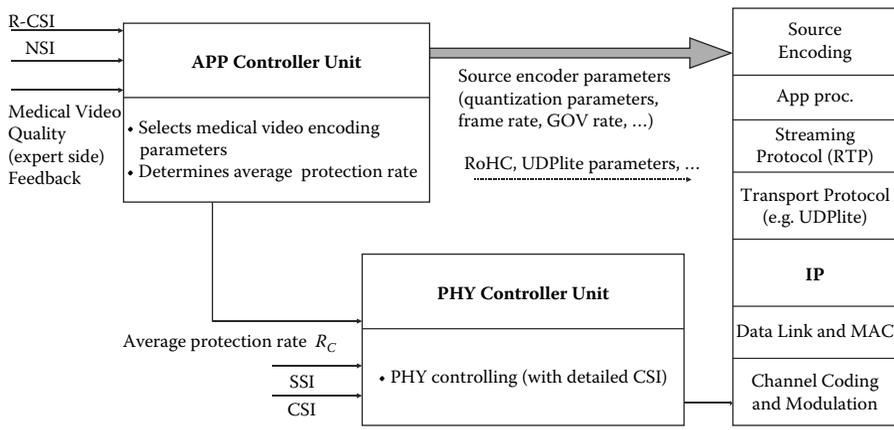


Figure 15.4 The ROAM cross-layer controller structure (patient side).

Table 15.2 I/O Parameters of System Controllers in the Described Scenario

	<i>APP Controller</i>	<i>PHY Controller</i>
Input	R-CSI (reduced channel state information)	CSI (channel state information)
	NSI (network status information)	SSI (source significance information)
	US video quality in the previous time step	Average channel code rate R_c (from the APP controller)
Output	Source encoder quantization parameters for I and P frames (qI, qP)	Code rates $R_{c,i}$ of each sensitivity class in the US video stream
	Source encoder frame rate	Adaptive bit-loading parameters for multicarrier
	Source encoder intra frames refresh rate (GOP length)	
	Average channel code rate R_c (derived from the previous parameters)	

As shown in Figure 15.4 and in Table 15.2, the application (APP) controller collects information from the network (NSI, in terms of packet loss, delay, and jitter) and has availability of reduced channel state information (only channel state information averaged over longer time steps is available at this layer) and the quality metric of the previously decoded frame (or group of frames) of ultrasound video. According to this cross-layer information, it produces controls for the source encoder block (e.g., quantization parameters, frame rate, error resilience tools to activate) and possibly the network. The controller has been modeled as a finite state machine: typically, a low ultrasonography video-quality value associated to a negative trend will cause a transition to a state characterized by a higher robustness, i.e., with a higher compression allowing invoking stronger error-resilient tools and lower-rate/higher-protection channel coding. When there is network congestion, the controller immediately sets the state to the one characterized by the lowest source bit rate (corresponding, for example, to the minimum requirements for OTELO medical video), in order to reduce as much as possible the amount of data that have to flow through the IPv6 network.

Given the bit rate associated to the chosen state, the code rate R_c available for signal protection is evaluated and such information is provided directly to the PHY controller, whose task is to provide controls to the physical layer blocks, i.e., the channel encoder, modulator and interleaver, by deciding on the channel coding rate for each different sensitivity partition of the source ultrasound video, with the goal

of minimizing the total distortion with the R_c constraint (for the general procedure, see Martini and Chiani¹⁴).

The source encoded medical video bit stream may in fact be separated in partitions or layers with different sensitivity to channel errors. A different protection can thus be used for the different partitions, when allocating the average channel coding rate available. As an example, in the case of MPEG-4 video the bit stream can be separated in packets and each packet can be separated in a header and two data partitions, with different error sensitivity. Packets from I (intra) frames can be separated in a first class related to DC DCT coefficients and a second class related to AC DCT coefficients, whereas P (predicted) frames packets can be separated in two partitions relevant to motion and texture data, respectively. This different sensitivity can be exploited to perform unequal error protection, either at the application layer or the physical layer. The video stream sensitivity can be modeled similar to Martini and Chiani¹⁴ in order to simplify the UEP policy. Similarly, in the case of H.264-based compression, the data partitioning tool or the granularity offered by scalable video coding (SVC) may be exploited.

Unequal protection based on ROI can also be considered, by exploiting the possibility offered by the MPEG-4 standard to separate any video sequence in video objects that can be differently managed. In this view, the identification of regions of interest allows dedicating a higher protection to the region of interest, allowing an increase in diagnostic accuracy, for a fixed available bandwidth.

Furthermore, the PHY controller subunit sets the parameters for bit-loading in multicarrier modulation, interleaver characteristics and performs a trade-off with receiver complexity. Again, the metric chosen for representing distortion should be representative of diagnostic accuracy.

15.4 Results and Discussion

In order to demonstrate the feasibility of the proposed framework and to evaluate the performance achievable, the proposed cross-layer controlled architecture has been implemented with its different sub-blocks, namely: application layer controller; source encoder/decoder (three possible codecs: MPEG-4, H.264/AVC and scalable video coding in H.264/AVC Annex G), where soft-input source decoding is also allowed for H.264/AVC; cipher/decipher unit; Real-Time Transport Protocol (RTP) header insertion/removal; transport protocol header (e.g., UDP-Lite, UDP, or datagram congestion control protocol [DCCP]) insertion/removal; IPv6 header insertion/removal; IPv6 mobility modeling; IPv6 network simulation; RoHC; DLL header insertion/removal; radio link, including physical layer controller, channel encoder/decoder (convolutional, rate compatible punctured convolutional [RCPC], low density parity check [LDPC] codes with soft and iterative decoding allowed), interleaver, modulator (also OFDM, TCM, TTCM, STTC; soft and iterative demodulation allowed) and channel (e.g.,

additive white Gaussian noise [AWGN], Rayleigh fading, shadowing, frequency selective channels).

The proposed structure is implemented in a simulated laboratory environment with images and video stream acquired from the real OTELO system. The test ultrasound video sequences acquired by the robotic sonographer are thus provided to the source codec, performing source (MPEG-4/H.264) encoding (according to the parameters suggested by the APP controller) by every controller time-step. The encoded bit stream is then processed by the lower layers and finally transmitted over the wireless channel model. The parameters of upper layers, down to the network, are determined by the application layer controller unit by every APP controller time-step. The parameters of lower layers, in particular of the physical layer, are determined run-time by the PHY controller unit with the relevant time-step (lower or equal to the one of the APP controller). In particular, the application layer controller unit performs source bit rate adaptation and the physical layer one provides UEP, according to the average bit rate suggested by the APP controller, and drives adaptive bit-loading for multicarrier modulation. Default parameter setting is considered in the initialization phase.

The 802.11e WLAN support was added at the radio link level, with a total bit rate of 12 Mbps. The ultrasonography video stream is coded according to the MPEG-4 standard and assumed to be multiplexed with other real-time transmissions, so that it occupies only an average portion of the available bandwidth corresponding to a coded bit rate of 650 kbps. The CIF image resolution has been selected. The MoMuSys MPEG-4 reference video codec is considered, with some modifications in the decoder to improve bit error resilience. The modified decoder is used both in the adapted and in the non-adapted system. A test echocardiography video sequence is considered in the example shown.

RoHC is also applied, in order to compress the transport and network headers by transmitting only non-redundant information.

Channel codes a regular repeat-accumulate (IRA) LDPC codes with a “mother” code rate of (3500, 10500), properly punctured and shortened in order to obtain different code rates. The resulting codewords are always 4200 bits long. The code rate is $2/3$ for the non-adapted system (EEP); in the adapted case the code rate can change according to source significance information (SSI) in order to perform UEP. The average coded bit rate is the same in both cases considered.

In the first case, the modulation is a “classical” OFDM with 48 carriers for data transmission and a frame duration of 4μ sec; margin adaptive bit-loading techniques managed by the PHY JSCC controller are considered in the adapted system.

The channel is obtained according to the ETSI channel A model, representing the conditions of a typical office (hospital) environment. It takes into account also a log-normal flat fading component with channel coherence time of 5 sec, to consider the fading effects due to large obstacles. A median signal-to-noise ratio of $E_b/N_0 = 13.2$ dB has been considered.

For the scenario considered, five different states have been chosen for the APP JSCC controller, each characterized by different sets of values for the above-mentioned parameters. State 1 corresponds to the lowest source data rate (lowest video quality) and highest robustness, whereas state 5 corresponds to the highest source data rate (highest video quality) and lowest robustness. Thus, increasing the state number means increasing the robustness transmission at a cost of loss in the error-free received video quality.

Table 15.3 shows the relevant frame rates and corresponding states with resulting data rates used in the simulation set-up.

The source bit rate after MPEG-4 compression depends on the APP controller status and ranges from 210 (state 1) to 384 kbps (state 5), taking into account also the overhead due to the various network headers, and it is thus in good accordance with the OTELO requirements, as shown in Table 15.1 and in Section 15.2.2. Note that in some cases, to keep an acceptable video quality in very deep fades or network congestion, the most robust states have to consider lower frame rates than those required by the OTELO system.

State 4 is the reference one, i.e., the one considered in the non-adapted case, whereas the controller switches among the states in Table 15.3 in the adapted case. The maximum bit rate over the channel is 450 kbps.

The simulation set-up is summarized in Table 15.4.

Figure 15.5 shows the comparative results of the system in terms of PSNR and structural similarity metric (SSIM),²⁹ a well known video quality metric used for better subjective video quality assessment. The quality curves reported in the graph have been obtained through the average of four distinct simulations, run with different noise speeds. Quality values averaged over 1 sec are reported in the results. Moreover, the quality values have been normalized with respect to the maximum value achieved in order to allow the comparison of different metrics in the same figure. An average gain of 4.4 dB in terms of PSNR is provided by the system, allowing the performance of the diagnosis with much higher accuracy than in the non-adapted case, as visual results confirm. It is clear from these results that even

Table 15.3 Sets of Source Encoder Parameter Values Used by the APP ROAM Controller (Medical Video Encoded According to MPEG-4 Standard)

<i>State</i>	<i>(q1, qP)</i>	<i>Frame Rate</i>	<i>Group of Pictures (GOP) Length</i>	<i>Resulting Bit Rate of Ultrasound Video (kbps)</i>
1	14,16	7.5	8	210
2	11,14	7.5	8	241
3	11,12	7.5	8	269
4	11,13	15	15	363
5	10,12	15	15	384

Table 15.4 Summary of the ROAM Functional Parameters Used for OTELO System Test

US Test Video Sequence	
US Video sequence	Echo-cardiogram video sequence
Video format	CIF (352 × 288)
Frame rate	30 frames/sec
Duration	4.5 sec (then looped)
Joint Control	
Overall coded bit rate	650 kbps (after channel coding)
Control update step	1 sec
Source Coding	
Encoded sequence frame rate	7.5/15 frames/sec
Intra frame refreshment period	8/15 frames
Source bit rate	Reference mode: average rate of 270 kbps Adapted mode: variable rate from 155 to 295 kbps
MPEG-4 packet average size	180 bytes
Encoding mode	Data partitioned mode
Extra header maximal size (for rate estimation)	96 bits (0 in classical mode)
Packetization and Transport	
Real-time management	RTP standard format
Transport protocol	UDP-Lite
Transport checksum coverage	UDP-Lite+RTP
PHY Layer	
Channel encoder	IRA-LDPC codes of "mother" rate (3500,10500)
Codewords length	4200
Supported code rates	1/3, 1/2, 2/3, 3/4, 5/6
Modulation	OFDM with 48 subcarriers for data, adaptive bit-loading in the adapted case
Frame duration	4 μsec
Number of RX/TX antennas	1/1
Maximum coded bit rate	650 kbps
(Continued)	

Table 15.4 Summary of the ROAM Functional Parameters Used for OTELO System Test (Continued)

Radio channel	Frequency selective (according to ETSI channel A model), with channel sample time of 1 ms (= block fading duration)
Slow fading	Uncorrelated log-normal distribution, $\sigma_{dB} = 4$ dB
Slow fading coherence time	5 sec (= block fading duration)
Median Eb/N0	13.2 dB

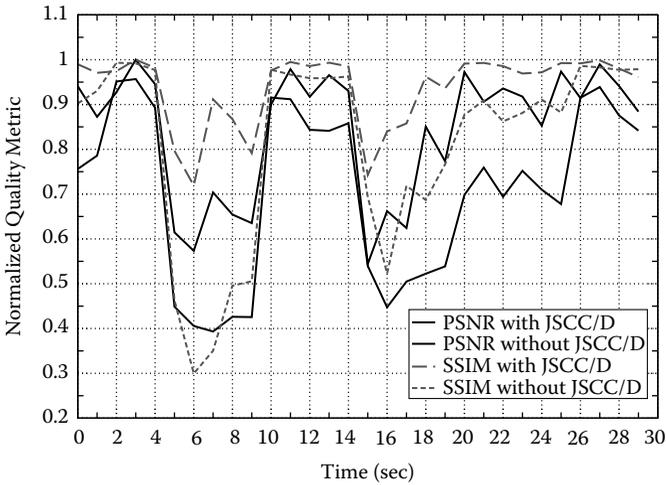


Figure 15.5 Comparative performance of the ROAM and reference architecture. Normalized PSNR and SSIM versus time.

in deep channel fade conditions the medical video quality in the ROAM adapted system is kept within acceptable levels. In particular if such fades happen in the first part of the ultrasonography, where the medical doctor is searching for the specific organ, this allows a reduction of the search time, avoiding the time where the quality of the communication is not acceptable.

Results in terms of complementary cumulative distribution function of video quality expressed as SSIM are reported in Figure 15.6. We may observe, for example, that the probability to have a video quality above 0.7 in terms of SSIM, which can be considered as a threshold for acceptable quality, is 0.95 in the cross-layer adapted case, whereas it is only 0.71 in the reference case. The relevant PSNR results are reported in Table 15.5.

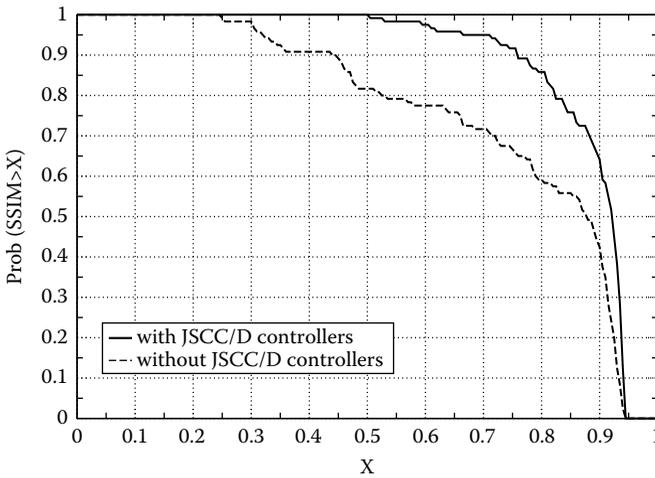


Figure 15.6 Comparative performance of the ROAM and reference architecture: Complementary cumulative distribution function of ultrasound video quality in terms of SSIM.

Table 15.5 Comparative PSNR Results in Subsequent 30 Sec of the Tested Ultrasound Video Sequence

ROAM	30.4	28.2	30	32.3	30.6	19.9	18.6	22.8	21.2	20.6
Reference	24.4	25.4	30.8	30.9	28.9	14.5	13.1	12.7	13.7	13.7
ROAM	29.1	31.7	29.7	31.2	30.1	17.6	21.4	20.2	27.5	25
Reference	29.6	29.5	27.3	27.2	27.7	17.4	14.5	16.3	16.9	17.4
ROAM	31.5	29.3	30.3	29.7	27.6	31.5	29.6	32	30.5	28.6
Reference	22.6	24.6	22.4	24.3	23	21.9	29.6	30.4	28.3	27.2

Figure 15.7 shows example comparative visual results for the echocardiography sequence acquired on the expert side, in accordance with average visual impact, in the tested set-up. The original frame (no. 422) of the U.S. test sequence is reported in Figure 15.7(a). The corresponding received video frame with the non-adapted system is reported in Figure 15.7(b); this figure clearly shows evident artifacts, in terms of light stripes, affecting the accuracy of the diagnosis. Figure 15.7(c) shows the corresponding received video frame with the adapted system, presenting a much higher visual quality, also reflected in very good diagnosis accuracy. These preliminary results highlight the successful application of the outlined framework on the OTELO system.

The main problem in advanced m-health applications involving medical video transmission is in acceptance from the medical community, because such

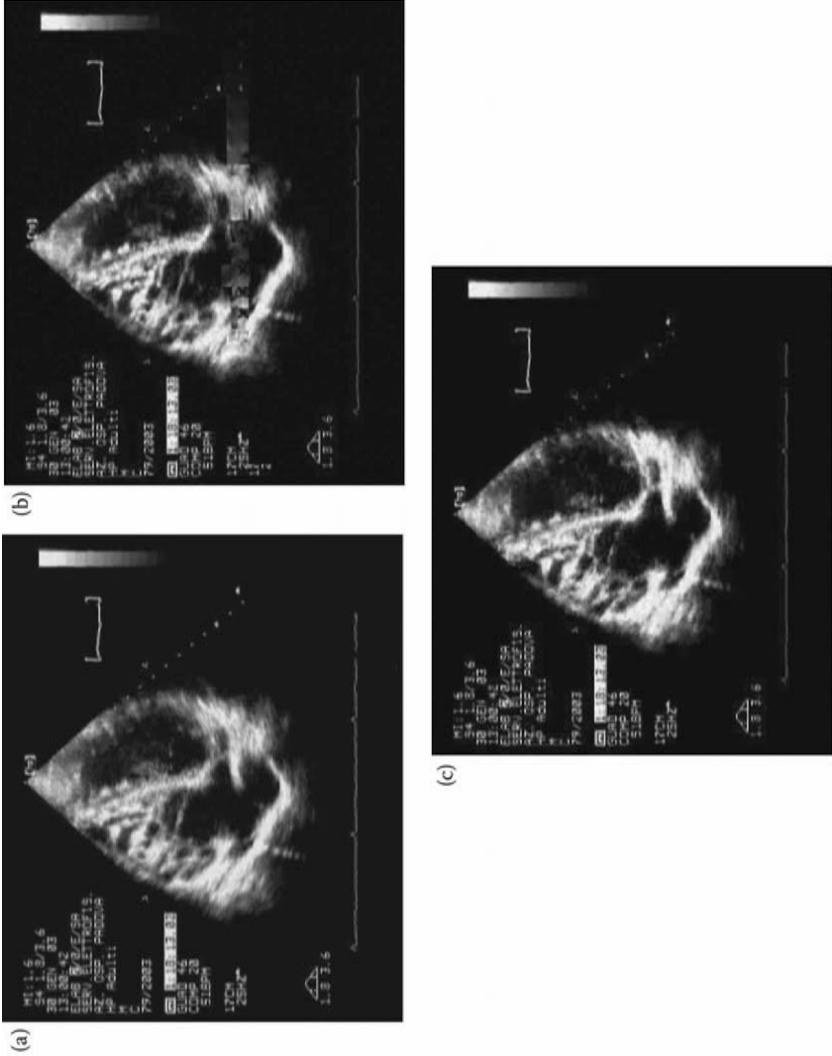


Figure 15.7 Visual results of received ultrasound video frames. (a) Frame no. 422, original; (b) frame no. 422, MPEG-4, ref. scheme; (c) frame no. 422, MPEG-4, cross-layer approach.

systems are often associated with poor video quality, not acceptable even for a first-level diagnosis. According to the obtained results, the presented approach allows an acceptable diagnostic accuracy for diagnosis, although a clinical performance evaluation should confirm it, and this is part of the ongoing work in this area.

To consider a nonobjective, automatically implementable, video quality assessment metric well representing diagnostic accuracy would help in improving system performance. In this case, in fact, the multi-layer controller would perform optimization according to a target better representing the medical goal and would use such a metric as input of its optimization algorithm.

15.5 Conclusions

In this chapter we introduced a cross-layer design for wireless medical video transmission from a mobile robotic tele-ultrasonography system. A cross-layer approach based on a multi-layer controller structure for enhanced medical video streaming in robotic tele-ultrasonography applications is presented.

In particular we presented a cross-layer design based on a dual application and physical controller units. The former drives the source encoder parameters with the knowledge of channel and network state information and of the medical video quality at the expert site, and the latter unit performs adaptation to the channel conditions and exploits the knowledge characteristics of the ultrasonography video stream to provide diagnostically acceptable received video streams.

The proposed framework is implemented in a simulated laboratory environment with images and video stream acquired from the real OTELO system. Results in the case of ultrasonography video transmission over a WLAN link show that a great improvement in terms of both objective medical video quality and of subjective quality is achieved with the proposed system.

Ongoing work is currently underway to test the performance of the proposed system in real medical and clinical settings to verify the performance of the robotic system in hospital and emergency situations. Further work is also planned to test the presented framework in 3.5G (HSDPA) systems.

Acknowledgments

The PHOENIX IST project (FP6-2002-IST-1-001812) is gratefully acknowledged by M.G. Martini and M. Mazzotti. Prof. Chiani (University of Bologna) is also acknowledged for the fruitful collaboration and support in the framework of the PHOENIX project. Co-authors R.S.H. Istepanian and N. Philip are grateful to the European Union for support for the EU IST-2001-32516 project “OTELO: Integrated, end-to-end, mobile tele-echography system.”

References

1. R.S.H. Istepanian, S. Laxminarayan, and C.C. Pattichis, *M-Health: Emerging Mobile Health Systems* (Springer, 2006).
2. R.S.H. Istepanian, E. Jovanov, and Y.T. Zhang, M-health: Beyond seamless mobility for global wireless healthcare connectivity [editorial], *IEEE Transactions on IT in Biomedicine*, 8(4): 405–414, 2004.
3. S.B.Z. Azami, P. Duhamel, and O. Rioul, Joint source-channel coding: Panorama of methods, in *Proceedings of CNES Workshop on Data Compression*, Toulouse, France (November 1996).
4. M.G. Martini, M. Mazzotti, C. Lamy-Bergot, J. H. uusko, and P. A mon, Content adaptive network aware joint optimization for wireless video transmission, *IEEE Communications Magazine*, 45(1): 84–90, 2007.
5. P.C. Cosman, Thoracic CT images: Effect of lossy image compression on diagnostic accuracy, *Radiology*, 190: 517–524, 1994.
6. H. Yu, Z. Lin, and F. Pan, Applications and improvement of H.264 in medical video compression, *IEEE Transactions on Circuits and Systems I, Special Issue on Biomedical Circuits and Systems: A New Wave of Technology*, 52 (12): 2707–2716, 2005.
7. S.A. Garawi, R.S.H. Istepanian, and M.A. Abu-Rgheff, 3G wireless communications for mobile robotic tele-ultrasonography systems, *IEEE Communications Magazine*, 44(4): 91–96, 2006.
8. F. Pereira and T. Ebrahimi, *The MPEG-4 Book* (Prentice Hall, 2002).
9. T. Wiegand, G.J. Sullivan, G. Bjntegaard, and G. Luthra, An overview of the H.264/AVC video coding standard, *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7), 560–576 (July 2003).
10. C.E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, 27(7), 379–423, 623–656 (July, October 1948).
11. J. Hagenauer and T. Stockhammer, Channel coding and transmission aspects for wireless multimedia, *Proceedings of the IEEE*, 87(10) (October 1999).
12. J. Hagenauer, N. Seshadri, and C.E. Sundberg, The performance of rate compatible punctured convolutional codes for digital mobile radio, *IEEE Transactions on Communications*, 38(7), 966–980 (July 1990).
13. J. Modestino and D. Daut, Combined source-channel coding of images, *IEEE Transactions on Communications*, 27(11), 1644–1659 (November 1979).
14. M.G. Martini and M. Chiani, Rate-distortion models for unequal error protection for wireless video transmission, in *Proceedings of IEEE VTC 2004 Conference*, Milan, Italy (May 2004).
15. D. Dardari, M.G. Martini, M. Mazzotti, and M. Chiani, Layered video transmission on adaptive OFDM wireless systems, *Eurasip Journal on Applied Signal Processing*, 2004(10), 1557–1567 (August 2004).
16. V. Srivastava and M. Motani, Cross-layer design: A survey and the road ahead, *IEEE Communications Magazine*, 43(12), 112–119 (December 2005).
17. M. van Der Schaar and N.S. Shankar, Cross-layer wireless multimedia transmission: Challenges, principles, and new paradigms, *IEEE Wireless Communications*, 12 (4), 50–58 (August 2005).

18. B.D.J. Sublett and A.C. Weaver, Design and implementation of a digital teleultrasound system for real-time remote diagnosis, in *Computer-Based Medical Systems*, 292–299 (June 2005).
19. R. Ribeiro, R. Conceicao, J.A. Rafael, A.S. Pereira, M. Martins, and R. Lourenco, Teleconsultation for cooperative acquisition, analysis and reporting of ultrasound studies, in *TeleMed 98*, London (November 1998).
20. D. DeCunha et al., The MIDSTEP system for ultrasound guided remote telesurgery, in *Proceedings of 20th International Conference of the IEEE Engineering in Medicine and Biology Society* (1998).
21. G. Kontaxakis, S. Walter, and G. Sakas, EU-TeleInViVo, an integrated portable telemedicine workstation featuring acquisition, processing and transmission over low-bandwidth lines of 3D ultrasound volume images, in *Information Technology Applications in Biomedicine* (November 2000).
22. A. Vilchis, J. Troccaz, P. Cinquin, F. Courreges, G. Poisson, and B. Tondu, Robotic tele-ultrasound system (TER): Slave robot control, in *1st IFAC Conference on Telematics Application in Automation and Robotics*, 95–100, Weingarten, Germany (July 2001).
23. K. Masuda, E. Kimura, N. Tateishi, and K. Ishihara, Development of remote echographic diagnosis system by using probe movable mechanism and transferring echogram via high speed digital network, in *Proceedings of IX Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON'01)*, 96–98, Pula (June 2001).
24. F. Courreges, P. Vieyres, R.S.H. Istepanian, P. Arbeille, and C. Bru, Clinical trials and evaluation of a mobile, robotic tele-ultrasound system, *Journal of Telemedicine and Telecare*, 2005(1), 46–49 (2005).
25. A. Gourdon, P. Poinet, G. Poisson, P. Vieyres, and P. Marche, A new robotic mechanism for medical application, in *Proceedings IEEE/ASME Conf. Advanced Intelligent Mechatronics*, 33–38, Atlanta, Georgia (September 1999).
26. S.A. Garawi, F. Courreges, R.S.H. Istepanian, H. Zisimopoulos, and P. Gosset, Performance analysis of a compact robotic tele-echography e-health system over terrestrial and mobile communication links, in *Proceeding of the 5th IEEE International Conference on 3G Mobile Communication Technologies—3G 2004*, 118–122, London (October 2004).
27. M.G. Martini and M. Chiani, Proportional unequal error protection for MPEG-4 video transmission, in *Proceedings of IEEE ICC 2001 Conference*, Helsinki, Finland (June 2001).
28. M.G. Martini, M. Mazzotti, C. Lamy-Bergot, P. Amon, G. Panza, J. Huusko, J. Pelto, G. Jeney, G. Feher, and S.X. Ng, A demonstration platform for network aware joint optimization of wireless video transmission, in *Proceedings of IST Mobile Summit 2006*, Mykonos, Greece (June 2006).
29. Z. Wang, L. Lu, and A.C. Bovik, Video quality assessment based on structural distortion measurement, *Signal Processing: Image Communication*, 29(1) (January 2004).

Chapter 16

Enabling Mobile Adaptive Computing Environments in Teleteaching and Telemedicine Applications

Tuan Cao-Huu

CONTENTS

16.1	What Is Mobile Computing?	320
16.1.1	Middleware for Mobile Systems	322
16.2	Adaptability: Critical for Mobile Computing	323
16.2.1	Transparency	324
16.2.2	Constraints of Mobile Computing Environments (MCE) ..	324
16.2.3	Application-Aware Adaptation.....	325
16.3	Mechanism of Adaptation	326
16.3.1	Adapting Functionality.....	326
16.3.2	Impact of Mobility on the CS Model	327
16.3.3	Adapting Data	327
16.3.4	Fidelity and Agility.....	328
16.4	How to Develop Adaptations.....	329

16.4.1	State-Based Approach	330
16.4.2	Where Adaptations Can Be Performed	330
16.4.3	Proxies	331
16.5	Support for Building Adaptive Mobile Applications	333
16.5.1	Odyssey.....	334
16.5.1.1	Odyssey Application Adaptation Model.....	334
16.5.1.2	Application Interaction with Odyssey.....	334
16.5.2	An Example Application Behavior	335
16.5.3	Odyssey Wardens.....	335
16.5.4	The Odyssey Viceroy.....	335
16.5.5	Rover	335
16.5.5.1	Optimize Use of Expensive Links	336
16.5.5.2	Make Use of Asymmetric Links.....	336
16.5.5.3	Stage Messages Near Their Destination	337
16.6	Conclusions	337
	References	337

The limitations of mobile computing environments (MCEs) for teleteaching and clinical medicine are discussed in this chapter. We examine the different mechanisms for adaptations that are critical for performance for mobile telemedicine. Our applications and implementation fall under the domain of mobile information access to a centralized server. Many algorithms are built in-house and involve interacting, annotating, constructing, and analyzing different kinds of information, medical images, and multidimensional, multimedia database and video streaming applications. We examine strategies that may enable these applications to continue working seamlessly in any mobile computing environment.

16.1 What Is Mobile Computing?

Basically, mobile computing systems are distributed systems. A distributed system is a collection of independent computers that appears to the users as a single coherent system. A network is required to communicate between different machines. Wireless communication is needed to enable mobility of communicating devices. In our context, we are concerned only about logical aspects of mobile communication. What is the difference between mobile computing and communication? Communication is necessary for computing. Many mobile computing tasks require mobile communication. But mobile communication does not solve all the problems. As we will see in this chapter, there are lots of issues that need to be resolved from a higher level perspective than just being able to exchange signals and packets. One example is to send a patient's diagnosis, data, and video images from Toronto to a specialist in Boston via a mobile phone, as shown in the TeleMedMail application in Figure 16.1.

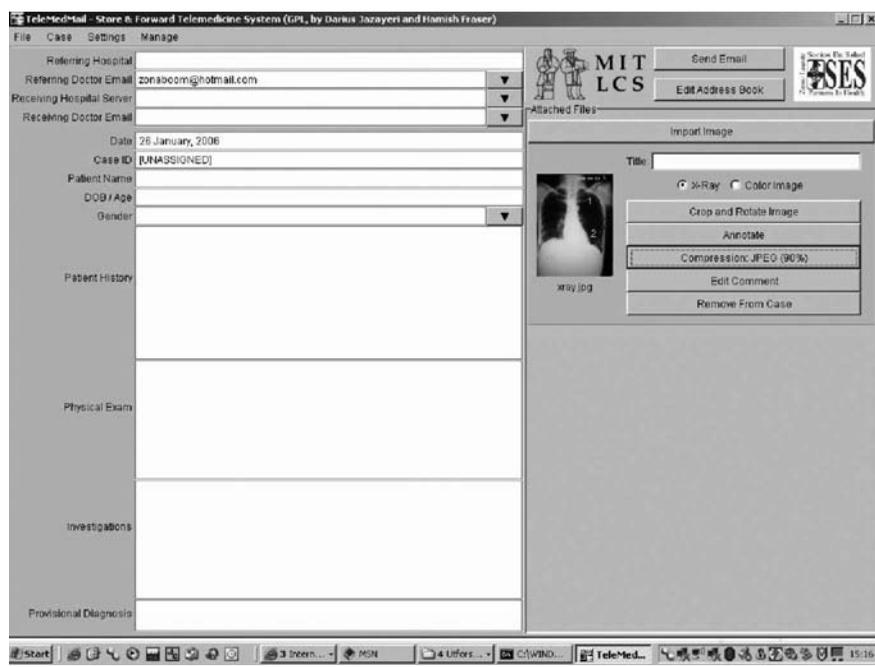


Figure 16.1 TeleMedMail is an open source mail client for telemedicine. Written in Java, it is a platform-independent, object-oriented development tool. TeleMedMail enables encryption and compression. When a case is sent through Email, the content, which is in the form of text, is changed to HTML by TeleMedMail and images are seen as thumbnails. TeleMedMail has the option of encrypting the data with an RC4 “secret key” algorithm. Our implementation of TeleMedMail offers additional support that allows resume-sending parts of a mail if a network connection is lost.

Another example is watching and interacting with video data streaming over the Internet. For telemedicine and clinical teaching, we probably wish to “interact” with the data set because if we can view a rotation, a cut through, or look from behind, the functional and anatomical images would be much more informative. Interactive engagement and fruitful interaction between members of a learning community, who may be globally distributed, may provide an environment for creative thinking. Suppose you are on the move and you are just watching a streaming video. Wireless communication is different from wired communication. When you have a wire you usually have a fixed bandwidth. Once the application is started and the movie begins you can watch the movie at a good quality of service (QoS). However, in wireless communication the wireless bandwidth is mostly shared among several users in a dynamic fashion. That is to say, there is no dedicated bandwidth available. Even if your application can reserve certain wireless bandwidth, due to the nature of wireless media, the usable bandwidth fluctuates. There are both

short- and long-term fluctuations. The question is how the application best responds to these fluctuations. One solution is the application responds in a uniform manner, irrespective of the data set. The other approach is to respond based on the type of “data set” you are studying. For example, suppose you are watching an action movie. The application can reduce the bandwidth requirement by switching from full-colored video to black-and-white or by reducing the resolution. On the other hand, if you are interested in only transcribing a diagnosis, the application may switch to simply audio streaming of the radiologist’s comments. So the decision is based on the content of the video, and the decision making may involve both the client and the server. The client needs to inform the server that it no longer wants the video frames.

In this chapter we will also re-examine computer (software) system design and see how it needs to be changed in order to accommodate mobility. As we will see, many of the changes have to do with providing mechanisms for adapting to changing environmental and system conditions such as location, available resources, etc. Mobile computing is about providing information anytime, anywhere, or more generally computing anytime and anywhere.

Mobile computing is also about dealing with limitations of mobile computing devices. For example, PDAs and laptops have a small interface and are battery powered. One of the major issues is how to do computation in an energy-efficient manner. The battery technology is not advancing at the same pace as the processor technology. One does not have to deal with these issues when designing systems for stand-alone systems or distributed systems. One may have to deal with issues of fault tolerance in distributed systems such as server crashes or network links failures. However, energy is usually not an issue. In mobile systems, energy becomes a resource like processing time or memory space. So now one has to design resource management techniques for energy as seen in traditional operating systems, which deal with process and memory management.

How about security or privacy? Wireless communication happens on an “open” wire and is relatively easy to tap. It might seem that traditional techniques of cryptography can be used to secure communication. However, the main problem is that the secure techniques designed for the wired networks are computational and communication intensive. Attempts to reduce their overheads lead to security schemes that are relatively easy to break. If there is little security, no serious medical application is viable.

16.1.1 *Middleware for Mobile Systems*

Different computers in a mobile computing environment may have different capabilities. A nycollaborative activity between these devices needs a n underlying software entity to deal with heterogeneity of the devices. This software entity is called middleware. Middleware may also allow a mobile device and a wired device to interact. When

a mobile client moves from one administration to another administration domain it may want to know what new services are available.

16.2 Adaptability: Critical for Mobile Computing

What are the techniques by which we are able to adapt to diverse environments in telemedicine and teaching? Can we incorporate some of these techniques into our computing systems? Anyone who has seriously used computers would like them to be more resilient and adaptive to our needs and circumstances. Computing systems and applications fail for various reasons. What is most frustrating is when they fail for no apparent reason. You install a new application and some other apparently unrelated application stops functioning. And sometimes, we would just desire that computers could learn from our past actions and act proactively and appropriately. Making systems resilient and adaptive is not a trivial task.

The vision of mobile computing is to be able to roam seamlessly with your computing devices anywhere while continuing to perform your computing and communication tasks uninterrupted. Many technological advances at various fronts such as security, privacy, resource allocation, charging, and billing have to be made in order to make this feasible. A quintessential characteristic for suitability to mobile computing of any solution for these problems is their ability to adapt to dynamic changes in computing and communication environment. The system's agility to react to the changes in the computing environment and to be able to continue the computing task uninterrupted is a new measure of performance in mobile computing environments.

Consider the scenario where you move from coverage area of one access point to another while, for example, a video streaming application is running on your computer. In order to continue receiving the video stream uninterrupted and possibly with no deterioration in the video quality, the video stream packets should now be automatically routed through the new access point. In an IP-based network this may involve the mobile client obtaining a new IP address in the new access point's IP network and informing the server so that it can send the packets to the new address. Many more sophisticated techniques have been developed. The point here is that the underlying protocols have to take many actions automatically in order to ensure continued connectivity and in this case uninterrupted viewing of the video stream. In essence the underlying system has to adapt to the changes in the environment such as the configuration, availability of communication, and computation resources and services. However, is this enough? More specifically, the above adaptation scheme did not take into account the applications' requirements and the applications themselves did not have any part to play in the adaptation. Does this application-transparent way of adapting suffice to meet the goals of mobile computing?

16.2.1 *Transparency*

Transparency is the ability of the system to hide some characteristics of the underlying implementation from the user. Much of the research effort in distributed computing has been devoted to developing mechanisms for providing various forms of transparencies. Examples include the following:

- Access transparency is the ability of the system to hide the differences in data representation on various machines and how a particular resource is accessed
- Location transparency is the ability of the system to hide where the resource is located; related to location transparency are
 - Name transparency (which ensures that the name of a resource does not reveal any hints as to the physical location of the resource)
 - User mobility (which ensures that no matter which machine a user is logged onto, she should be able to access resources with the same name)
- Failure transparency is the ability of the system to hide failure and recovery of a system component

Mobile computing systems can be viewed as a form of distributed system and attempts can be made to provide “mobility transparency,” which would encompass the transparencies mentioned above. This would in essence support application-transparent adaptation. But is this an achievable or even desirable goal for building mobile computing systems and applications? Let us closely look at the characteristics of the mobile computing environment and their implications.

16.2.2 *Constraints of Mobile Computing Environments (MCE)*

There are many constraints for a mobile computing environment:

1. Mobile computers can be expected to be more resource-poor than their static counterparts. With the continued rapid improvement in hardware technology, in accordance with Moore’s law, it is almost certain that a laptop purchased today is more powerful than the desktop computer purchased just a year or even a few months ago. However, mobile computers require some source of electrical energy, which is usually provided by a battery pack. Because batteries store a finite amount of energy, they need to be replaced or recharged. The first option costs money and the second option, although cheaper in terms of money expended, requires plugging in the computer for recharging — restricting mobility. This has impact on the design of mobile computers — all the hardware and software components in the mobile computers are designed to reduce energy consumption and increase the lifetime of the batteries. For example, processors on mobile computers are designed to consume less energy — consequently achieving lower computation performance (MIPS).

2. Mobile computers are inherently hazardous, less secure and less reliable.
3. Mobile connectivity is highly variable in performance and reliability. Disconnections (voluntary and involuntary) are common. The link bandwidth can vary by orders of magnitude.
4. In general resource availability and quality varies dynamically.

These characteristics of MCEs require rethinking of how mobile applications and systems should be designed. Resource paucity and lower reliability of mobile devices point towards designing the system in such a manner that more reliance is put on the static infrastructure. On the other hand the possibilities of disconnections and poor connectivity point towards making the system less reliant on the static infrastructure. Further, as the mobile moves around (or even if it does not) its situation over time keeps changing. Hence, depending upon the situation the mobile should change its behavior so as to be either more or less reliant on the static infrastructure.

16.2.3 Application-Aware Adaptation

Who should be responsible for adaptation? Application or system? There are two extreme approaches to designing adaptive systems: application-transparent (the system is fully responsible for adaptation) and *laissez faire* (the system provides no support at all).¹ Obviously, the *laissez-faire* approach is not desirable because it puts too much burden on the application developer. Further, no support from the underlying system restricts the types of adaptations that can be performed. However, as the following example points out, application-transparent approach is not sufficient either. Consider two different multimedia applications: in one you are videoconferencing using a mobile device, and in the other you are watching a live video stream from a remote server on a mobile client. Now consider the following scenarios: (1) you move from an area where sufficient bandwidth for your application is available, to an area where the amount of bandwidth is lower than needed by your applications; and (2) your battery power level drops considerably. Both scenarios have to do with change in availability of resources. How would you like your system/application to behave under each scenario?

In the application-transparent (user) approach the system/application may behave the same, irrespective of which application is running. For example, in the first scenario a non-adaptive system may just do nothing and let the audio/video quality drop. In the second scenario, the system may just give a warning to the user without any assistance in how to deal with the situation. In an adaptive system, various behaviors can be envisioned. For example, the system may try to do best in both situations. However, the system's adaptation does not take into account the type of the application that is running. For example, in the first scenario the system may try to adapt by requesting the server or other peers to start to send lower quality video — in effect requiring lower bandwidth. In the second scenario, the

system may try to conserve energy by reducing the intensity of the backlight of the display (besides warning the user of the lower battery power level). A still more adaptive approach is possible in which the system interacts with the user/application in deciding how to adapt.

In the application-transparent approach, the responsibility of adaptation solely lies with the underlying system. On the other hand, in the application-aware approach the application collaborates with the underlying system software. The underlying system provides status information about the resources. The application takes this information and makes a decision on how to adapt to changes in the resource availability. Each application can adapt in its own way.

16.3 Mechanism of Adaptation

What can be adapted? As we will see in this section both the functionality of various components in the mobile application and data that is delivered to the application can be adapted. The next question is, how to adapt? In the context of the client–server (CS) model, functionality can be adapted by varying the partition of duties between the client and the server, e.g., during disconnection, a mobile client works autonomously, while during periods of strong connectivity, the client depends heavily on the fixed network, sparing its scarce local resources. In the following we look at these approaches in more detail.

16.3.1 *Adapting Functionality*

The first approach is to change dynamically the functionality of the computational entities involved in response to the change in the operating condition. An example of this approach is the extended client–server model.¹ The client–server paradigm is the most widely used architecture for distributed computing. In the standard CS model the roles of the client and the server are defined usually at the design time and remain fixed during (run-time) execution of the system. Servers (usually a small number) provide some services (such as access to database, Web pages, allocation of temporary IP address, name translation, etc.) to (usually a larger group of) clients. A client (or the underlying system — middleware) may dynamically select the server from which to request the service. A server may or may not maintain information (or state) regarding the clients it is providing service to. The state information may be maintained as soft state or hard state. Once installed, soft state has to be updated periodically to avoid automatic deletion by the state maintainer (in our case, a server) whereas hard state once installed requires explicit deletion. Soft state is useful in systems with very dynamic configurations, such as mobile systems. The reason is that soft state requires no explicit action to make the state information consistent with dynamic changes in the system.² Soft state is used in various protocols such as RSVP

(Resource Reservation Protocol)³ and IGMP (Internet Group Management Protocol)⁴ to adapt gracefully to the dynamic changes in the system state.⁵ Specifically, in case of data servers (such as file servers) the client–server model (as implemented by CODA⁶) has the following characteristics:¹

- A small number of trusted server sites constitute the true home of data
- Efficient and safe access of data possible from a much larger number of untrusted client sites
- Techniques such as caching and pre-fetching are used to provide good performance
- End-to-end authentication and encrypted transmission are used to preserve security

Developers of the CODA system point to the advantages of the CS model: good scalability, performance, and availability:

The CS model decomposes a large distributed system into a small nucleus that changes relatively slowly, and a much larger and more dynamic periphery of clients. From the perspective of security and system administration, the scale of the system appears to be that of the nucleus. From the perspective of performance and availability, a client receives service comparable to stand-alone service.³

16.3.2 Impact of Mobility on the CS Model

The CS model permits a resource-poor client (mobile) to request a resource-rich server to perform expensive computations on its behalf. For example, the client can send a request to the server, go to sleep (to conserve energy), and later wake up to obtain the result from the server. For the sake of improved performance and availability, the boundary between the clients and servers may have to be dynamically adjusted. This results in an extended client–server model. In order to cope with the resource limitations of clients, certain operations that are normally performed at the client may have to be performed by resource-rich servers. Conversely, the need to cope with uncertain connectivity requires the clients sometimes to emulate the functions of the servers, resulting in short-term deviation from the classic CS model. However, from the long-term perspectives of system administration and security, the roles of servers and clients remain unchanged.

16.3.3 Adapting Data

Another way to adapt the resource availability is by varying the quality of data (fidelity) made available to the application running on the mobile client. Fidelity

is defined as the “degree to which a copy of data presented for use at the client matches the reference copy at the server.”⁷ This kind of adaptation is extremely useful in mobile information access applications. The quality-of-service requirements for such applications are

- *Information quality*: Ideally, a data item being accessed on a mobile client should be indistinguishable from that available to the application if it were to execute on the server storing the data.
- *Performance from client’s perspective*: Latency of data access should be within tolerable limits.
- *Performance from system’s perspective*: Throughput of the system should be maximum.

In general it is difficult to provide both high performance and highest-quality information in a mobile computing environment. In some cases the information quality can be traded off against performance. The idea behind data adaptation is as follows. Assume that any data item accessed by a client has a reference copy (at a remote server) that is complete and current. When resources are plentiful the client will access and manipulate the reference copy. However, when resources are scarce the mobile client may choose to access or manipulate a data item that has been degraded, consuming fewer resources.

16.3.4 Fidelity and Agility

Data fidelity is a property of many dimensions. One common dimension is consistency that is shared by all data items, irrespective of their type. The other dimensions are type dependent:

- *Video data*: Frame rate and image quality
- *Spatial data such as topographical maps*: Minimum feature size
- *Telemetry data*: Sampling rate, timeliness

Dimensions of data fidelity can be exploited for the sake of adaptation required for handling mobility. For example, a mobile client can choose to use the locally cached stale copy when it is disconnected from the server and a possibly more current copy at the server is inaccessible. Fidelity of data can be changed in several ways and requires knowledge of data representation. For example, a video stream can be degraded by reducing the frame rate, reducing the quality of individual frames, or reducing the size of the individual frame. Another point to note is that different applications using the same data may exploit different trade-offs among dimensions of fidelity. For example, a video editor may choose to slow the frame rate whereas a video player may choose to drop the frames. When developing different strategies

for trading-off data fidelity dimensions against performance, an issue that arises is how to determine which strategy is better. Developers of the Odyssey system have evaluated their system using agility as a metric.

Agility is defined as the speed and accuracy with which an (adaptive system) application detects and responds to changes in its computing environment, e.g., change in resource availability. The larger the change is, the more important the agility is. For example, for an adaptive system that tries to adapt to availability of connection bandwidth one can try to determine how well the system reacts to sudden changes in bandwidth. One issue is how to model changes. Developers of Odyssey have used reference waveforms: set-up, step-down, impulse-up, impulse-down. They generated these waveforms using a trace modulation technique, which emulates a slower target network over a faster wired LAN. The results obtained from such studies should be interpreted by keeping in mind that an adaptation strategy is strictly better if it provides better fidelity with comparable performance or better performance with comparable fidelity. Further comparison must take into account application goals.

16.4 How to Develop Adaptations

In general it is difficult to enumerate all the mechanisms that can be employed to construct adaptive programs. However, it should be intuitively clear that all adaptive programs must adapt to some detectable “change” in their “environment.” Either a program can implement its own mechanisms to detect the changes or mechanisms may be provided by some other entity (e.g., middleware or operating system) to make the program aware of these external changes. In general, we can view these entities as software “sensors” (as opposed to hardware sensors we will discuss later in the chapter). For example, a TCP client adapts its transmission window size by (indirectly) monitoring the congestion level in the network. Conceptually, it maintains a software timer for each packet sent, and as long as it receives an acknowledgment for a packet before its timer expires, it keeps increasing the size of its transmission window (up to a maximum allowable window size). However, upon a loss event (timeout or receipt of triple-ack for a packet) it assumes that the loss is due to buildup of congestion in the network and so it backs off by reducing its transmission window size.*

* As a side note, this behavior is not suitable in wireless networks since the loss of packets may be due to high-error rate in a wireless link on the delivery path or may be because an endpoint has moved. In such cases the TCP client should not back off but continue trying to push the packets through the network. Many techniques have been developed to “adapt” TCP for wireless networks. Examples include TCP-snoop and Indirect TCP.

16.4.1 State-Based Approach

In this approach, changes in the MCE are viewed as state transitions. Irrespective of how the state of the environment is sensed, the adaptation (of functionality or data) can be performed when a state transition occurs. Logically, each system state corresponds to an “environmental state.” Each system state is associated with some appropriate functionality. As long as the environment remains in a particular state, the system behaves according to the functionality associated with that state. When environment state changes, the system may have to perform some chapter-keeping functions associated with state transition before assuming the functionality associated with the new state. In order to perform these operations the system may have some additional states.

For example, consider the functionality adaptation in the CODA (continued data availability) distributed file system developed at Carnegie Mellon University. CODA is designed to maximize the availability of data at the expense of possible access to stale data. Each CODA client (called Venus) maintains a local cache. Venus adapts its functionality based on the state of the connectivity between the client and the server. Venus uses the following four states:

1. *Hoarding:* Venus is in hoarding state when it has strong connectivity with the server. In this state the client aggressively pre-fetches files from the server to store locally. Files to be pre-fetched are decided based on user preference and access pattern.
2. *Emulating:* Venus is in emulating state when it is disconnected from the server. In this state, the client emulates the server by optimistically allowing both read/write access to local files. In order to update the primary copy of the files on the server and detect any conflicting updates, the client maintains a log of all the file operations.
3. *Write-disconnected:* Venus is in write-disconnected state when the client has a weak connectivity to the server. In this state, a CODA client decides whether to fetch files from the server or to allow local access.
4. *Reintegration:* Venus enters this state when the connectivity improves to “strong connectivity.” In this state, Venus resynchronizes its cache with the accessible servers. Log of operations is used for this purpose. If any conflict is detected, then user assistance may be required. Upon completion of resynchronization, Venus enters the hoarding state.

Note that the first three states correspond to some environmental state but the last state corresponds to a state transition (Figure 16.2).⁷

16.4.2 Where Adaptations Can Be Performed

In a distributed application, in particular a client–server application, the adaptation can be performed at the client and/or at the server or at both. Further, there are

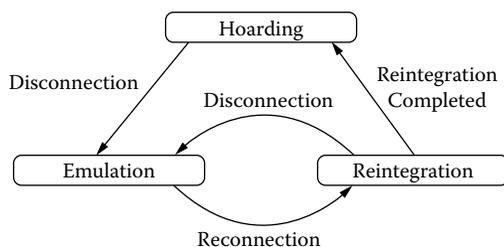


Figure 16.2 The state-transition diagram of a CODA client with respect to a volume or CODA state transition diagram.

additional possibilities. The adaptation can also be performed within the network, e.g., at an intermediate software entity called the proxy. For example, consider a typical client–server application: a video streaming application. Several different adaptations may be performed at different components located in different points in the data and control path between the client and the server:

- *Adapting to the hardware/software capabilities of the mobile device:* In the proxy or at the server
- *Adapting to the connectivity of the mobile device:* At the server or client
- *Adapting to the resource availability at the mobile device:* At the client

Let us look at some concrete examples to get a better understanding of incorporating adaptations in mobile applications.

16.4.3 Proxies

Proxies have been used by many applications to perform various tasks such as filtering data and connections (e.g., security firewalls), and modifying control data (network address translators [NATs] change the IP fields). Of particular interest to data adaptation are transcoding proxies that modify (adapt) a data flow to suit the end mobile device. For example, if the end device is not capable of handling full motion video, a transcoding proxy may convert it to a form that can be displayed on the end device (Figure 16.3).

Browsing over wireless networks can be expensive and slow due to characteristics of pay-per-minute charging (in cellular networks) and characteristics of wireless communication. Additionally, Hyper Text Transport Protocol (HTTP) was not designed for wireless networks and suffers from various inefficiencies: connection overhead, redundant transmission of capabilities, and verbose protocol. We have experimented with techniques and algorithms that reduce user cost and response time of wireless communications by intercepting the HTTP data stream and performing various optimizations on it.

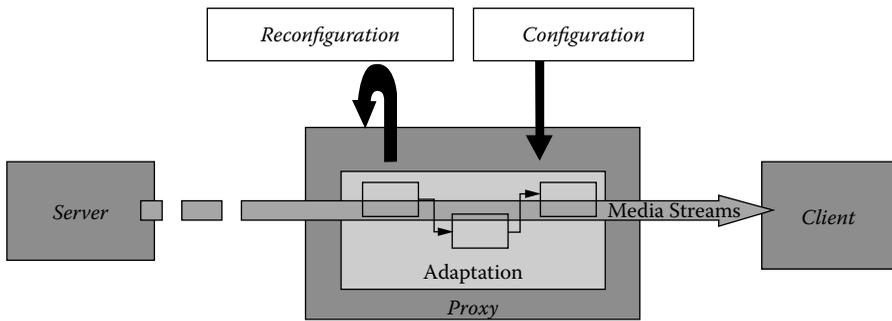


Figure 16.3 Adaptation using proxies.

WebExpress is a client/intercept system for optimizing wireless Web browsing. It utilizes proxies (called intercepts) that allow it to be used with any Web browser and any Web server. It enables WebExpress to intercept and control communications over the wireless link for the purpose of reducing traffic volume and optimizing the communications protocol to reduce data transmission. WebExpress architecture consists of two components that are inserted into the data path between the Web client and Web server: a client-side intercept (CSI, also known as client-side proxy) and server-side intercept (SSI, also known as server-side proxy). CSI is a process that runs in the end user client mobile device; SSI is a process that runs within the wireline network. One of the features of this client-proxy-server model (also called intercept model) is that the proxies are transparent to both Web browsers and servers. This makes this adaptation technique insensitive to the evolution of technology. This is a very important advantage since HTML/HTTP technology was (and still is) rapidly maturing when WebExpress was developed. Another advantage is highly effective data reduction and protocol optimization without limiting browser functionality or interoperability. WebExpress employs several optimization techniques such as caching, differencing, protocol reduction, and header reduction:

- *Caching:* WebExpress supports both client and server caching using a least recently used (LRU) algorithm. Cache objects persist across browser sessions. Caching reduces volume of application data transmitted over wireless link through cross-browser sessions.
- *Differencing:* Caching techniques do not help in CGI processing where each request returns a different result, e.g., a stock-quote server. However, different replies from the same program (application server) are usually very similar. For example, replies from a stock-quote server contain lots of unchanging data such as graphics. For each dynamic response from a CGI (HTML file) cached at the SSI, the SSI computes a base-object for page before sending it to CSI. If SSI receives a response from the CGI server and the CRC received does not match the CRC of the base object, SSI returns both difference stream and base object. This is called a basing operation in WebExpress parlance.

Rebasing is carried out in the same fashion when the SSI detects that the difference stream has grown beyond a certain threshold.

- *Protocol reduction:* Repeated TCP/IP connections and redundant header transmissions present additional overhead. The WebExpress system utilizes two main techniques to reduce this overhead and optimize browsing in a wireless environment.
 - *Reduction of TCP/IP connection overhead using virtual sockets:* WebExpress establishes a single TCP/IP connection between the CSI and SSI. CSI sends requests over this connection. SSI establishes a connection with the destination server and forwards the request. Thus overhead is incurred between SSI and Web server but not over the wireless link. Virtual sockets are used to provide multiplexing support. Virtual sockets are implemented in the following manner. Data sent is prefixed by a virtual socket ID, command byte, and a length field. At CSI the virtual ID associates with a real socket to the browser. At SSI the virtual socket ID is mapped to a socket connection to a HTTP server. This mechanism permits efficient transport of HTTP requests and response while maintaining protocol transparency.
 - *Reduction of HTTP headers:* HTTP request headers containing lists of MIME content types can be hundreds of bytes in length. CSI allows this information to flow in the first request and saves the list. On subsequent requests CSI compares the list received with the saved one. If the two lists match, the MIME content-type list is deleted from the request. SSI inserts the saved one if none is present. Transparent (proxy-based) architecture of WebExpress allows the operation of commercial Web application of wireless networks. Differencing and virtual sockets offer the most critical optimizations in the WebExpress system.

16.5 Support for Building Adaptive Mobile Applications

Adaptations should cater to the needs of individual applications. We have argued that applications are in a better position to perform application-specific adaptations (than system alone). But what does this mean with regard to where adaptations can be performed? Not just that the application makes local adjustments. The application should also collaborate with other adaptation technologies that are available in other components of the system. For example, in the CS scenario both client and server may need to adapt. The advantage of the application-aware adaptation is that the application writer knows best how to adapt. However, does the application writer know networks as well as his application? Further, the application-aware adaptation tends to work only at the client or perhaps at the server. If this is the only mechanism for adaptation, without any monitoring on resource usage by each

application, this may result in selfish behavior by the applications. In the following we look in detail at some efforts for developing adaptive applications.

16.5.1 *Odyssey*

Odyssey aims to provide high fidelity and support concurrent mobile applications with agility. It emphasizes collaboration between application and OS in performing adaptation to handle constraints of MCE, especially those imposed by presence of wireless link.

Imagine a user with a lightweight/wearable mobile computer with ubiquitous wireless access to remote services, unobtrusive heads-up display, microphone, earphones, speech for computer interactions, and online language translation. The user has ubiquitous connectivity but the quality varies as he moves and different networks can be accessed. The user simultaneously gets voice, video, and other data sent to him. When a user moves to a relatively shadowed area and network bandwidth drops dramatically, Odyssey informs the video, audio, and other applications of the changes, thus allowing them to make proper adaptations in their network usage.

The Basic Odyssey Adaptation Model is OS support on the portable machine monitor's condition. Each application interacts with OS tools to negotiate services. When things change, the OS notifies applications of what has happened.

Why use application-aware adaptation here? The assumption is that, for this environment, only the application knows what to do. If the OS makes the decision, it may do the wrong thing. But the OS must be involved to ensure fairness between competing applications.

Odyssey architecture consists of two main components: Viceroy and Wardens. Viceroy performs centralized resource management and monitors the availability of resources notifying applications of changes. Wardens provides (data) type-specific operations (tsop) to change the fidelity. They are also responsible for communicating with the servers and caching data.

16.5.1.1 *Odyssey Application Adaptation Model*

Odyssey applications do not interact directly with their remote servers. Applications talk to their wardens, and wardens talk to the servers. Applications tend to have limited roles in actually adapting transmissions. They may know about different formats and tolerances and accept data in its different adapted versions.

16.5.1.2 *Application Interaction with Odyssey*

All data to and from the server flows through Odyssey. Also, applications must register their preferences and needs with Odyssey in the form of requests. A request

specifies that an application needs a particular resource within certain limits, e.g., between 100 kbps and 1 Mbps of bandwidth. If the request can currently be satisfied, it is. If things change later, the app is notified using an upcall to the applications. When previously satisfied requests can no longer be satisfied, Odyssey performs an upcall to the application. In response, the application can adjust itself and make another request. Note that upcalls can occur because things got worse or got better.

16.5.2 An Example Application Behavior

A video application requests enough bandwidth to receive 20 fps in color. Odyssey says “in your dreams.” The application requests bandwidth sufficient for 10 fps in black and white. Specifying the minimum and maximum needed for this can improve quality if it is worthwhile. The channel gets noisy and bandwidth drops. Odyssey makes an upcall saying that the bandwidth is outside the limits. The app requests lower bandwidth suitable for the lower fps rate.

16.5.3 Odyssey Wardens

Odyssey wardens mediate server/application interaction. A warden is a data-type specific module capable of performing various adaptations on that data type, e.g., caching and pre-fetching is done by a warden. If you want Odyssey to handle a new data type, you not only need to alter the application but you also need to write a new warden. The better the warden you write understands the data type, the better adaptivity will be possible.

16.5.4 The Odyssey Viceroy

The central controlling facility that handles sharing of resources, the viceroy notices changes in resource conditions. If they exceed pre-set limits, it informs the affected applications using the upcall mechanism.

Developers of Odyssey have evaluated their system to answer the following questions: how agile is Odyssey in the face of changing network bandwidth? How beneficial is it for applications to exploit the dynamic adaptation made possible by Odyssey? How important is centralized resource management for concurrent applications? Interested readers should refer to Noble et al.⁸ for details.

16.5.5 Rover

Rover is a non-object-based software toolkit for developing both mobility-aware and mobility-transparent client-server distributed applications.^{6,9} It provides the application developers with relocatable dynamic objects (RDOs) and queued

remote procedure calls (QRPCs), two programming and communication abstractions specifically designed for assisting applications in harsh network environments such as mobile computing environments. RDOs can be used to reduce interaction between two weakly connected entities such as a client on the mobile device and server in the wireline network. Rover RDOs are objects with well-defined interfaces dynamically loadable from a server to a client. This in essence moves objects to the client machine, thus avoiding the client having to communicate with objects at the server. QRPCs can be used to handle disconnections. Rover QRPCs are essentially non-blocking RPC calls and support split-phase operations. That is, it allows an application to make an RPC call without worrying about whether the destination is reachable. If the destination of the RPC call is not reachable at the time of the call, the call is queued up. Upon availability of connection to the RPC's destination, RPC is performed. The result of the RPC call is delivered asynchronously to the application.

How does one use Rover? By writing (and perhaps rewriting) your application using the tools of the toolkit — invoke these tools in situations where network connectivity is bad. This, however, requires a good understanding of programming networks and mobility. To use RDOs application (usually clients) import RDOs from a server. It invokes methods on imported RDOs. When done, RDOs are exported back to the server. Rover can cache copies of objects. This allows clients to use the cached copy, instead of fetching it from the server. Updates to objects are handled by an optimistic client-server replication method. RPCs are queued only at the client. They are stored until the client can handle them. When an appropriate level of connectivity is established, Rover clears the queue intelligently using an RPC prioritization mechanism. It can also batch related requests. If the client is not available when the response comes back, the server drops the response. Queued requests at the client will eventually be replayed. This may cause some inefficiencies but simplifies application design. A mobile application can employ QRPCs in various situations dynamically to optimize cost to the user or the performance of the application.

16.5.5.1 Optimize Use of Expensive Links

Consider the situation where the mobile user pays for wireless connectivity based on the duration of usage, e.g., pay-per-minute plans for cellular phones. To optimize the monetary cost to the mobile user, QRPC can batch several requests and disconnect from the network after invoking all the batched QRPC calls through a single connection.

16.5.5.2 Make Use of Asymmetric Links

The queued RPC requests are not associated with a particular network interface. Thus, responses can be obtained over any network device, including a different one.

This permits an application, for example, to launch requests over expensive links and receive responses over cheaper links. This is beneficial for situations where the request size is much smaller than the expected response size (e.g., in Web browsing) and the response can be used incrementally as it arrives.

16.5.5.3 Stage Messages Near Their Destination

Arrange for RPC queuing to occur just before a “bad” link. If link quality improves, RPC queue will be cleared out. Meanwhile, the transmitter is not blocked on a bad link.

16.6 Conclusions

We discussed the limitations of mobile computing environments (MCEs). We examined and explained how adaptations can be done to enable these kinds of applications to continue working seamlessly in any mobile computing environment. Many applications are based on the client–server model. We explained that in order to enable adaptation we will have to extend this model. We saw how and why we need to extend the CS model. Once we established the need for adaptation, we looked at who should be responsible for adaptation — the system or the application. Then we looked at application-aware adaptations and illustrated examples.

Approaches to developing application-aware adaptive applications consist of purely internal, layered outside application, using special OS features and libraries, and interacting with other mechanisms, e.g., intelligent use of proxies.

One simple solution is to build application-aware Web browsers’ interaction with HTTP proxy. A browser can get any adaptation it wants, provided it can name a proxy that will do it. This works well mostly at the client side only. So the key question for future research in this direction is how to assign an application the control over the entire path of its data transmission.

References

1. M. Satyanarayanan, Mobile Information Access, *IEEE Personal Communications*, 3(1), 1996.
2. R. Han, P. Bhagwat, R. LaMaire, T. Mummert, V. Perret, and J. Rubas, Dynamic adaptation in an image transcoding proxy for mobile Web browsing, *IEEE Personal Communications*, 5(6): 8–17, 1998.
3. M. Gouda and T. Herman, Adaptive programming, *IEEE Transactions on Software Engineering*, 17: 911–921, 1991.
4. E. Pitoura and G. Samara, *Data Management for Mobile Computing*, Kluwer/Academic Press, 1998.

5. M. Satyanarayanan, *Fundamental Challenges in Mobile Computing*, PODC, 1996.
6. A. Joseph, A. F. deLapinasse, J. A. Tauber, D. K. Gifford, and M. F. Kaashoke, Mobile computing with Rover toolkit, *IEEE Transactions on Computers*, 46(3): 337–352, 1997.
7. B. C. Housel, D. B. Lindquist, and G. Samaras, WebExpress: A client/intercept based system for optimizing web browsing in a wireless environment, *Journal of ACM/Baltzer Mobile Networking and Applications (MONET)*, Special Issue on Mobile Networking on the Internet, 3(4): 419–431, December 1998.
8. B. Noble, M. Satyanarayanan, D. Narayanan, J. E. Tilton, J. Flinn, and K. Walker, Agile Application-Aware Adaptation for Mobility, SOSP'97, December 1997.
9. A. Joseph, A. F. deLapinasse, J. A. Tauber, D. K. Gifford, and M. F. Kaashoke, Rover: A Toolkit for Mobile Information Access, SOSP'95, December 1997.
10. B. Noble, M. Price, and M. Satyanarayanan, A programming interface for a application-aware adaptation in mobile computing, in *Proceedings of the 1995 USENIX Symposium on Mobile and Location-Independent Computing*, April 1995.

Chapter 17

Building a Mobile Healthcare Network within Public Networking Infrastructures

Ziad Hunaiti

CONTENTS

17.1 Introduction	340
17.2 Mobile Networks and Health Care	340
17.3 Implementation of Mobile Health Care Networks	342
17.4 Conclusions and Outlook.....	345
References	345

This chapter discusses a complete solution concerning the establishment of a mobile health care network for use in the health care sector. The solution will allow the construction of an autonomous network dedicated for health care purposes using the available broadband public mobile networks and public Internet. Virtual private network (VPN) techniques will be utilized in order to ensure secure and reliable communication channels for medical data anywhere, anytime. Building

a mobile health care network will contribute to the enhancement of health care sector performance. In addition, it will be vital in many critical scenarios, such as pre-hospital treatment, remote consultation by emergency services, and other applications that require remote access of medical data.

17.1 Introduction

During the mid- to late 1990s the world witnessed a great revolution in the field of information and communication technologies. The development of the Internet and mobile communications has redefined the ways of exchanging information. This is even extended to the redefinition of other aspects of life such as business and education. It has become much easier than ever before for people to access communication media. As a result, these technologies became an integral part of most businesses today. Document exchange and business meetings have evolved from the physical world to the digital, virtual world. Email has replaced faxes and most paper-based mail, and videoconferencing reduces the need to travel for meetings. This enabled effective operation and data exchange in national and international organizations while reducing the costs of managing these complex organizations. Virtual private networks (VPNs) provide the security infrastructure that businesses can rely on. Furthermore, the introduction of mobile communication, in particular, third-generation mobile networks, wireless local area networks (WLANs), and World Interoperability for Microwave Access (WiMAX) had a great impact on the productivity of businesses. This is mainly due to wireless broadband connectivity from local vicinity to a wider scale, which makes it possible to perform tasks while users are traveling from one place to another or while they are in a café, a restaurant, or an airport. The increase in the use of mobile communication for information exchange has increased the need for more bandwidth to handle new applications.^{1,2} This in turn led mobile operators to expand their networks' capacity to attract new customers and meet the ever-increasing demand from current customers. One of the potential customers is the health care sector.³ This chapter reviews the way mobile communications has been used in this sector and suggests a solution for building a mobile health care network utilizing the current and future mobile networks.

17.2 Mobile Networks and Health Care

Along with an increasing demand for bandwidth to transfer data over the mobile networks for businesses, several attempts have been made to use public mobile networks in the health care sector as a complementary network to other public infrastructure such as Internet and satellite communications. Internet and satellite communications have been used successfully to transfer medical data (telemedicine or telehealth⁴).

Researchers have been attempting to transmit medical data over mobile links since the deployment of second generation mobile networks [(2G), sometimes referred to as Global System for Mobile communication (GSM)], which for most applications was limited to a bandwidth of around 10 kbps.

Some researchers have managed to transmit vital signs and ECG over this link.^{4,5} However, the small bandwidth limited the ability of efficient transmission of more sophisticated data, which required a higher data rate (see Table 17.1).⁶ The anticipation was to use mobile networks to achieve what has been achieved over fixed-line networks and satellite links.

The development of the GSM, 2.5G, or General Package Radio Service (GPRS) promised an increased bandwidth. GPRS operates on the concept of packet switching service over circuit switching service. Nevertheless, the bandwidth increase that GPRS offered was limited due to the fact that it was based on the same technology as GSM. Service providers offered to dedicate four GSM slots for downlink and one for uplink. Theoretically, this offered 10 kbps of uplink speed and a maximum of 40 kbps on the downlink speed.⁷ Despite this, users still experienced high delay, low and fluctuated bandwidth, packet loss, and temporary loss of connection (link outages) using the GPRS.⁷ In most mobile networks GPRS did not offer any improvement on the uplink speed. Therefore, it was only useful in cases where medical data is needed to be transferred from the hospital to mobile medical staff.³

Despite the fact that GPRS did not offer a great Internet experience in terms of quality of service (QoS), its real significance lies in the fact that it was a transitional stage towards the third-generation (3G) system, also known as the Universal Mobile Telecommunications System (UMTS). The UMTS used a completely different air

Table 17.1 Some of the typical data rates required by telemedicine devices

<i>Device</i>	<i>Required Data Rate</i>
Digital blood pressure monitor	<10 kbps
Digital thermometer	<10 kbps
Digital audio stethoscope with ECG	<10 kbps
Ultrasound	256 kb per image
MRI	384 kb per image
Scanned x-ray	1.8 Mb per image
Digital radiography	6 Mb per image
Mammogram	24 Mb per image
Compressed and full motion video	384 - 1544 kbps

Source: M. Ackerman, R. Craft, F. Ferrante, M. Kratz, S. Mandil, and H. Sapci, Telemedicine technology, *Telemedicine Journal and e-Health*, 8, 1, 71–78, 2002. With permission.

interface and modulation techniques. This made it possible to provide a bandwidth of up to 64 kbps for uplink speed and 384 kbps for downlink speed. In addition, UMTS had shown a much more stable performance than GPRS.⁸ UMTS has a direct impact on the amount of medical data transmission. A very useful high-bit data rate can be transmitted easily over available networks. One example of this is the transmission of a live video image from an ambulance to a receiving hospital for emergency treatment. Transmission of live video images was not possible over 2G and 2.5G, whereas 3G may enable up to 15 frames per second^{9,10} to be transmitted. This increase in transmission speed will enable hospital medical staff to carry out remote diagnosis and thereby save valuable time in dealing with critical medical cases. Recently, many networks have started upgrading their UMTS to a true wireless broadband network, 3.5G High Speed Downlink Packet Access (HSDPA).¹¹ HSDPA promises to provide increased downlink speeds beyond 10 Mbps (initially 1.8 Mbps) while increasing the uplink speed up to 384 kbps. In addition, HSDPA promises to provide lower delay and higher reliability. Within the next few years, a further upgrade is expected: the High Speed Uplink Packet Access (HSUPA) or 3.75G,¹² which is expected to increase the uplink speed to 5.7 Mbps. High-speed OFDM Packet Access (HSOPA) is a further development that could result in an increase in the bandwidth of up to 37 Mbps. 3G LTE (Long-Term Evolution) is expected to be available within three years to provide speeds that can reach 100 Mbps.¹³ WiMAX is another solution that has become available recently.^{14,15} WiMAX is mainly designed for fixed and mobile data transmission. It can provide up to 75 Mbps data rates on both directions and 15 Mbps for mobile users. These recent developments in mobile communication technologies gives us the opportunity to develop a far more reliable remote system that will be able to transmit higher quality images for remote medical services using high-quality real-time video and image transmission.

17.3 Implementation of Mobile Health Care Networks

Most mobile health care networks consist of a mobile device linked via the public mobile networks to a stationary computer at a hospital.¹⁶ Sharing the network with other subscribers can have negative impact on the transmission of medical data. For instance, if the network is congested, the chances of packet dropping, delay, and reductions in the bandwidth can increase.^{7,8} Other concerns about the use of public mobile networks for health care applications are the security and authentication issues. This can be avoided within the existing wireless packet data technologies by using tunneling mechanisms, i. e., establishing VPNs within UMTS cellular systems.¹⁷ VPNs have been used traditionally in wireline networking technologies. Currently, VPN is widely used over the Internet, allowing reliable, secure remote access to the users of private networks. In the case of wireless networks, mobile VPNs (MVPNs) can be used to add extra immunity to the links.¹⁸ MVPNs provide the health care sector with constant media-independent connectivity to hospital

sites. As with most businesses, the use of MVPNs will have the following advantages for health care establishments:

- Improved global connectivity and better reliability, while having capabilities such as secure extranet communications.
- Mobile remote access for transferring medical data can be outsourced, thus eliminating the costs of purchasing and supporting the infrastructure while maintaining full control authentication and security.
- The public Internet can be controlled to provide anytime/anywhere secure connectivity to remote medical personnel, devices, etc. By providing constant connectivity, real-time medical information can be shared, which will improve health care performance.
- Different protocols can be used to implement MVPN IP security (IPSec) protocols, Multi-Protocol Label Switching (MPLS), and Layer Two Tunneling Protocol (L2TP). IPSec-based firewalls are the most popular systems used and are appropriate to meet most sector needs.

In recent years wireless networks have begun to offer end-to-end IPSec non-transparent GPRS/UMTS access, which can be utilized effectively in reinforcing mobile health care networks.¹⁸ In non-transparent access service operation, the GPRS or UMTS network needed to allocate access point name (APN) network identifiers to companies. These APNs are employed by the serving GPRS support node (SGSN) to assign the gateway GPRS support node (GGSN) to be allocated for a number of mobile users within a company as well as being used for billing. The GGSN verifies the IP addresses of the GGSNs to which mobile users will be linked. Another VPN tunnel established over the public Internet ensures a secure route between the GGSN and the health care network (hospital network). In this way, the health care sector can set a mobile network with secure mobile access to private medical data and applications. This can be very beneficial to build a complete health care network for the use of many health care applications.

Figure 17.1 illustrates how a mobile health care system could develop. The establishment of a mobile health care network will enable an extra degree of freedom for health care staff by allowing remote access of medical data anytime/anywhere.⁵ The solution makes it possible to perform useful tasks remotely such as:

- Remote consultation/diagnoses
- Patient monitoring
- Accessing of patient data
- Health care education
- Healthcare management

Access can be achieved within the public mobile network coverage, i.e., nationwide, as if it is happening locally within the hospital site. As a result both medical

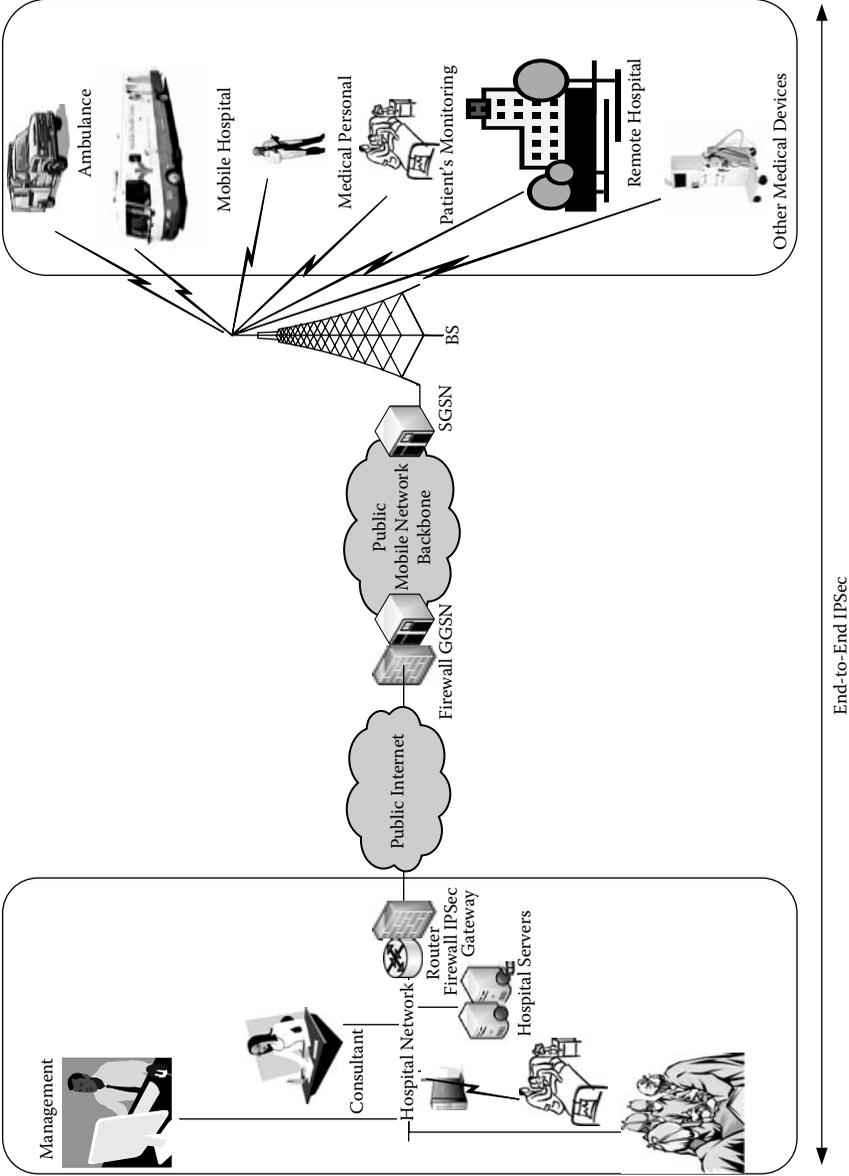


Figure 17.1 Illustration of how mobile healthcare might look.

staff and patients will benefit. Medical staff can perform tasks quicker with less risk to the patient as well as staying in contact with specialist colleagues in the hospital. Patients can expect a higher quality of treatment delivered as quickly as possible when they are in a critical, life-threatening condition.

17.4 Conclusions and Outlook

The idea of implementing telemedicine over mobile networks (also known as mobile health care) has emerged alongside the deployment of mobile networks. However, mobile health care development has been limited by narrow mobile network bandwidth capabilities, thus only low-rate medical data could be handled by 2G and 2.5G mobile links.

With the launching of the first mobile broadband networks, UMTS, mobile health care has become possible by allowing high-rate medical data and images to be transmitted over 3G mobile links. Moreover, ambitious mobile health care applications have emerged such as the mobile tele-echography system.¹⁹ 3G mobile technology was only the beginning of other mobile broadband technologies (3.5G, 3.75G, and WiMAX), which boost both bandwidth and performance, enabling mobile health care networks to be a superior solution for the health care sector.

This chapter highlights the possibility of constructing a mobile health care network within existing mobile network infrastructures. Using the VPN mechanism we can assure integrity, security, and confidentiality of patient health information. Such a solution will allow a fast and cheap deployment of mobile health care networks, and therefore, improve the quality of health care services.

References

1. Z. Hunaiti, A. Rahman, Z. Huneiti, and W. Balachandran, Evaluating the Usage of Wireless Broadband Hotspots, The 2nd International Conference on E-Business and Telecommunication Networks (ICETE), Reading, U.K., 3–7 October 2005.
2. D. Renaudeau, D. Boettle, and H. Steyaert, WiMAX: From fixed wireless access to Internet in the pocket, *Alcatel Telecommunications Review*, 2nd Quarter, 2005.
3. Z. Hunaiti, A. Rahman, Z. Huneiti, and W. Balachandran, Mobile Medical Data Access System, The 15th International IEEE Conference on Electronics, Communications, and Computers (CONIELECOMP 2005), Puebla, Mexico, February/March 2005.
4. C. Pattichis, E. Kyriacou, S. Voskarides, and R.S.H. Istepanian, Wireless telemedicine systems: An overview, *IEEE Antennas and Propagation*, 44, 2, 143–153, 2002.
5. Z. Hunaiti, A. Rahman, Z. Huneiti, and W. Balachandran, 3G Mobile Health System, The 2nd International Conference on Cybernetics and Information Technologies, Systems and Applications (CITSA 2005), Orlando, Florida, 14–17 July 2005.

6. M. Ackerman, R. Craft, F. Ferrante, M. Kratz, S. Mandil, and H. Sapci, Telemedicine technology, *Telemedicine Journal and e-Health*, 8, 1, 71–78, 2002.
7. Z. Hunaiti, V. Garaj, W. Balachandran, and F. Cecelja, An Assessment of GSM/GPRS Link in a Navigation System for Visually Impaired Pedestrians, The 14th International IEEE Conference on Electronics, Communications, and Computers (CONIELECOMP 2004), Veracruz, Mexico, February 2004.
8. Z. Hunaiti, V. Garaj, W. Balachandran, and F. Cecelja, An Assessment of 3G Link in a Navigation System for Visually Impaired Pedestrians, The 15th International IEEE Conference on Electronics, Communications, and Computers (CONIELECOMP 2005), Puebla, Mexico, February/March 2005.
9. Z. Hunaiti, V. Garaj, and W. Balachandran, Assessment of the video image quality in a remote vision guidance system for visually impaired pedestrians, *Journal of Telemedicine and Telecare*, 12: 400–403, 2006.
10. B. Konstantinos, P. Konstantinos, T. Sapal, and K. Dimitrios, Use of 3G mobile phone links for teleconsultation between a moving ambulance and a hospital base station, *Journal of Telemedicine and Telecare*, 12, 1, 23–26, 2006.
11. 3GPP TS 25.308, High Speed Downlink Packet Access (HSDPA), overall description, September 2004.
12. 3G Americas, Mobile Broadband: The Global Evolution of UMTS/HSPA 3GPP Release 7 and Beyond, July 2006.
13. 3GPP, TR 25.913, Requirements for Evolved UTRA and Evolved UTRAN, www.3gpp.org, 2005.
14. Intel White Paper, IEEE 802.16 and WiMAX: Broadband Wireless Access for Everyone, Intel, 2003.
15. Mobile WiMAX, The Best Personal Broadband Experience! WiMAX Forum, June 2006.
16. B. Woodward, R.S.H. Istepanian and C.I. Richards, Design of a telemedical system using a mobile telephone, *IEEE Transactions on Information Technology in Biomedicine*, 5, 1, 13–15, 2001.
17. A. Shneyderman, A. Bagasrawala, and A. Casati, Mobile VPNs for Next Generation GPRS and UMTS Networks, white paper, Lucent Technologies, Inc., 2000.
18. C. Andersson, *GPRS and 3G Wireless Applications*, John Wiley & Sons, 2001.
19. S. Garawi, R.S.H. Istepanian, and M.A. Abu-Rgheff, 3G wireless communications for mobile robotic tele-ultrasonography systems, *IEEE Communications Magazine*, 44, 4, 91–96, 2006.

**CHALLENGES
AND
OPPORTUNITIES**

6

Chapter 18

Telemedicine Research: Opportunities and Challenges

Pennie S. Seibert, Tiffany A. Whitmore, Carin M. Patterson, Caitlin C. Otto, Patrick D. Parker, Nichole Whitener, Michael J. Ward, Jean Basom, and Christian G. Zimmerman

CONTENTS

18.1 I ntroduction.....	350
18.2 R esearch Methods.....	351
18.3 R ecrutment.....	352
18.4 P artnering Institutions.....	353
18.5 I nstitutional Review Boards.....	354
18.6 E t hical Issues.....	355
18.6.1 R isk Management.....	355
18.6.2 A ccuracy.....	356
18.7 L egal Issues.....	356
18.7.1 C onfidentiality.....	356
18.7.2 L icensing.....	357
18.7.3 I nformed Consent.....	358

- 18.8 Perceptions and Satisfaction.....359
- 18.9 Opportunities in Research 360
 - 18.9.1 Design..... 360
 - 18.9.2 Longevity of Program361
 - 18.9.3 Rural Areas362
- 18.10 Conclusion363
- References363

Telemedicine is emerging as an effective tool that can potentially enhance the ability to provide quality health care in hospitals, professional offices, and homes. Applications have demonstrated success in a wide range of disciplines and are particularly important for rural communities in hopes of advancing health care practices to these areas that may not have the same advantages, resources, or specialists. Research into this exciting field has led to challenges rather unique to telemedicine. The demands for appropriate research methods and tools for each application have become more complex because of the variety of fields involved and issues specific to telemedicine.

Telemedicine research commonly encompasses various sites and review boards, which introduces more challenges. Adding to the complexity of telemedicine research are the ethical and legal issues associated with the use of technology, enrolling participants, transmitting data, confidentiality, and licensing. Illuminating perceptions and level of satisfaction could help identify methods of facilitating acceptance of telemedicine and expedite the transition from using solely traditional methods of health care to including telemedicine. Despite all the challenges, telemedicine presents many opportunities and brings such promise to the increasing shortage of health care providers. Accordingly, it is vital for researchers to address and overcome these challenges.

18.1 Introduction

There are many ways technology can supplement health care to provide the best service to a dispersed population. This application of technology is referred to as telemedicine, telehealth, or e-health. The development of telemedicine is driven by the increasing need to provide medical services to remote and underserved areas.

Numerous health inequalities exist for people residing in rural and underserved locations, including limited access to health services and its related decrease in overall health.¹⁻³ Rural residents potentially experience higher rates of preventable risk factors for disease, such as obesity, smoking, poor diet, and lower rates of activity than people living in urban areas.^{1,2} Socioeconomic variables also contribute to health differences, as people in rural communities are more likely to be uninsured, and to have lower education and income levels.^{1,2} Further complicating rural health

are the numerous challenges for health care delivery in rural settings. Rural residents face limited access to providers¹⁻⁴ and specialists are rarely available in rural settings. Adequate rural health care services are difficult to maintain for a variety of reasons, including difficulties recruiting and retaining specialists as well as low patient volumes.⁴⁻⁷ These factors result in people having to travel to urban areas to receive specialty care along with increased transports to larger facilities in emergent/critical cases. Travel is difficult for the elderly⁸ and for people with chronic health conditions. The length of travel time required for people to obtain a specialist frequently hinders adequate disease management.⁹ In emergent cases, such as stroke, valuable time for life-saving interventional treatments is lost during a assessment and transport. Also, research indicates that many rural hospitals lack the resources necessary to facilitate best practice guidelines.^{10,11} These issues underscore the importance of researching ways to provide better access to health services in rural locations. While there is potential for telemedicine and telehealth programs, many challenges exist for developing and implementing these types of programs. Accordingly, further research is needed to investigate the viability and outcomes of varied applications.¹²

Challenges also exist for implementing research projects. Current problems in telemedicine research can be identified in methods, recruitment, reporting ethical standards, confidentiality, licensing, and the perception and satisfaction of patients and practitioners. Additionally, cost effectiveness, safety, and accuracy are difficult to analyze. Opportunities for research lay in the numerous applications of telemedicine. Researchers need to consider these aspects as they pursue the investigation into this exciting field.

18.2 Research Methods

The versatility of the technology provides opportunities for implementation of telemedicine into almost every facet of health care. Thus, researchers need to consider both the specialty of interest and the factors associated with varied applications. It is important to address the accuracy of equipment, outcomes obtained, acceptance, cost-effectiveness, and satisfaction. Researchers from a wide range of fields (e.g., psychological science, health sciences, law, human factors, business, marketing, computer science, engineering, etc.)¹³ need to contribute to advancing knowledge in telemedicine. Previously established methods in each field could be utilized to ensure accurate and reliable results. Additionally, for telemedicine to progress, novel interdisciplinary research methods need to be developed for specific areas of interest within telemedicine.¹⁴

A recent article, for example, described the use of telemedicine in pathology for intraoperative diagnosis over the course of four years.¹⁵ As illustrated by the authors, further research in this area would need to account for a robotic telepathology microscope, its ability to allow diagnostic accuracy by pathologists, the time required for diagnosis as a whole, as well as specific processes (cutting, staining, and

other smear preparations). Thus, it would be useful to analyze problems resulting from the particular telepathology equipment in conjunction with traditional diagnostic methods.

Each of these aspects needs to be addressed individually using appropriate designs from the respective fields. The mechanical and esthetic design of the microscope would best be investigated using engineering techniques. On the other hand, the comparison of diagnostic methods between traditional and telemedical applications would be better assessed by pre-test/post-test or experimental designs. Relevant research techniques need to be employed to ensure valid and reliable findings. Indeed, issues such as these demonstrate the critical need for novel interdisciplinary approaches to advance telemedicine.

Consistent and comparable findings can be obtained by using validated techniques and questionnaires. Standardized questionnaires are frequently not available for telemedicine research.¹⁴ According to the American Telemedicine Association, the best way to construct a survey is to look for standards in the field of interest.¹⁴ If the questionnaire is addressing cost analysis, a marketing research model could be most beneficial.¹⁴ However, there are times when precedence cannot be used. For example, a study by Tudiver et al. identifies the lack of a valid survey pertaining to the primary care provider's attitudes and satisfaction.¹⁶ Therefore, they developed a unique survey assuring content validity and reliability by testing for face and content validity.¹⁶ To validate the study, they used factor analysis and internal consistency reliability (Cronbach's alpha).¹⁶ This exemplifies the point that it is important for researchers to identify and describe newly designed questionnaires to be properly identified as valid and reliable.¹⁶

Telehealth is going to transform health care, from the way it is delivered to the way it is paid for.¹⁴ Although telemedicine is an emerging field, researchers must be mindful of the historically validated methods. Though the face of health care is constantly changing, research would benefit from using established methods while simultaneously employing novel, innovative applications.

18.3 Recruitment

Enrollment is a major challenge in telemedicine research. Rates of refusal to join a study have been reported as high as 75%.¹⁷ Reviews show over one third of research studies are reporting sample sizes of 22 participants or less.¹⁸ Consequently, the findings from small unrandomized samples are difficult to validate.¹⁸ Without a large enough sample size and a truly representative group, results are even more difficult to generalize to the population. It is also the case that when conducting research in rural areas, the limited population base is not conducive to obtaining large sample sizes. There are not enough people available to participate.¹⁹

One study attempted to analyze why people decline to participate. According to a survey given to those who refused to participate, the top reasons included being

uncomfortable with technology, being too busy, and the belief that technology would not help them. This information has limitations because only one third of those who refused to participate answered the questionnaire, most all of whom were rural candidates.²⁰ However, these factors can be used to emphasize the importance of education and user friendly technology in designing and assessing telemedicine programs and applications.

In another study, non-participants reported the lack of perceived benefit, existing health care routines being sufficient, the equipment would be a burden, and unwillingness to become involved in anything else as the major reasons to refuse participation.¹⁷ This study, involving a metropolitan community, used an open-ended question and then categorized the responses. Possibly, respondents would not have answered in the broad categories if given more limited choices. Nevertheless, the differences between rural and urban populations can be used to adjust each study to the target population appropriately.

Other factors contribute to enrollment rates. By using multivariate analysis, Palmas et al. reported rural participants having many independently associated variables. Younger age, male gender, better health status, and knowing how to use a computer, among other factors, were associated with enrollment.²⁰ It seems in this study, that only the more healthy and knowledgeable group is evaluated. This may not create a clear representation of the entire population.

Additionally, recruiting could be improved by increasing support for the project. Utilizing physicians who support telemedicine may improve the acceptance rate. A physician's endorsement of a program potentially increases patient's respect and willingness to try the equipment.¹⁷ Physicians' personal preferences also impact their willingness to endorse a telemedicine program. For instance, a physician may believe face-to-face patient meetings are necessary. Indeed, some physicians insist that there is something special about in-person interactions that cannot be captured from a distance.¹⁸ Many physicians share the general population's distrust of technology. Some physicians and other care givers express concerns that telemedicine is a way to foster "big brother" watching over them. Clearly, further research is needed to assess the effects of physician and other care giver endorsement on enrollment rates.

Lastly, telemedicine technologies need to become more user friendly. By implementing innovative ideas like robotic pets, patients who are unfamiliar with technology could become more comfortable and more likely to utilize it. We suspect that as the technologically savvy generation ages, telemedicine and research into telemedicine applications will be more widely accepted.

18.4 Partnering Institutions

Telemedicine projects typically involve multiple research sites. Regardless of whether a telemedicine study focuses on home health care or delivery of health services to remote areas, the distance between sites presents unique challenges to

the field of telemedicine research. Further, challenges arise in developing research protocols that account for the locations' varying needs. For example, if a major trauma center partners with a rural critical access hospital on a telemedicine project, protocols have to be designed to work across facilities. This means that a great deal of collaboration and trust is necessary between institutions throughout the development and implementation of the research project. Procedures for scheduling health services, documentation of processes, clinical documentation systems, study enrollment procedures, and staffing availability to assist with the research are just a few of the issues that have to be addressed to ensure that the research protocol is compatible with the policies and procedures across research sites.

18.5 Institutional Review Boards

Once research protocols are written, institutional review board (IRB) approval must be obtained prior to initiating the research. Any institution engaged in human subject research must designate an IRB to review, approve, and monitor the research. While many larger health care institutions have an IRB, smaller facilities rarely have an IRB of record. This presents a unique challenge for partnering facilities when establishing a program of research, as an IRB must be designated to review the research for each facility. If one facility has an IRB, the institution can choose to serve as the IRB of record for all partnering institutions, or any institution without an IRB may choose to secure the services of an independent IRB.

For one institution's IRB to serve as the IRB of record, partnering sites must enter into a legal agreement. This means that one facility assumes the responsibility for ensuring that the research is scientific, ethical, and meets regulatory standards at all institutions. One institution may be hesitant to assume this responsibility for all facilities, particularly in studies that include higher levels of risk for patients. In telemedicine research, where the distance between partnering institutions may be great, an IRB may be even more hesitant as oversight may be more difficult.

If an institution elects to use the services of an independent IRB, an additional set of expenses may be acquired. IRBs typically charge a one-time review fee (approximately \$2000); however most independent IRBs charge per review. These charges may result in multiple institutions paying fees for the review of one protocol if different IRBs are used (e.g., institutional versus independent). Further, when multiple IRBs review the same protocol, one IRB may request changes to the protocol that differ from another IRB's requests. As the protocol must be the same across institutions, the protocol might be modified multiple times to obtain all IRBs' approvals for the same document. Not only may this process be expensive, but lengthy as well.

In cases of telemedicine research projects funded by the U.S. Department of Health and Human Services (HHS), an additional review by the Office for Human Research Protections is necessary once all local IRB approvals are obtained. If the

Office for Human Research Protections requests changes, the protocols must be resubmitted for approval to the local IRBs. This once again extends the time and expense necessary to initiate the program of research.

While it is ideal for partnering institutions to designate one IRB to review the research, challenges may be present in determining which IRB will conduct the review. This is particularly true in cases where more than one partnering facility has an IRB, or when an IRB is not willing to undertake the responsibility to review the research for the other institutions. As approval is needed for all human subject research, and legal or financial agreements may be needed prior to review, it is important for an IRB to be designated early in the development of the telemedicine project.

18.6 Ethical Issues

Research ethics is a key component in the design of a research project. An IRB is charged with assuring that ethical standards are maintained in the design and implementation of a project. HHS has formulated guidelines for IRBs under the Code of Federal Regulations, Title 45: Public Welfare; Part 46: Protection of Human Subjects.²¹ As with all other research programs, telemedicine programs are subject to complying with HIPAA regulations in dealing with human participants. Maintaining these standards presents unique challenges in the research of telemedicine. In a literature review by Marziali et al., 99% of the telemedicine articles surveyed did not sufficiently report the use of research ethics.²² This is not to imply that ethical research is not occurring; it simply may not have been adequately reported. In addition, there are no clear standards specific to telemedicine. Telemedicine has to meet the same standards as all other types of research, but some grey areas are left to IRB interpretation. That interpretation may vary across IRBs.

18.6.1 Risk Management

Telemedicine research groups must be particularly cautious about minimizing risks or attributing reasonable risks in relation to benefits. The safety and accuracy of the technology must be carefully analyzed before any human participants are involved. Careful investigation into the technology used for teleconsultation is needed so that patients do not suffer harm.²¹ For example, the use of robotics to perform surgeries presents a vulnerable situation. In one study, two out of 128 various surgeries performed with the da Vinci™ robotic system failed because of technical problems.²³ Researchers must consider how to handle the situation in the event something does go wrong. Prior to the procedure, patients must be fully informed of the risks and also how they will be compensated if there is a complication. A clear emergency plan should be designed to ensure the protection of the patient. This may or may

not differ from standard emergency protocols, depending on the form of telemedicine being implemented.

Serious consideration to risk assessment is not only essential for high risk procedures, but also for telehealth, in-home applications. There is an increasing need for home health care systems as the population ages and serious health problems require a greater need for temporary and permanent home care services.²⁴ With this sort of research, there are still risks involved that need to be analyzed properly. As with all electronics, technical failures can occur. For example, an electronic blood pressure cuff could malfunction and possibly remain inflated, distressing or injuring the patient. The European Union has strict quality control standards for the medical devices that are used in the home of a patient communicating with their health care provider through telemedicine. Despite these strict quality control standards, it is still possible for patients to suffer harm due to a technical failure.²⁵

18.6.2 Accuracy

The accuracy of the telemedicine equipment needs to be scrutinized carefully by the research team. The equipment needs to be tested against traditional measures and possibly calibrated for optimum accuracy. One study examined the differences between traditional colposcopy and telecolposcopy by analyzing the discrepancies between distant and site evaluations.²⁶ They found that the accuracy for diagnosis was maintained, but the examination was not adequate using telecolposcopy alone.²⁶ This study and others similar in nature illuminate the importance in analyzing the applicability and accuracy of the device. However, it can be a challenge to determine the accuracy of the equipment when the abilities of the specialists and local health care providers vary.

The accuracy of a device is commonly described by the company that designs and produces the system. However, no independent studies comparing the accuracy of similar systems from different manufacturers were found. This creates a challenge for designing and implementing research studies and discerning which product is most effective and appropriate for a particular study.

18.7 Legal Issues

18.7.1 Confidentiality

Confidentiality is both a legal and an ethical issue in the field of telemedicine. Implementing telemedicine programs introduces new liability issues regarding confidentiality. It is the duty of medical support staff to maintain complete confidentiality for their patients. The issue of maintaining confidentiality is especially important to address when developing research through telemedicine applications.

Medical information has traditionally been kept in paper files in the offices of health care professionals. Although patient information has never been completely confidential, the difficulty of access to paper records often protected patients' privacy.²⁷ Currently, new technology allows physicians and researchers to gather and store medical records through electronic means. This ease of access creates vulnerability to patients' confidentiality.²⁸

Patient confidentiality is a primary concern in medical research. Because there is no current consensus for maintaining confidentiality in telemedicine research, there are several concerns that need to be addressed for a smooth integration of telemedicine systems. It is mandatory for telemedicine practices to adhere to HIPAA privacy standards, as well as standards set by the AMA for maintaining patient confidentiality in the clinic and research environments. However, there is currently no protocol for maintaining confidentiality during doctor–patient communication via e-mail, and research via telecommunication applications. Therefore, a necessary next step is the development of standard protocols for telecommunication systems. Some of the following areas will need to be addressed for this protocol:

- How will confidentiality be maintained when communicating with patients via e-mail?
- How will confidentiality be maintained when communicating with patients via teleconferencing?
- Where and how will the data be stored?
- Who can access this data?
- How is patient information secured through the network?
- Does information technology support staff need to sign a confidentiality non-disclosure form if they are going to be in contact with the patient or patient's data?
- Are the measures being taken to ensure confidentiality?

Once these protocols to maintain confidentiality have been established, support from the legislative system is needed to ensure accountability for maintaining confidentiality.²⁹ There is no current, adequate legislation that applies specifically to managing confidentiality of electronic medical records.²⁸ In fact, only about half the states within the United States have a law that prohibits disclosure of health information without patient authorization, and even some of these laws are ambiguous.²⁷

18.7.2 Licensing

Issues of licensing and malpractice are of more concern in the implementation of telemedicine and have to be taken into account when designing research programs within telemedicine. In the United States, each state has its own statutes for medical licensing and practicing. Telemedicine brings forth an interesting issue with cross-border consultations.

Currently, if a physician consults with a patient in a certain state, he or she is legally required to be licensed to practice in that state. With the current protocol of medical practices, it is assumed that the licensed physician only practices in the state he or she is licensed in. With the implementation of telemedicine programs, the lines of licensing are blurred.

Another issue related to licensing is malpractice insurance. Presently, malpractice insurance only covers physicians in the state in which they are licensed to practice. If a physician then consults with a patient across a state border, the physician may risk being charged with practicing without insurance. This presents a complicated situation for telemedicine because each state may not have the specialist needed for every case. Telemedicine has the potential to improve the accessibility for specialists in rural areas if this issue were resolved. Some states have passed legislation that requires physicians to be fully licensed to practice in that state if they wish to consult a patient via electronic media.³⁰ One solution would be a national licensing for practitioners in telecommunications. In the European Union, health care providers are licensed in all participating states.³⁰ Some legislation is being developed to allow a license in teleconsultation. This way the practitioner would have a national license, and could consult across state borders legally and remain insured.

18.7.3 Informed Consent

The following three issues should be addressed when approaching the subject of consent related to research in telemedicine:

1. What is a valid consent?
2. What are the parameters for valid consent?
- 3 . Who defined these requirements?

To gain a valid authorization to proceed with a research study the patient must be informed of his or her diagnosis, the nature and purpose of the proposed procedure, alternatives to the procedure, and the risks and benefits of the proposed procedure, alternative treatments, and refusing treatment.³¹ Also, according to the HHS's Basic Policy for Protection of Human Research Subjects, the subjects must be informed of how their confidentiality will be maintained.²¹

The aspect of obtaining informed consent in research is extremely important as it relates to telemedicine. Informed consent documents will need to be adapted for use in telemedicine. The following list addresses some of the particular aspects of telemedicine that the patients or participants will need to be informed of, if they are to participate in the research study³²:

- Who will be at the other end of a videoconference?
- What is their role in the consult?
- What are their credentials?

- If the videoconference is recorded, how will the information from it be used (i.e., clinical audit or research purposes)?
- How will their confidentiality be maintained via telemedicine services?

The patient will need to be informed of these and other variables specific to telemedicine in addition to the normal variables of obtaining informed consent prior to the research group creating an electronic file.³² This raises a question as to the patient's knowledge of information that is transferred over the Internet and whether or not the patient is aware that his or her privacy is at risk because of this transfer. Safeguards should be implemented and described in research reports. Some of the most common Internet safety devices are firewalls, passwords, and encrypted Websites.²²

18.8 Perceptions and Satisfaction

A predominant theme within telemedicine research is the investigation of perceptions and satisfactions with programs and devices. Identifying perceptions could help identify methods of increasing acceptance of telemedicine and expedite the transition from using solely traditional methods of health care to including telemedicine. Perceptions of telemedicine can also have a direct benefit to research. Many studies within the literature report low enrollment numbers or problems with gaining participants for their studies. Hopefully, by assessing perceptions of participants, researchers can develop programs in which participants will be more willing to enroll. One obvious drawback is researchers can only gain perceptions of those who have already agreed to participate. Perceptions of those who are not willing to participate are critical to obtain a full understanding of the issues involved.

Several models have been constructed to identify perceptions and satisfaction as an essential element of successful telemedicine programs.^{18,33,34} In fact, it has been proposed that resistance of staff could have a very large impact on program failures.³⁵ A review by Jennett et al. suggests resistance to telemedicine programs may stem from a lack of readiness to take on new procedures or new programs.³⁵

Contrasting data has been found in the literature regarding patients', physicians', and staff's perceptions of telemedicine and also with their level of satisfaction. Examples of findings include patient preference both for and against telemedicine, while physicians within the same study preferred face-to-face meetings in clinical practice.¹⁸ Patient satisfaction was attributed to such variables as quality service, convenience, communication, and human connection through videoconferencing.^{18,36} Physicians' satisfaction can possibly be attributed to their exposure to both face-to-face meetings and videoconferences. One article even reported prisoners preferring the use of telemedicine over leaving the institution.¹⁸

18.9 Opportunities in Research

18.9.1 Design

A major issue in telemedicine design is who should lead programs or who is best able to guide telemedicine practices for proper health care. Possible candidates that have been suggested are administrative liaisons,³³ physicians,³⁷ or telemedicine “champions.”³⁸ Research has already begun to identify models for the implementation and integration of telemedicine. This potentiates the opportunity to identify successful components or models of telemedicine design in an effort to develop consistent practices and standards for telemedicine.

The title of administrative liaison does not necessarily refer to someone of administrative stature within the ranks of a hospital or facility. Rather, this refers to the notion that telemedicine incorporates multiple aspects of experience and expertise, ranging from IT personnel to surgeons. For a telemedicine program to reach its full potential, someone needs to assume the role of liaison effectively to ensure channels of communication are open, efficient, and effectual. It is essential to note: the key element to success is appropriate communication.

The article regarding telemedicine “champions” was referring to the use of identifying personnel who were best qualified to lead specific functions of telemedicine programs.³⁸ The example given was the use of a nurse practitioner to lead staff at peripheral sites. Potentially the use of “champions” and an administrative liaison could be combined by appointing a central program leader while delineating duties to those with specific education and skills as pertaining to various facets of the program. Nevertheless, one dilemma with this has been identified. In the case of one rural telemedicine program, the “champion” was the only clinical provider for an entire island in Michigan.³⁹ Thus, an overabundance of duties were imposed on one individual.

An article by Yellowlees takes an entirely different approach to leadership for telemedicine programs.³⁷ In this case it is suggested that physicians and clinical practitioners direct clinical telemedicine programs. Yellowlees states, “It is crucial that clinicians continue to ‘own’ telemedicine systems and to be involved in all stages of planning, implementation, and evaluation.”³⁷ The claim is that many bureaucratically driven programs do not have the insight to apply programs to the “real clinical world,” and further, that telemedicine should not be so concerned with confidentiality and legality issues.³⁴ This perspective is contradictory to most of the literature considering confidentiality and legality issues. In addition, Yellowlees claims experiential learning by clinicians could have benefits to telemedicine design.³⁸

As already stated, research is currently identifying models for the implementation and integration of telemedicine. Even within these, differences of opinion still exist. Jennett et al. reviewed models of readiness, wherein success of a program depends on modifying established procedures prior to implementation in order to accept telemedicine.³⁵ On the contrary, others suggest shaping telemedicine programs and

adapting equipment to existing clinical programs.³⁷ As suggested before, the integration of telemedicine may differ across particular areas of telemedicine.

By identifying or constructing successful models, research has the potential to improve telemedicine practices as well as create consistency within these practices. A lack of consistent training of telemedicine program staff has been reported as a problem by several sources.^{37,38,40} This poses threats for the accuracy, validity, and the ability to generalize research findings in telemedicine. Combining inconsistencies in training and practice with already-presented inconsistencies in legislation and regulation, along with confounding theories in literature results in a very complex myriad of challenges for research to first untangle and then address.

18.9.2 Longevity of Program

Issues of funding and lack of consistency of design have lead to a short life span for many telemedicine programs. Reportedly, most programs have experienced problems with continued funding beyond two to three years after implementation of the program.^{18,37} This is detrimental to telemedicine programs, especially with one study reporting “no successful clinical based or official home health teleconsultations were conducted during the course of the grant period” of only one year.³⁹

Accordingly, long-term benefits of health care and quality of life for telemedicine patients have not been determined.³³ Though a call exists for longitudinal studies to assess long-term benefits more accurately,¹⁸ this is difficult to do given the prevailing norm of short life spans observed in many telemedicine programs. Turn-over rates for project staff have been as high as 100% within a single year.³⁹ Continued education designed to combat attrition rates³⁸ and to resolve inconsistencies in program procedures could help strengthen longevity of telemedicine programs. Also, enhancements within technology and telemedicine equipment can have the effect of a double-edged sword. On the one hand, the equipment becomes more user friendly and efficient, while on the other hand there is the quandary that when much-needed upgrading occurs, existing equipment may be rendered obsolete and outdated. This creates further challenges for budgeting and funding.

Another possible factor related to the duration of telemedicine programs is communication between IT staff and others in the interdisciplinary team. Complaints regarding IT communication styles have long been an issue as technology becomes an integral part of life. It is common to hear complaints that IT folks speak a different language that no one can understand and further that those within IT have difficulty understanding the language of those without benefit of experience in the IT field. Faulty, outdated, or misunderstood information from IT departments can cause delays,³⁹ wasting valuable time because of poor communication.³⁷ As already stated in this chapter, design models are being developed with communication issues in mind. These models need to account for the IT versus non-IT communication problems.

Finally, funding and re-funding, sustainability, and medical reimbursement issues must be addressed for telemedicine to experience longevity. Unfortunately, many telemedicine projects have remained incomplete because funding allowed the inception but was insufficient to complete the project. These kinds of failures contribute to disparaging views about the potential for telemedicine. Without a doubt, up-front costs to establish telemedicine programs are very high while the benefits are not as clear. The paucity of information regarding cost-balance effectiveness of the investment is also of concern. Establishing an effective, viable telemedicine program requires long-term commitment. This is difficult to obtain when the current research data cannot ensure the original investments will pay off in the end.

18.9.3 Rural Areas

The use of telemedicine in rural or remote areas is a topic of great interest and can be found throughout the literature.^{33,35,41–45} The hope is to advance health care practices in areas that may not have the same advantages, resources, or specialists. Two questions rise out of this:

1. Do the benefits of telemedicine in rural areas outweigh the costs?
2. Are telemedicine practices in rural areas comparable to more traditional methods?

Literature reports positive benefits of telemedicine such as saving patients' time and money spent on travel to larger cities or travel to more equipped facilities.³³ This has the potential to make health care more accessible in rural areas, and decrease costs and time for patients. Also, it is possible that telemedicine could provide an avenue for patients who need medical assistance and would not have otherwise sought it to benefit because of improved access.⁶ Telemedicine could also be implemented as a tool for physicians and care givers in remote areas to continue education as is critical for accreditation and staying current with the field.³⁸ However, there is clearly a time burden associated with establishing a telemedicine program. Indeed, it is often far greater than expected.⁴⁶ Several articles have reported telemedicine adding additional work on top of already-busy clinical schedules rather than in place of clinical work.^{37–39}

A major concern of telemedicine, and a large opportunity for research, is quality of care. Obviously the health care community does not want to implement a tool that will decrease standards of care in any domain. A recent article by Hailey states that even if telemedicine does have weakness, it could still be a useful instrument of health care. Along with the advantages, the disadvantages need to be assessed thoroughly.³³

There are also suggestions that directing telemedicine towards rural areas has significant drawbacks. More precisely, those who could receive the greatest benefits

from telemedicine may have the greatest barriers to its implementation.³⁸ These communities often are associated with a lack of technology, finances, resources, and unfavorable perceptions that can impact telemedicine success. Perhaps this is why Yellowlees suggested that developing telemedicine programs may be more reasonable in areas with more expertise or areas that are more equipped to take on the burdens of creating a program.³⁷

18.10 Conclusion

Opportunities for research lay in the numerous applications of telemedicine. However, research into this exciting field has led to a host of challenges unique to telemedicine. The demand for appropriate research methods and tools for each application have become more complex because of the variety of fields and technology specific to the various fields. Striking communication and education challenges are also apparent. Person–technology interactions are teeming with problems to be resolved, ranging from fears of the unknown to fears related to actual adverse events. Cost–benefit analyses often fail to demonstrate telemedicine success. Issues regarding safety, ethics, and confidentiality have yet to be resolved. Research design and implementation will need to combine existing formats with novel ones. Yes, the challenges are humbling. Nevertheless, we are steadfast in our belief that telemedicine is certainly the wave of the future and has the potential of providing efficient, effective, and humane health care to those who are currently underserved. Moreover, we are convinced that telemedicine can enhance medical practice for those in both rural and urban settings.

References

1. Eberhardt MS, Pumuk ER. The importance of place of residence: examining health in rural and nonrural areas. *Am J Pub Health*, 94: 1682–1686, 2004.
2. Hartley D. Rural health disparities, population, and rural culture. *Am J Pub Health*, 94: 1675–1678, 2004.
3. Johnson ME, Brems C, Warner TD, Roberts LW. Rural-urban health care provider disparities in Alaska and New Mexico. *Adm Policy Ment Health*, 33: 504–507, 2006.
4. Nesbitt TS, Marsin JP, Daschbach MM, Cole SL. Perceptions of local health care quality in 7 rural communities with telemedicine. *J Rural Health*, 21: 79–85, 2005.
5. Ricketts TC. The changing nature of rural health care. *Annu Rev Public Health*, 21: 639–657, 2000.
6. Hart LG, Salsberg E, Phillips DM, Lishner DM. Rural health care providers in the United States. *J Rural Health*, 18: 211–32, 2002.
7. Seibert PS, Whitmore TA, Parker PD, Payne K, Grimsley FP, O'Donnell JE. The emerging role of telemedicine in diagnosing and treating sleep disorders. *J Telemed Telecare*, 12: 379–381, 2006.

8. Goins RT, Williams KA, Carter MW, Spencer SM, Solovieva T. Perceived barriers to health care access among rural older adults: a qualitative study. *J Rural Health*, 21: 206–213, 2005.
9. Iezzoni LI, Killeen MB, O'Day BL. Rural residents with disabilities confront substantial barriers to obtaining primary care. *Health Serv Res*, 41: 1258–1275, 2006.
10. Hess DC, Wang S, Gross H, Nicholas FT, Hall CE, Adams RJ. Telestroke: extending stroke expertise into underserved areas. *Lancet Neurol*, 5: 275–280, 2006.
11. Burgin WS, Staub L, Chan W, Wein H, Felberg RA, Grotta JC, Demchuck AM, Hickenbottom SL, Morgenstern L B. A cute stroke care in non-urban emergency departments. *Am Acad Neurology*, 57: 2006–2012, 2001.
12. Bashur RL, Mandil SH, Shannon GW. Chapter 8: Executive summary. *Telemed J e-Health*, 8: 95–107, 2002.
13. Seibert PS, Patterson CM, Whitmore TA, Parker PD, Whitener N, Zimmerman CG. Opportunities for psychological science in telemedicine. Western Psychological Association; May 3–7, 2007; Vancouver, BC, Canada.
14. Krupinski E, Dimmick S, Grigsby J, Mogel G, Puskin D, Speedie S, Stamm B, Wakefield B, Whited J, Whitten P, Yellowlees P. Research recommendations for the American Telemedicine Association. *Telemed J E Health*, 12(5): 579–589, 2006.
15. Hutarew G, Schlicker HU, Idriceanu C, Strasser F, Dietz O. Four years experience with teleneuropathology. *J Telemed Telecare*, 12(8): 387–391, 2006.
16. Tudiver F, Wolff LT, Morin PC, Teresi J, Palmas W, Starren J, Shea S, Weinstock R. Primary care providers' perceptions of home diabetes telemedicine care in the IDEA-Tel project. *J Rural Health* 23(1): 55–61, 2007.
17. Subramanian U, Hopp F, Lowery J, Woodbridge P, Smith D. Research in home-care telemedicine: challenges in patient recruitment. *Telemed J E-Health*, 10: 155–161, 2004.
18. Thurmond VA, Boyle DK. An integrative review of patients' perceptions regarding telehealth used in their health care. *Online J Knowl Synth Nurs*, 9(2), 2002.
19. Seibert PS, Whitmore TA, Parker PD, Patterson CM, Ward MJ, Basom J. Telemedicine facilitates CHF home health care. American Telemedicine Association 2007: 12th Annual International Meeting; May 13–15, 2007; Nashville, TN.
20. Palmas W, Teresi J, Morin P, Wolff TL, Field L, Eimicke JP, Capps L, Prigollini A, Orbe I, Weinstock RS, Shea S. Recruitment and enrollment of rural and urban medically underserved elderly into a randomized trial of telemedicine case management for diabetes care. *Telemed J E-Health*, 12(5): 601–607, 2006.
21. United States Department of Health and Human Services. OHRP — Code of Federal Regulations, Title 45: Public Welfare; Part 46: Protection of Human Subjects. June 23, 2005. <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.101> Accessed: March 22, 2007.
22. Marziali E, Dergal Serafini JM, McCleary L. A systematic review of practice standards and research ethics in technology-based home health care intervention programs for older adults. *J Aging Health*, 17(6): 679–696, 2005.
23. Bodner J, Augustin F, Wykypiel H, Fish J, Muehlmann G, Wetscher G, Schmid T. The da Vinci robotic system for general surgical applications: a critical interim appraisal. *Swiss Med Wkly*, 135: 674–678, 2005.

24. Lamothe L, Fortin JP, Labbe F, Gagnon MP, Messikh D. Impacts of telehomecare on patients, providers and organizations. *J Telemed E-Health*, 12(3): 363–369, 2006.
25. Stanberry B. Legal and ethical aspects of telemedicine. 4: Products and jurisdictional problems. *J Telemed Telecare*, 4(3): 132–139, 1998.
26. Ferris DG, Macfee MS, Miller JA, Litaker MS, Crawley D, Watson D. The efficacy of telecolposcopy compared with traditional colposcopy. *Obstet Gynecol*, 99(2): 248–254, 2002.
27. United States Department of Health and Human Services. Standards for Privacy of Individually Identifiable Health Information. December 28, 2005. <http://www.hhs.gov/ocr/part1.html>. Accessed: March 22, 2007.
28. Chin JJ. The use of information technology in medicine: defining its role and limitations. *Singapore Med J*, 44(3): 149–151, 2003.
29. Hodge JG, Gostin LO, Jacobson PD. Legal issues concerning electronic health information privacy, quality, and liability. *JAMA*, 282(15): 1466–1471, 1999.
30. Stanberry B. Legal and ethical aspects of telemedicine. *J Telemed Telecare*, 12(4): 166–175, 2006.
31. American Medical Association, Office of the General Counsel, Division of Health Law. AMA (legal issues) informed consent. September 1998. <http://www.ama-assn.org/ama/pub/category/4608.html>. Accessed: March 16, 2007.
32. Stanberry B. Telemedicine: barriers and opportunities in the 21st century. *J Intern Med*, 247: 615–628, 2000.
33. Hailey D. Technology and managed care: is telemedicine the right tool for rural communities? *J Postgrad Med*, 51(4): 275–278, 2005.
34. Jen-Hwa Hu, P. Evaluating telemedicine systems success: a revised model. *Proceedings of the 36th Hawaii International Conference on Systems Sciences*, 2002.
35. Jennett PA, Gagnon MP, Branstadt HK. Preparing for success: readiness models for rural telehealth. *J Postgrad Med*, 51(4): 279–285, 2005.
36. Stevens A, Doidge N, Goldbloom D, Voore P, Farewell J. Pilot study of televideo psychiatric assessments in a underserved community. *Am J Psychiatry*, 156(5): 783–785, 1999.
37. Yellowlees P. Successful development of telemedicine systems — seven core principles. *J Telemed Telecare*, 3(4): 215–252, 1997.
38. Brebner JA, Brebner EM, Ruddick-Bracken H. Experience-based guidelines for the implementation of telemedicine services. *J Telemed Telecare*, 11: S1:3–5, 2005.
39. Whitten P, Adams I. Success and failure: a case study of two rural telemedicine projects. *J Telemed Telecare*, 9(3): 125–129, 2003.
40. Hopp F, Whitten P, Subramanian U, Woodbridge P, Mackert M, Lowery J. Perspectives from the Veterans Health Administration about opportunities and barriers in telemedicine. *J Telemed Telecare*, 12(8): 404–409, 2006.
41. Gagnon MP, Cloutier A, Fortin JP. Quebec population and telehealth: a survey of knowledge and perceptions. *J Telemed E-Health*, 10(1): 3–12, 2006.
42. Gagnon MP, Duplantie J, Fortin JP, Landry R. Exploring the effects of telehealth on medical human resources supply: a qualitative case study in remote regions. *BMC Health Serv Res* [serial online]. January 2007; issue 7:6.

43. Gagnon MP, Fortin JP, Landry R. Telehealth to support practice in remote regions: a survey among medical residents. *J Telemed E-Health*, 11(4): 442–450, 2005.
44. Gagnon MP, Duplantie J, Fortin JP, Landry R. Implementing telehealth to support medical practice in rural /remote regions: what are the conditions for success? *Implementation Sci* [serial online]. August 2006; issue 1:18.
45. Campbell JD, Harris KD, Hodge R. Introducing telemedicine technology to rural physicians and settings. *J Fam Practice*, 50(5): 419–424, 2001.
46. Seibert PS, Payne K, Hudspeth R. Rising to the challenge of telemedicine research. American Telemedicine Association. May 7–10, 2006 San Diego, CA.

Chapter 19

Conventional Telemedicine, Wireless Telemedicine, Sensor Networks, and Case Studies

Lauren Biggers, Yang Xiao, and Fei Hu

CONTENTS

19.1 I ntroduction	369
19.2 C onventional Telemedicine.....	369
19.2.1 H istory of Telemedicine.....	369
19.2.2 C urrent Uses of Telemedicine	370
19.2.3 C urrent State of Telemedicine.....	371
19.2.4 T he Problems of Conventional Telemedicine	371
19.3 W ireless Telemedicine	372

19.3.1	What Is Wireless Telemedicine?	372
19.3.2	Advantages of Wireless Telemedicine	372
19.3.3	Hardware and Software	374
19.3.3.1	Connectivity	374
19.3.3.2	Camera Phones	374
19.3.3.3	Tablet PC	375
19.3.3.4	Laptop PC	376
19.3.3.5	Web Camera	376
19.3.3.6	Digital Peripheral Devices	376
19.3.3.7	Sensors	377
19.3.3.8	Data Compression Techniques	377
19.3.3.9	RTB2400 Wireless Router	378
19.3.3.10	IP Telephony Software	378
19.3.4	Disadvantages of Wireless Telemedicine	378
19.4	Sensor Networks	379
19.4.1	Applications	380
19.4.2	What Is a Sensor Network?	381
19.4.3	Anatomy of a Sensor Network	381
19.4.4	Security Problems	383
19.5	Case Studies	384
19.5.1	Mobile Telemedicine Systems	384
19.5.2	Wireless Telemedicine Kit for Nursing Homes and Retirement Centers	385
19.5.3	Wireless IP for Telemedicine	386
19.5.4	LINCOS Project	386
19.6	Conclusions and Future Research	387
	Acknowledgment	388
	References	388

Telemedicine has been in use for many years, and is the use of telecommunications technologies to consult with a remote physician. Conventional telemedicine has a 30-year history. This chapter first explores conventional telemedicine in some detail — its advantages, disadvantages, hardware, and software. By studying conventional telemedicine, some reasons why it is not ideal for all medical situations should become apparent. This chapter then explores other areas of telemedicine and their applications. These areas include wireless networks and sensor networks and how each can advance current telemedicine solutions. This chapter describes the feasibility of each new technology for telemedicine and the hardware and software required to implement telemedicine in these networks.

19.1 Introduction

Telemedicine can be defined several ways; however, the core concept remains the same. Telemedicine is the use of modern telecommunications technologies to deliver health care remotely to individuals who are unable to attend a physical appointment with their doctor or specialist.¹⁻⁵ In a perfect world, telemedicine can be employed to provide excellent health care to people who would otherwise be unable to talk with a specialist about their problems.² The health care a patient receives by telemedical means should be as good as the health care a patient would receive in an office visit.

Over its 30-year history, conventional telemedicine has offered patients many advantages. However, conventional telemedicine also has several drawbacks. While conventional telemedicine systems are well defined and dependable, they are also large, immobile, and expensive.^{3,5}

Many researchers have begun investing their time into the research of wireless telemedicine systems. Proponents of wireless systems claim that the increased mobility and the lower cost of the systems are highly beneficial to telemedicine. Mobility and lower cost health care solutions are benefits of new telecommunications technologies. Wireless systems are not a far-fetched idea. Their usefulness and reliability have been tested in several trials.¹⁻⁷

The objective of this chapter is to present an overview of telemedicine: past, present, and future. Therefore, the structure of this chapter is as follows. Section 19.2 covers conventional telemedicine. Section 19.3 focuses on wireless telemedicine applications while Section 19.4 deals primarily with sensor networks and their applications in telemedicine. Section 19.5 is dedicated to case studies and some of the telemedicine systems that have been implemented. Section 19.6 provides the interested reader with possible areas of research. Finally, concluding remarks are in Section 19.7.

19.2 Conventional Telemedicine

19.2.1 History of Telemedicine

Doctors and researchers began experimenting with telemedicine about 30 years ago. At that time, telephones and fax machines were the new telecommunication technologies. Conventional telemedicine was successfully implemented over wired communication technologies, such as plain old telephone service (POTS) lines, integrated services digital networks (ISDN), or T1-class data links.^{2,3,5}

In 1959, two-way videoconferencing began taking place between the Nebraska Psychiatric Institute and Norfolk Hospital, over 100 miles away. This link was comprised of a closed-circuit television link in order to allow communication between

the doctors at both locations.^{3,5} The link had two primary uses, as follows. The first was to allow the doctors at the hospital to consult with the doctors and students at the institute about patients. The second use was to allow doctors at the institute to observe and give psychiatric consultations to patients at the hospital without the necessity of traveling to the hospital.⁵

Another later use of conventional telemedicine was put into practice in Massachusetts. In April 1968, a microwave video link was created between the Massachusetts General Hospital and Logan Airport in Boston. This link had two significant advantages. The first advantage was that immediate health care could be given to airport employees as well as passengers. With the introduction of this telecommunication link came a second advantage: physicians no longer needed to be at the airport to provide health care to employees and passengers. Until this time in order to receive immediate health care at the airport, a physician needed to be on duty. However, now a physician at the hospital could use the microwave video link to diagnose a patient at the airport. Rather surprisingly, Logan Airport was also equipped to handle cardiology, dermatology, and radiology examinations.⁵

19.2.2 Current Uses of Telemedicine

The primary consumers of conventional telemedicine services today are hospitals, clinics, prisons, and any other installation where the equipment will not need to be moved once it has been installed. Conventional telemedicine lends itself particularly well to these locations because doctors and patients are already in these locations and travel for either the doctor or the patient would be inconvenient.³

In prisons or penitentiaries, it is often difficult, costly, and risky to bring a medical specialist on-site or to move an inmate to see a specialist. When moving an inmate patient to see a doctor, precautions must be taken to ensure that the prisoner will not be able to flee once he is outside the prison. This requires planning and manpower. Also there is a risk associated with bringing a physician to a prison. If given the opportunity, an inmate could attempt to hold the physician hostage for his own benefit. By using telemedicine, there is no need to move the inmate to see the doctor or for the doctor to come to see the patient. Therefore, prisons are ideally suited to telemedicine because patient and doctor interactions can occur with much less cost and risk.³

In hospitals, it is often a financial burden to transport a patient to see a specialist or to call in a remote specialist to see a patient. It is often much easier, in terms of time and money, to move a patient and specialist into rooms equipped with telemedicine equipment in their current respective locations. The physician and the patient can then have a medical consultation without the burden and cost of traveling to meet one another. Therefore, hospitals are also prime candidates for conventional telemedicine systems.³

19.2.3 Current State of Telemedicine

Currently, telemedicine is still very much conventional telemedicine. Conventional telemedicine consists of at least two distinct locations connected by a wired link. Because this link is wired, conventional telemedicine does not lend itself to mobility, flexibility, or portability. This lack of mobility, flexibility, and portability are common reasons for researching wireless telemedicine.¹⁻⁷ As telemedicine equipment becomes more mobile, flexible, and portable, telemedicine consultations will become increasingly popular alternatives to traveling to a medical appointment and for use in emergency situations. Emergency cases, disaster areas, and locations with high patient-to-doctor ratios will benefit from the use of portable equipment.¹⁻⁵

Conventional telemedicine currently uses T1-class data links, POTS, and ISDN lines to transmit user input data securely and provide two-way videoconferencing, high-resolution digital photography, and data capture from medical devices in a real-time environment.^{2,3}

For the locations mentioned above, conventional telemedicine provides an adequate consultation between a patient and a doctor. However, if the patient or doctor is not at a location that has telemedicine equipment installed, then these people must still incur the cost and burden of travel. While many telemedicine patients have responded favorably to the current system,³ there is still much work to be done in order to make telemedicine useful to a much larger population.¹⁻⁷

19.2.4 The Problems of Conventional Telemedicine

A major problem with conventional telemedicine is that it is highly dependent on wired data links.^{2,3,5} These links are often rented from communications companies, and therefore can become very expensive.³ Also, the hard-wired links used in conventional telemedicine do not lend mobility or portability to the conventional telemedicine systems.¹⁻⁷

Because the links are immobile, many of these conventional telemedicine systems are comprised of large, expensive machinery. Since the machinery must remain immobile because of the T1-class data links, it does not matter how large or bulky the equipment associated with the system becomes. However, since many places that use conventional telemedicine equipment are not concerned with having a portable telemedicine kit, they have upgraded their equipment to very sophisticated technology. Therefore, although this equipment is immobile, it is very powerful, very effective, and can be very expensive.³ Unfortunately, the immobility of these systems inhibits real-time telemedicine at the patient's location.³

19.3 Wireless Telemedicine

19.3.1 *What Is Wireless Telemedicine?*

Wireless telemedicine is a new form of telemedicine that can be conducted over a wireless network. A more detailed explanation of wireless telemedicine is a consultation between a remote physician and a patient that uses wireless telecommunication technologies to link the end systems of both the physician and the patient to the Internet in order to allow this remote consultation to take place from any geographic region covered by a wireless network.¹⁻⁵ A wireless system is ideal for solving the problems of immobility, inflexibility, and importability posed by conventional telemedicine.¹⁻⁷

19.3.2 *Advantages of Wireless Telemedicine*

Wireless telemedicine does not sound very different from conventional telemedicine systems. A person may be wondering why researchers are spending so much time studying wireless telemedicine systems. After all, conventional telemedicine has been in use for over 30 years, so theoretically the research community should be knowledgeable on the subject of telemedicine.^{2,3,5} These kinds of thoughts are why this next section of the chapter is included. Before getting into the implementation and hardware issues associated with wireless telemedicine, it is appropriate to study why wireless telemedicine is good: what are its advantages and how can these advantages change the conventional telemedicine system?

First, there is the issue of mobility. In a conventional telemedicine system, the telemedicine equipment is generally large and immobile.³ However, a wireless telemedicine system will allow a medical consultation to be taken to the participants as long as the geographic region is covered by a wireless network.¹⁻⁵ This makes wireless telemedicine ideally suited to retirement homes, hospitals, or any place where the mobility of a patient is a constraining factor. It is often much easier to move the system than it is to move a highly immobile patient.³

Since wireless telemedicine kits can be mobile, it also makes sense for the kits to be portable.¹⁻⁵ Just because a telemedicine kit uses wireless technology does not mean the kit is portable. Portability allows the equipment used in a telemedicine consultation to be moved easily from one area to another. Ease of movement is based on how easily a system can be brought into a room, set up, a telemedicine consultation can be given, and the tear down of the kit for movement to a new location.^{3,5} Adding portability to mobility allows many changes to evolve in telemedicine systems. Some of these changes will require further research but many of the changes are beneficial to doctors and patients who currently use telemedicine systems.¹⁻⁷ For example, consider an elderly person with advanced Parkinson's disease living in a retirement home. This disease can severely restrict the mobility of the patient and Parkinson's patients frequently have medical consultations with their

doctors to control or modify the types of medicines he or she consumes. A mobile and portable telemedicine system can be taken to the patient's room to allow the patient to have a medical consultation without the need to travel either out of his room at his current location or to another geographically distant location. This use of a portable, wireless telemedicine kit is extremely beneficial to the staff at such an institution by allowing the patient to have frequent medical consultations without the need to travel any considerable distance to see his or her doctor.³

Another benefit of portable, wireless telemedicine systems is that some patients feel confined to their hospital rooms. A feeling of confinement can discourage or even depress patients. However, if a patient who is not in need of constant medical care is equipped with a wireless, patient-worn telemedicine kit, this patient can return home, or to an area covered by a wireless network, to enjoy as much of a normal routine as possible.¹ The patient's doctors can then obtain real-time medical data from the patient-worn telemedicine kit via a gateway PDA.¹ In this way, a patient can still be under a doctor's care without the confining restriction of constantly staying in or frequently visiting a hospital.^{1,3}

Second, the tools needed to construct a wireless telemedicine system are often much more inexpensive than the arrangement used for conventional telemedicine systems.^{3,5} In many cases, the entire telemedicine system is composed of different components but the idea is the same: to provide health care to those individuals who are unable to attend a regular, in-person appointment with a doctor or specialist.¹⁻⁵ One reason the wireless telemedicine system is more inexpensive than a conventional telemedicine system is that there is no wired link between the wireless system and the Internet as there is with the conventional system. Therefore, there is no large cost associated with renting a high-bandwidth link from an Internet Service Provider (ISP).³

Another reason is the different types of equipment that are being used. Large, conventional telemedicine systems often use very expensive equipment. For smaller institutions, the cost of these conventional telemedicine systems is too high for the organization to absorb.³⁻⁵ However, competitively priced software and hardware for wireless telemedicine systems have been developed. This software and hardware will allow the spread of portable, wireless telemedicine kits in smaller institutions that simply cannot afford the equipment associated with conventional telemedicine.³⁻⁵ Some kits consist of no more than a wireless Internet connection, a laptop or tablet PC, a digital camera, some simple digital peripherals, and a connected Webcam with an integrated microphone.³

Finally, there is the issue of time and resource management. In any society, doctors are precious resources because they are healers. A doctor's skills are always in demand. In the United States, for every doctor there are 200 to 500 patients.⁵ Clearly, doctors are a limited resource.¹ By employing wireless telemedicine, society can make better use of such a limited resource. Most patients must currently travel to see a doctor, so this involves a factor of time. However, if the patient is too ill to travel, then the doctor must travel to the patient. In this scenario, time that could

be spent with one patient must now be spent traveling to visit another patient. For this situation, if telemedicine were conveniently available to both parties, then telemedicine would be a viable option and it could save time for both the doctor and the patient.^{1,3-5} In such a way, telemedicine can make much more efficient use of one of society's most valuable resources — the doctor.

19.3.3 Hardware and Software

Special tools are needed to conduct a wireless telemedicine session.¹⁻⁷ Many different wireless systems and prototypes have been developed. This section of the chapter discusses many of the hardware and software requirements of these different types of wireless telemedicine systems. Not every system is composed of the same hardware and software. A variety of hardware and software is presented in this section. Each piece of hardware or software presented has been used in a developed wireless telemedicine system. These various pieces of hardware and software are being presented to offer the reader a greater understanding of the available options for wireless telemedicine kits.¹⁻⁷

19.3.3.1 Connectivity

First and foremost, there must be access to a wireless network session, but the bandwidth must never have less than 100 kbps to devote to the session.^{1,3} If the bandwidth drops below 100 kbps, the real-time remote consultation will not be transparent to the users of the telemedicine system. The purpose of real-time telemedicine is to allow a patient and a doctor to interact with one another as normally as possible although being separated by some distance.¹⁻⁵ In order to ensure as normal a doctor-patient visit as possible, video is highly relied upon. Good video and good audio quality allow the patient to feel that his doctor is in the room with him, fostering an important virtual presence. Therefore, any interruptions in the audio or video will make the telemedicine session apparent to the patient and to the doctor.³

19.3.3.2 Camera Phones

Now that the importance of the link has been stressed, we should talk about some of the other hardware components currently used in telemedicine. Japan is generally three to five years ahead of Europe in technological trends. As such, Europe and America can expect to see these same technologies in the near future.

In Japan, good quality camera phones, such as the KDDI/AU Infobar with an automatic focus, flash, and global positioning systems (GPS), are becoming an

increasingly important part of telemedicine. These phones can be used to take pictures of an accident and transmit this data to a doctor in a hospital using a commercial wireless carrier such as Verizon Wireless.²

Japanese police have also used these phones to monitor stalking victims in an attempt to help ensure the victim's safety at all times.

The civilian population has also put these phones into use to help take care of the elderly and easily confused. With this phone, if an elderly person were to become lost, he could take a picture of his surroundings and send this picture along with GPS data to a family member or someone else who can come find him and pick him up.² The receiver can then use the GPS data and the picture to determine the location of the lost person, and therefore go to pick him up.²

In 2004, third-generation (3G) mobile communication networks covered 99% of Japan. These communication networks are the reason cellular phones are able to play such a large part in telemedicine in Japan.² Because of the almost complete coverage of the nation, Japanese citizens with chronic diseases are able to take advantage of these technological advances in telecommunications in order to monitor their condition more easily.²

19.3.3.3 Tablet PC

Another tool that is very beneficial in telemedicine is a tablet personal computer (tablet PC). Tablet PCs are different from regular desktops and laptops. A user can interact with these PCs in the same way that he can interact with a regular PC. However, these computers also function as pen-input devices. A pen-input device allows the user to write on the screen with a stylus. The goal of a pen-input device is to provide a familiar way to interact with an unfamiliar object, such as the PC.³

When pen-input devices were first introduced, they did not explode in the market. These input devices contained a major flaw: the associated handwriting recognition software just did not perform well enough.³ The poor performance of handwriting recognition software severely inhibited the growth of pen-based input products.

Today some tablet PCs offer an extended version of the Windows XP operating system.³ The beauty of this design is that a user can have all the Windows functionality that he is accustomed to, but he can also have the added functionality of useable pen-based input methods. For example, suppose that a nurse needs to draw a diagram for the remote doctor. Some people find drawing in the Windows program Paint to be awkward and unnatural. However, if this nurse is using a tablet PC, she can rotate and flip the screen so that she can write on the screen with a stylus as if the screen was a pad of paper — hence the name tablet PC.³

19.3.3.4 *Laptop PC*

If a tablet PC is not available, another wireless solution would be to make use of a laptop personal computer.³ Although a normal laptop cannot provide pen-based input methods, it can allow a user to connect to a wireless network to access the Internet. Once connected to the Internet, data can be transferred between a remote patient and doctor.³

19.3.3.5 *Web Camera*

Another popular item for a wireless telemedicine kit is a Webcam. Webcams are used to transmit images across the Internet. While a Webcam might not be appropriate for emergency use, it would be more than adequate for a general checkup with a doctor.³ Some of these Webcams come with an integrated microphone. Such a device would allow audio and video to be streamed over the Internet and to a remote physician. This one device is not too expensive. Therefore, the Webcam is perfect for smaller institutions that cannot afford a larger, more conventional telemedicine system.³

19.3.3.6 *Digital Peripheral Devices*

As important as videoconferencing is, it is not the final component of a wireless telemedicine kit. There are many digital medical instruments that are needed to complete a thorough telemedicine session. These instruments allow the recording and transmission of medical data to a remote physician. Digital stethoscopes, sphygmomanometers, and cardiovascular and respiratory monitoring systems allow a remote physician to assess the current health situation of one of his patients.

In many cases, an attending nurse or the primary care physician operates the wireless telemedicine kit.^{1,3,5} In an attempt to make use of the current available technology, digital medical peripherals that can interface with a computer have been designed and manufactured. These digital peripherals interface with a tablet PC or laptop to send a patient's vital signs to the remote physician without the necessity of manually inputting all of the patient's medical data into the computer by hand. Several standard medical tools have become digital peripherals including electronic stethoscopes, sphygmomanometers, electrocardiogram recorders, and digital thermometers.^{1,3,5} These peripherals allow the capture and transmission of patient vital signs without the need for the nurse or doctor to input this information manually.^{1,3,5} This eliminates any human error associated with inputting data and transmission.

The sophistication of the particular wireless telemedicine kit being used dictates which of these peripherals are included in the kit. Some kits will not contain an electrocardiogram recorder because it performs functions outside of the intended use of a particular wireless telemedicine system.^{3,5}

Some peripherals still use the RS232 serial interface to transfer the data collected by a peripheral device to the computer.³ However, there are peripherals that run the current universal serial bus (USB) interface.⁵ The USB interface is likely to become the standard peripheral interface in the future because RS232 serial interfaces are an aging standard.³

These digital peripheral devices are moderately priced.⁵ Therefore, these devices can be integrated into a wireless telemedicine kit without significantly increasing the cost of the system.

19.3.3.7 Sensors

Sensor networks have been suggested for use in telemedicine kits intended to monitor patients with chronic illnesses at home.^{1,7} A more complete discussion of sensor networks will follow in the next section of this chapter. However, sensor networks can be used in conjunction with wireless telemedicine systems.⁷ For the moment, it is enough to know that sensors can be used to monitor and collect patient data.¹ This data can then be sent to a cellular telephone or a PDA with an Internet connection.¹ The information collected by the sensor can then be aggregated and relayed to the primary doctor or to a database established as a medical information repository.^{1,7}

19.3.3.8 Data Compression Techniques

In 1983, the American College of Radiology (ACR) and the National Equipment Manufacturers Association (NEMA) formed the Digital Imaging and Communications in Medicine (DICOM) committee.² DICOM adopted several formats for the transmission of still images.² The committee decided on two types of compression techniques for still image JPEGs: lossless JPEG and JPEG-LS.² The effectiveness of these compression techniques determines how quickly still images can be sent to and received by the remote attending physician.²

In the past, compression techniques for medical images have focused on lossless compression methods so that the original image could be reconstructed exactly at the receiver side.² Unfortunately, these techniques do not provide much compression of the image; in fact the compression ratio is normally between 2 and 3.7 for lossless methods.² These ratios offer only a limited improvement over sending the image without any compression.

However, lossy compression methods have been introduced so that the image is only approximately reconstructed.² With an approximate image reconstruction, several areas must be studied in order to ensure that the image still contains diagnostic value. One area is the effect of compression on diagnostic performance. Two other areas are optimal performance of the compression technique and the impact of a lossy compression technique on different diagnostic situations.²

In order to determine the impact of compression on diagnostics, a simple test can be performed. To perform this test, testers create a set of images comprised of compressed and reconstructed images and the original uncompressed image. If the original image cannot be identified in the set of compressed images, then the compression technique works sufficiently well in reconstructing the image such that the reconstructed image is as good as an uncompressed original image.²

Another useful compression concept is the region of interest (ROI).² The region of interest is the area of the image that contains a specific diagnostic value to a physician. By limiting compression techniques to only the ROI, there can be a drastic improvement in the compression ratio.² For example, if the ROI only covers 20% of the image, then a 15.1 average compression ratio has been reported for JPEG-LS compression techniques. If the entire image were the ROI, then the average compression ratio would only be 2.58.²

By improving compression techniques, the transmission rate of an image can be decreased.² A decreased transmission rate allows the end users to interact with the system in a transparent manner. Accordingly, appropriate compression techniques can improve the value of telemedicine by improving the overall experience of the system users. With increased response times comes a more transparent user session.³

DICOM has not yet adopted a standard for digital video transmission, so those compression techniques will not be discussed.

19.3.3.9 RTB2400 Wireless Router

This wireless router can receive and transmit data, and therefore it can also act as a repeater. When the router is used in compliance with Japanese regulations, it can achieve a circular coverage area with a radius of 3 km with transmission speeds of up to 2 Mbps on a TCP/IP platform.⁴

19.3.3.10 IP Telephony Software

IP telephony software is becoming of greater interest to some researchers in wireless telemedicine systems.⁴ IP telephony, also known as voice over IP, can provide real-time, two-way, synchronous voice and data traffic over a packet-switched IP-based network.⁴

19.3.4 Disadvantages of Wireless Telemedicine

Wireless telemedicine is not always transparent to the users. Sometimes, because of bandwidth fluctuations or lost, late, or unordered packets, artifacts can be introduced into the video streams.^{1,3} Other times audio quality is not as high as

it should be, and therefore the users are either distracted by the disturbance in the audio data or unable to understand the audio at all.^{1,2-4} Still at other times, the connection is not able to support all the incoming and outgoing data in real-time. Some data packets may even be lost, dropped, or arrive out of order because of network congestion.¹ If any of these things happen, the telemedicine session is no longer transparent and can even become a hindrance to the medical consultation.³

Another disadvantage is the misuse of telemedical equipment. For example, a wireless telemedicine kit ideal for use in a retirement home might include a Webcam and a tablet PC.³ However, this equipment may not be suited for use in an emergency situation. In an emergency situation, like inside an ambulance, still images might be better along with data collected by digital instruments.¹ For this application, doctors are able to monitor the patient's vital signs and view the extent of the damage to a patient without worrying too much about the quality of a video stream.¹

Security is another problem in wireless telemedicine.^{3,7} A person's medical records will ideally remain private at all times. The Health Insurance Portability and Accountability Act (HIPAA) was developed to protect confidential health care information through improved security standards and federal privacy legislation.⁷ HIPAA dictates how electronic medical data must be protected before, during, and after electronic transmission.⁷

Security is another current problem in sensor networks that can be used in telemedicine.⁷ Security in sensor networks will be discussed in greater detail in the next section. However, sensor networks are not the only point of interest. Wireless transmissions are much more easily intercepted than are wired transmissions.³ In this aspect, conventional telemedicine systems have an advantage over wireless telemedicine systems.

One final disadvantage of telemedicine systems is the human element. The administrators and operators of the system must be properly trained to operate the equipment. Currently, many institutions are understaffed and do not have time to train their employees in a new skill set unless the new technology has been tried and proven.⁷

19.4 Sensor Networks

Sensor networks can be used as a key component in some wireless telemedicine systems.^{1,7} The sensors of a sensor network can monitor a patient's blood pressure, heart rate, oxygen saturation, and temperature.⁷ Some people may consider the digital medical peripherals discussed earlier to be sensors. In fact, sensor networks are not an application of telemedicine at all. Telemedicine is one application of sensor networks.⁷ This section considers several topics including their applications, what a sensor network is, the anatomy of a sensor network, and security problems.

19.4.1 Applications

The medical field currently has several active databases for electronic patient records. Some of these databases are the Picture Archiving Communication System (PACS), Order Communication System (OCS), and the Electronic Medical Record (EMR). Information from these separate databases and information a doctor obtains from a patient visit can be combined into one of two databases, or both if both are applicable to the patient.⁷

The first major database is the Integrated Medical Information System (IMIS). The IMIS can combine the patient's medical information from PACS, OCS, and EMR along with his primary physician's notes. This can make the IMIS very useful to other physicians.⁷ Suppose that a person was in an accident and required medical care. A doctor other than the person's primary physician could then access the IMIS database to check for any allergies this person might have before administering any medication.

Another advantage of telemedicine applications is to integrate all the information from PACS, OCS, and EMR into an Integrated Disease Surveillance System (IDSS).⁷ The IDSS is particularly useful when monitoring a patient with a chronic disease. The IDSS can be used within a telemedicine application and in conjunction with sensor networks to monitor the progress of the patient's disease.⁷ As previously mentioned, a sensor can collect patient data and transmit this data to the Internet via an Internet-enabled cellular phone or PDA.⁷ For a patient with a chronic illness, once this data has entered the Internet it will be sent to the IDSS. By allowing sensors to perform this function, it is possible for physicians to obtain more accurate information about the patient's current health status.⁷ The physician also has the option to review his patient's current condition any time he is able to do so.

As telemedicine becomes more popular, there will be a proliferation of patient-worn wireless telemedicine devices.⁷ Instead of requiring a patient to make daily visits to a doctor for disease surveillance, patient-worn wireless telemedicine can transmit the needed data to the IDSS for later review by his primary physician.⁷

Patient-worn wireless telemedicine will help to alleviate the burden in other areas of medicine.⁷ Currently, there is a nursing shortage and the shortage does not seem to be coming to an end. In fact, the nursing shortage is only becoming more severe. The AHA has predicted that there will be over 600,000 unfilled nursing positions in 2008.⁷ Such a serious shortage will require new models of patient care and greater workplace efficiency.⁷

Patient-worn wireless telemedicine systems will help to alleviate patient problems. If the patient opted to wear a wireless telemedicine device, then he could experience greater mobility, comfort, and flexibility.⁷ The patient-worn telemedicine kit would allow him to move about more freely, not just within the walls of a hospital. The only constraint on the patient would be that he must remain in an area covered by a wireless network.

Now that the importance of a patient-worn wireless telemedicine device has been introduced, a discussion of the sensor network will follow.

19.4.2 What Is a Sensor Network?

A wireless sensor network consists of spatially distributed, autonomous devices. These devices use sensors to monitor the physical and environmental conditions of the object under study.⁸

The purpose of using sensors in patient-worn telemedicine devices is to gather data automatically without inconveniencing the patient.⁷ The process of gathering data and transmitting it from the source to a receiving station is called telemetrics.⁹ Currently, most of the research being done with sensor networks is telemetric in nature.⁷

For telemedicine, a telemetric sensor network can be thought of as a personal area network (PAN). The PAN should cover only a small area of the patient's body. The total number of sensors in a PAN can vary. Too many sensors will make the wearer uncomfortable but too few sensors will not effectively monitor the wearer's health.⁷ Therefore, the number of necessary sensors varies based upon the condition of the patient and the vital signs data the patient's doctor would like to collect.

As with any wireless technology, security is critical. Security is especially important in regard to the wearer's personal medical information.⁷ The security and security problems of a sensor network will be discussed later in a following subsection.

In order for transparency to occur in a patient-worn telemedicine kit, bidirectional nodal communication is necessary.⁷ By allowing the sensors to send and receive messages, a physician can make changes to how often a sensor collects patient data without the necessity of disturbing the patient. Also, the sensor networks described are hierarchically ranked.⁷ Therefore, bidirectional communication can occur between a leaf node and a parent node, and parent nodes can then communicate with the root node to access instructions.⁷

Bidirectional communication makes data aggregation much easier to perform. It is essential for a multi-sensor network to perform optimal data aggregation.⁷ If the data is never pieced together, it will be of no use to the physician or the patient. In order to fully comprehend the current physical status of a patient, the data gathered by the sensors must be fused together into a cohesive unit.

19.4.3 Anatomy of a Sensor Network

An example of the sensor network overview is illustrated in Kim.⁷ The sensor networks described in this paper conform to this system. A root node is in control of the sensor network at all times. The root node decides when and how to distribute its processing bandwidth based on the weight of the patient nodes and the weight of the sensor nodes.⁷

The weight of a node is determined by the node's relative criticality.⁷ The weight for each node is greater than or equal to zero and less than or equal to one.⁷ The root node assigns a weight for each lower node, whether the node is a patient or a sensor.⁷

If the root assigns a weight of zero to any node, that node is considered disconnected from the sensor network.⁷ The root assigns lower weights to nodes with less data. By determining node weight this way, the sensor network will always prioritize nodes with the greatest amount of aggregated data that needs to be processed by the root. Once the root has processed the information contained in the node, it assigns a lower weight to the node. Assigning weights based upon the amount of unprocessed information a node contains will allow each node to be processed, namely, there is no starvation or deadlock of the system.

The root node will be the starting point for the following discussion. The root node is a very important node because it assigns weights to all the lower nodes.⁷ The root contains the control information for the given sensor network. It is the root that decides the buffer length and weights that can be used by each patient node.⁷ Ultimately, the root is in charge of the entire sensor network.

Although the root has primary control of the network, the patient node is an intelligent node.⁷ When the weight of the patient node has been set to zero by the root, the patient node can begin or continue to perform data aggregation from its many sensor nodes. Data fusion and reference signal storage will increase the criticality of a patient node.⁷ Once the criticality of a patient node has begun to increase, the patient node can reconnect to the sensor network. When the patient node's criticality has reached a threshold, the root node will begin to assign a weight to the patient node.⁷ At some point, the patient node will have the greatest weight, the greatest amount of data unprocessed by the root node in the network, and the root node will access this patient node to process the information contained within the node. Once the information has been processed, the root node will set the weight of the patient node to zero and the process will start over.⁷

At the bottom of the tree lie the sensor nodes. Each patient has a unique set of sensor nodes. These nodes monitor a patient and collect medical data. The sensor nodes then send their signal information up the tree to their parent node. The parent node then uses its multi-sensor data fusion algorithm to order properly and aggregate the data it has received from the sensor nodes.⁷

The process described above is a variable bit-rate process. The variable bit-rate process was implemented to solve two major issues. A variable bit-rate process is energy efficient and helps to prevent bandwidth bottleneck.⁷

Energy efficiency is highly valued in a sensor network because a patient node has a limited amount of battery power.⁷ Allowing the patient node to connect to the network only when it has information to send will help the patient node to conserve its battery power.⁷

The bandwidth bottleneck is actually two different bottlenecks: a communication bottleneck and a processing bottleneck.⁷ If all the patient nodes were connected to the network at once, there would not be enough bandwidth to accommodate both the communication between the nodes and the processing of nodal information.⁷

19.4.4 Security Problems

Security in a sensor network is of the utmost importance. Some of the general security problems lie in the structure of the sensor network itself. First, sensor networks are made up of many nodes. Because of the network's construction, it would be impractical to try and monitor each node within the network. However, by monitoring only a subset of sensor nodes, another subset of sensor nodes is vulnerable to attack. A person attempting to hack into a sensor network can take one of two courses of action: he can compromise a subset of the sensor network's nodes or he can trick the sensor network into accepting an illegitimate node.⁷

A sensor network has limited hardware resources.⁷ Because of the limited resources of the network, a conventional approach to security is too heavy for this type of lightweight system. Therefore, lightweight but resilient security measures are needed to ensure the safety of the system.⁷

A security breach in the system can be classified into one of two broad categories: malicious use of a commodity sensor network or eavesdropping.⁷ Malicious use of a commodity security network includes such crimes as using sensors to detect an individual's presence at a private home and to obtain passwords or other private information from an individual's cellular telephone or personal computer.⁷

To counter a malicious commodity sensor network attack, an individual might install sensor detectors. Sensor detectors will not block a n intruder and they are expensive but they can give the owner or user of the sensor network important information.⁷ A sensor detector will detect the presence of potentially hostile wireless communications within an area but to do this the detector must be able to distinguish between authorized and unauthorized sensor networks and devices.⁷ Although the sensor detector can provide information about an attack, it is a costly defense when prevention of the attack cannot be guaranteed.

Eavesdropping, also known as sniffing, is the act of gaining private information through the unauthorized monitoring of nodal transmissions, accessing stored sensor data, and querying the sensor network.⁷ Once the sniffer has access to the stored data, he can perform data aggregation operations to determine private information about the individual who owns or operates on the sensor network.

As with the malicious use of a commodity sensor network, there are some available defense mechanisms.⁷ To protect a sensor network from an eavesdropper, the network can use encrypted communications. Once again, because of the lightweight nature of a sensor network the conventional end-to-end data encryption techniques are ill suited. The best communication encryption techniques rely on hop-by-hop encryption and multi-path routing.⁷ A n other good defense against sniffers is to query in a distributed manner. By querying across the system no single node can have access to the complete query results.⁷ Therefore if an eavesdropper is listening to the sensor network, he will be unable to obtain the entire query result by listening to a subset of the sensor network's nodes.⁷ Finally, a system can anonymize any sensed data within the sensor network by removing the identifying details

contained in the data.⁷ This technique will make it extremely hard for a sniffer to make sense of any data he can pick up.

The HIPAA standards mentioned earlier in the chapter apply to all electronic transmissions of personally identifiable health care information, whether the transmission is wired or wireless.

19.5 Case Studies

Earlier in this chapter, many different types of hardware and software associated with various telemedicine kits were listed. However, not every telemedicine system uses every piece of equipment mentioned. This section of the chapter will look at some of the systems that have been implemented by researchers in wireless telemedicine and explain why each system needs its particular subset of hardware and software.

19.5.1 *Mobile Telemedicine Systems*

Chu and Ganz implemented a portable, wireless telemedicine kit to be used in an emergency vehicle such as an ambulance.¹ They named their implementation of this wireless telemedicine kit the teletrauma system. The teletrauma system is composed of a teletrauma patient unit and a hospital unit.¹

The teletrauma patient unit consists of a laptop or a tablet PC, connected digital devices for vital signs monitoring, portable ultrasound equipment, and a video camera.¹ The teletrauma unit compresses the multiple multimedia streams from the attached peripherals and the video camera so that the amount of data that is sent can flow over the wireless link. In order to alleviate congestion and allow all the data streams to transmit information, different transmission methods are applied to different multimedia streams. Chu and Ganz use reliable Transmission Control Protocol (TCP) to send electrocardiogram (ECG) waveforms and images because TCP will ensure the delivery of the data packets. However, because videos can often tolerate some frame loss, the authors decided to implement user datagram packets (UDP) to transmit videos. Once the patient information is compressed and the transmission method is selected, the information is transmitted over a 3G wireless link and the Internet to reach the hospital unit.¹

In this case, the hospital unit is not wireless. The hospital unit is simply a PC connected to the Internet.¹ The PC receives the information being sent by the teletrauma unit and decompresses the information according to the media type. Now, a remote physician can examine an ECG waveform, medical images, and real-time video.¹

The authors believe that the teletrauma system can help reduce mortality and morbidity rates by allowing the hospital to prepare for the patient's arrival with prior knowledge of the patient's diagnostics.¹ The teletrauma system not only allows

the hospital to prepare for the arrival of the patient, but it also provides the emergency medical technicians (EMTs) with a specialist's advice before arriving at the hospital, which can improve the health care a patient receives in transit and thus improve the patient's chance for a full recovery.¹

Chu and Ganz also implemented a telepatient monitoring system. This system provides around-the-clock health care monitoring while a patient lives at home or in another area covered by a 3G wireless network.¹ The telepatient system requires the use of two networks. The first network is a wireless PAN whose main function is to collect a patient's medical data.¹ The second network, the 3G network, transmits the collected data to the health care provider.¹ A PDA acts as a link between the two networks. The PDA gathers the medical information collected by the PAN, aggregates the data, and uses the 3G wireless network to transmit the data to the health care provider.¹

19.5.2 Wireless Telemedicine Kit for Nursing Homes and Retirement Centers

Hackney built a wireless telemedicine kit for use in nursing homes and retirement centers.³ His goal for designing this kit was to provide wireless telemedicine for institutions that could not afford a conventional telemedicine system.³ The wireless telemedicine kit consisted of four key components: tablet PC, connectivity solution, digital medical peripherals, and customized software.³

The tablet PC employed by Hackney has a rotating screen that will flip down so that the operator can write on the screen as if writing on a pad of paper. Some of the nice features of the tablet PC are that it is lightweight and easily transported, contains wired and wireless Ethernet interfaces, can easily connect to peripherals via a USB port, and can run any commercially available software.³

Hackney suggests using a broadband connection if a suitable connection exists. However, if the connection does not exist, the tablet PC can connect to a local area network (LAN) and send the information through the LAN if one is available.³ The type of Internet connection used often depends on what the specific institution can afford. However, this connection must be able to transmit at least 100 kbps to ensure a high-quality, real-time telemedicine session.³ He does suggest using the H.323 videoconferencing standard. H.323 is a newer videoconferencing standard that defines a suite of protocols for videoconferencing using Internet Protocol (IP).³

Digital peripherals are used in the kit to collect standard diagnostic information.³ The kit uses USB connected Webcams with built-in microphones for the telemedicine consultation images. The Webcam allows the remote physician and the patient to interact while also being useful for minor diagnostic images.³ However, Hackney believes the kit should also contain a digital still camera for diagnostic images that require a high-resolution image.³

A superset of Windows XP is run on a tablet PC to allow for pen-based inputting.³ This operating system can run software written in .Net and is H.323 compliant. Customized software for each institution might be needed, depending on the various forms the institution would like to employ. The suggested software would be able to view and collect data from the peripheral devices. It might also automate any standard forms used by the institution for medical consultations.³

This kit is designed to be mobile and portable. The mobility and portability will allow the kit to be taken into the patient's room for the medical consultation.³

19.5.3 Wireless IP for Telemedicine

Zhao et al.⁴ implemented a wireless telemedicine kit over a wide area network (WAN), using the RTB2400 wireless router and IP telephony software to build their system.

The router was configured to a wireless IP WAN topology. This topology provides a flat connectivity between the nodes rather than a hierarchical connectivity. Having a flat network topology will reduce bottlenecks between the sender and the receiver because there are multiple paths between the pair that packets can travel.⁴

The IP telephony software allows the integration of voice traffic, images, and data.⁴ Some telemedicine systems would require a different stream for each type of data being sent. By integrating these three multimedia data types, the number of streams being sent and received over the Internet can be significantly lower than the number of streams implemented by other telemedicine kits.⁴

At least three routers and three IP telephony gateways are required to implement a minimum network with this hardware and software.⁴

19.5.4 LINCOS Project

The Little Intelligent Communities (LINCOS) project is an application of telemedicine in developing countries.⁵ In the United States, there is an approximate doctor-to-patient ratio of 1:200. However, in developing regions such as East Africa, where health care is trailing behind systems in industrialized nations, the doctor-to-patient ratio can be as high as 1:40,000.⁵ The LINCOS project began as a way to bring better health care to thousands of people who would not otherwise have access to a doctor.⁵

The LINCOS units are built from recycled ISO containers. These containers are subdivided into three sections: a computer lab, an information center, and a telemedicine center.⁵ The LINCOS unit is designed to be a digital town center. The computer lab is open to public use.⁵ The telemedicine station is where patients go to take part in a telemedicine consultation.

Because the LINCOS project is aimed at developing countries, several issues need to be taken into consideration. First, many of the towns and villages where the

LINCOS units will either have limited or no electrical connectivity. If no electrical connectivity is available, the unit can be powered by an alternate energy source such as solar power or gasoline generator.⁵ A second issue is Internet connectivity. It is likely that a village will not have either phone lines or a broadband connection.⁵ When this is the situation, a Very Small Array Terminal (VSAT) antenna can provide satellite communication to the LINCOS unit.⁵

The telemedicine kit developed for this unit is similar to the telemedicine kit for nursing homes designed by Hackney. It contains a large set of peripheral medical devices such as an electronic stethoscope, an ECG recorder, a medical imaging system, and devices for taking a patient's blood pressure, temperature, and other measurements.⁵

Unlike the telemedicine kit designed by Hackney, this telemedicine kit does not run in real-time. Instead it uses the asynchronous store-and-forward method.⁵ The store-and-forward method is ideal for the LINCOS units because it is inexpensive to maintain.⁵ It is not always possible to make a remote appointment with a specialist, even for a telemedicine consultation; sometimes there are conflicting schedules between the patient and the doctor.⁵ With the store-and-forward technique, an appointment with the remote physician is not necessary. The local physician can collect the patient's medical data and store this data in a computer. At a convenient time, the local doctor can forward this data to a specialist. The remote physician is then able to review the data at his convenience.⁵ After the remote physician has viewed the data, he can either give a diagnosis to the local doctor or he can ask the local doctor to perform further tests.⁵

19.6 Conclusions and Future Research

Telemedicine will become more common in the future. As protocols emerge that can ensure the quality of service and the security of the patient's medical information, many of these consultations will be taking place in real-time. Reliable real-time telemedicine technology will decrease the mortality rates associated with injuries that require a large amount of time in transit to a hospital.

Also, with the rise of home patient-monitoring technology, more patients will be sent home to recover from hospital visits. This technology will also become increasingly common as a monitoring device for individuals who suffer from chronic diseases.

There are many areas of research open in wireless telemedicine. Some suggested areas are mobility issues of current wireless IP telephony, public to private network interconnectivity issues, performance, reliability, and the integration of 3G, IPv6, and signal process for certain applications.⁴ Sensor networks also contain areas of research; for example, secure protocols and secure functions are needed for a new system model design.⁷ Other areas of research in sensor networks are aggregation of data, energy efficiency, and possibly a hard-wired approach.⁷

Acknowledgment

The work was partially supported by the U.S. National Science Foundation (NSF) under grants CNS-0716211 and CNS-0716455.

References

1. Y. Chu and A. Ganz, Mobile Telemedicine Systems Using 3G Wireless Networks, *Business Briefing: U.S. Healthcare Strategies*, 2–6, 2005.
2. C.S. Pattichis, E. Kyriacou, S. Voskarides, M.S. Pattichis, R. Istepanian, and C.N. Schizas, Wireless Telemedicine Systems: An Overview, *IEEE Antennas & Propagation Magazine*, 44, ED-2, 143–153, 2002.
3. D. Hackney, Wireless Telemedicine for Nursing Homes and Retirement Centers, B.S. thesis, Dept. of Computer Science, University of Virginia, 2005.
4. Y. Zhao, Y. Yagi, H. Juzoji, and I. Nakajima, A study of wireless IP for telemedicine, unpublished.
5. A. Adler, A Cost-Effective Portable Telemedicine Kit for Use in Developing Countries, M.S. thesis, Dept. Mechanical Engineering, Massachusetts Institute of Technology, 2000.
6. G. Fasol, Wireless telemedicine technology trends, unpublished.
7. M. Kim, Telemedicine: An application of sensor network and its security, unpublished.
8. Wikipedia, the Free Encyclopedia. 2006, November. Wireless Sensor Network. http://en.wikipedia.org/wiki/Sensor_network
9. The Free Dictionary. 2006. Telemetric, definition of telemetric by the Free Online Dictionary, Thesaurus, and Encyclopedia. <http://www.thefreedictionary.com/telemetric>
10. Eurotechnology, Wireless Japan 2002 Tradeshow (Tokyo BigSite). Report: <http://www.eurotechnology.com/wirelessjapan2002/>

Chapter 20

Telemedicine for Pervasive Healthcare

Quinton Alexander, Yang Xiao, and Fei Hu

CONTENTS

20.1 Introduction.....	390
20.2 Telemedicine for Pervasive Health Care.....	390
20.3 Medical Standards and Telemedicine.....	392
20.4 Applications and Emerging Technologies.....	393
20.4.1 The CodeBlue Infrastructure	393
20.4.2 Heart Monitoring by CardioNet	394
20.4.3 Tele-Ultrasonography with OTELO.....	394
20.4.4 Biomedical Sensor Networking Platforms.....	396
20.4.5 Next-Generation Network-on-Chip Protocol	398
20.4.6 Wireless LANs and Patient Monitoring.....	399
20.4.7 Un-cooperation of Routers in Wireless Patient Monitoring	400
20.5 Conclusions	403
Acknowledgment	404
References	404

This chapter provides a survey of telemedicine. Telemedicine deals with the application of technology and telecommunication resources to the practice of medicine. The materials surveyed by this chapter provide a general overview of this topic. This survey touches on issues concerning patient monitoring, emergency services, standardization, and mobility. Several other areas of interest are covered including reliability, routing concerns, implementation challenges, system analyses, and interoperability concerns.

20.1 Introduction

The average person has been to a health care facility or treated by a medical professional at least once in his or her lifetime. In fact many individuals visit a hospital or doctor's office at least once a year for regularly scheduled examinations. Health care is clearly an important aspect of the modern lifestyle. Because of its importance to society the health care system must constantly be evaluated and improved to help continue its usefulness to the society it supports. An effective health care system must provide reliable services and be responsive to the needs of the people it covers.

Telemedicine can be presented as a means of improving current health practices. It is a concept born from the emergence of this new age of technology and can be seen as the next logical improvement on an ever-evolving industry. Telemedicine is the basis for a pervasive health care system. It can provide better quality of service and reduced costs. Along with those improvements to the current system the increased efficiency will allow for the coverage of more individuals.

Several issues must be overcome to provide the benefits of telemedicine. A large number of concerns are centered on the use of wireless communication channels. Closely related is the application of sensor devices to the health care industry. System standards are also needed to establish interoperability between different service providers and approved devices.

A couple of services have been implemented on a limited basis and others are well along in the stages of testing. A communication platform named CodeBlue^{1,2} attempts to allow simple devices to operate in a decentralized manner that would support several types of services. An OTELO robot³ is being created to help medical experts give examinations to patients across long distances.

20.2 Telemedicine for Pervasive Health Care

Several challenges face the health care system of the United States and the world as a whole. The cost of health care is currently at 15% of the gross national product for the United States.⁴ Hospital errors account for 98,000 deaths each year in the United States.⁵ In a study conducted at the Hackensack University Medical Center

it was statistically determined that the health care industry would save nearly 140 billion dollars yearly through the increased utilization of existing technologies.⁵ The major issues affecting health care include rising costs, a shortage of qualified workers, increases in the number of medical mishaps, and poor health coverage in rural and impoverished areas.⁶ Pervasive health care would improve productivity and allow for a wider range of medical services.⁶

Telemedicine has been defined as the “use of advanced telecommunication technologies to exchange health information and provide health care services across geographic, time, social, and cultural barriers” in Choi et al.⁵ Advances in wireless technologies have opened the door to a number of powerful medical applications.⁶ A pervasive health care system would integrate mobile devices, wireless networks, and middleware in ways that would improve on the quality and availability of medical services.⁶

The widespread deployment of wireless networks has been proposed as a way to improve communication among patients, physicians, and other health care workers.⁶ Patients are often misdiagnosed because of the lack of complete and correct information at the time of service.⁶ In emergency situations the individuals attending to a victim may lack experience with that particular situation and will often have no knowledge of the victim’s medical history.⁶ Being able to transmit a patient’s medical data will allow hospitals to be better prepared for the patient’s arrival. Specialists would also be able to make recommendations on a condition from miles away.⁶

Not only could updates on a patient’s condition be sent directly from the ambulance, but paramedics and hospital staff would be able to retrieve the patient’s medical records electronically and review information on drug allergies and pre-existing conditions.⁶ If a medical facility has a wireless local area network, the doctors on staff could use a small handheld device to review and update medical records. Physicians would also be able to transmit prescriptions directly to a pharmacy, which would save time and avoid errors due to misunderstood handwritings.⁶ By combining WLANs and personal area networks a patient’s health could be monitored from most anywhere and workers quickly notified of any status changes or emergency service needs. This would allow more freedom for noncritical patients to move about the facility or even return home early and still be monitored by health care professionals.⁶

A pervasive health care system would allow for network- or even satellite-based positioning, which could help direct users to the nearest medical facility (even voice systems for blind users). Locating organ donors or people with matching blood types, and pinpointing any other need becomes easier.⁶ Tracking services would help ensure that a patient who is restricted to a particular area does not wander off or is not accidentally removed from the location. This can be achieved through the use of radiofrequency IDs and network sensors.⁶ Emergency rooms and other restricted areas would benefit from the ability to identify anyone entering a secure area without the proper ID.

In emergency situations, location-based tracking can help identify multiple reports of the same event and thus lower the burden on first responders and prevent multiple dispatches for the same emergency.⁶ Paramedics could be sent real-time traffic data that would allow patients to be routed more efficiently.⁶ Increasing efficiency in these areas would make it possible to handle more emergency calls sooner.⁶

Pervasive access to medical data gives the patient the ability to provide insurance data or medical history via a handheld device. This would save on the time needed to enter the information manually.⁶ Alerts could be sent to a patient's mobile device as a reminder to take prescribed medicines or to help keep up with appointments to see a doctor.⁶ Portable devices exist that can detect medical conditions like pulse, blood pressure, and breath alcohol level. The device might also be configured to measure anxiety according to the patient's keystroke patterns through the use of pervasive wireless networks; all of this information can be streamed to the nearest health care provider.⁶ Health care organizations can help offset the patient's cost of subscribing to a pervasive system by offering small rebates and discounts.⁶ These incentives would be based on a mobile device recording the patient making healthy decisions like exercising or eating nutritious meals.

Wireless devices are currently being used in the health care industry on a small scale.⁶ Many of the current systems implemented do not interoperate well. The creation and transition to a more pervasive health care infrastructure should not interfere with the basic operations of the current system.⁶ Consumers, health care providers, government entities, employers, and insurers have to be persuaded about the benefits of a pervasive system.⁶ Pervasive health care information would be subject to abuse by cyber-criminals or even corporations and insurance companies attempting to deny coverage to eligible individuals who may be considered too "risky." Strict and well-defined guidelines regarding the storage and access rights of this information are necessary to limit misuse.⁶

20.3 Medical Standards and Telemedicine

The lack of standardization in telemedicine raises a number of questions regarding the integrity and security of data transfers. Most interoperability issues can be traced to the lack of standardization in telemedicine.⁵ The definition of a standard is "a document, established by consensus and approved by a recognized body that provides, for common and repeated use, rules, guidelines, or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context" in Choi et al.⁵ Difficulty in creating a standardized telemedicine infrastructure could be linked to the disjointed use of standards in current medical systems.

An organization that creates standards for the health care industry must be accredited by the American National Standards Institute. There are many accredited

organizations, like Health Level 7; however, the U.S. Department of Health and Human Services (HHS) estimates that around 400 different formats are used for health care related transactions. Consider the coding of drugs and biologics where the National Council for Prescription Drug Program uses the National Drug Code but most institutions and professionals use the Health Care Financing Administration Common Procedure Coding System.⁵

Standards concerning health care providers include the National Provider Identifier (NPI) and the Health Industry Number. HHS established the NPI, a ten-digit number that, once implemented, will be the only identifier used in standard transactions. Based on the current growth of service providers the NPI will not have to be adjusted for the next 200 years. Other unique identifiers are being developed for both patients and physicians.⁵

The International Organization for Standardization made an attempt to solve the interoperability problem in telemedicine. The group concluded that the definition of telemedicine is too broad, and more collaboration is needed between the information technology and the telecommunications industries.⁵ Popular forms of telemedicine include “store-and-forward” and “two-way interactive.” The store-and-forward technique (as the name suggests) gathers some information, stores it, and passes it along to some other entity. With two-way interactive technologies each participant has a channel for communicating in real-time (e.g., videoconferencing).⁵ Telemedicine practiced at Johns Hopkins was used to perform surgery on a man in Bangkok, Thailand.⁵ As new medical standards like NPI are implemented, they may provide insight into how the use of telemedicine can be better structured.

20.4 Applications and Emerging Technologies

20.4.1 *The CodeBlue Infrastructure*

The resource limitations of sensors pose challenges to the design of applications for such devices. The CodeBlue software infrastructure is designed to help deal with these challenges.¹ CodeBlue integrates the sensors and other devices into a disaster-response setting. It has the ability to provide ad hoc network formation and other things like resource discovery and network security.¹ Initial testing of CodeBlue involved two wireless vital sign monitors and a PDA-based application for first responders as well as a radiofrequency-based localization system.¹

Sensor nodes are typically small devices with low power and low capability. A popular node design is the Mica2 developed at UC Berkeley. It uses a 7.3 MHz controller with 4 kbps of RAM and 128 kbps of ROM. The Mica2 includes a Chipcon CC1000 single chip radio capable of operating at 77 kbps within a range of 30 m. The device is not much larger than the two AA batteries needed to power it and can run for weeks at constant power.¹

Many security issues cannot be handled by traditional means because rescue and other emergency personnel should not be required to type in passwords or take time logging in during situations where every second counts.¹ CodeBlue allows for the seamless transfer of credentials, so one health care worker can provide access rights to another during patient transport. This handoff is possible in the absence of a pre-existing relationship between the workers and without the exchange of digital keys or certificates. Public encryption keys may be needed at certain times, so CodeBlue explores using elliptic curve cryptography (ECC). The ECC keys are smaller than their RSA equivalent and the current implementation can generate a key in 35 seconds.¹

The CodeBlue localization system, called MoteTrack, is a decentralized approach using radiofrequency beacons and single-chip sensor nodes. The system achieves an 80% accuracy for a density of 0.011 beacons/m². This is close to the accuracy provided by 802.11-based tracking systems but MoteTrack does not require a powered infrastructure. Cricket, an ultrasound-based system, achieves higher accuracy but requires a much higher beacon density. The MoteTrack system's decentralized architecture can tolerate a large number of beacon failures.¹

20.4.2 Heart Monitoring by CardioNet

Cardiovascular disease and stroke are the world's leading causes of death, causing more than 30% of all deaths worldwide.⁷ Many of these deaths can be prevented by reliable monitoring and response systems.⁷ The electrocardiogram (ECG) signal is the most commonly used means for monitoring the heart, and the 12-lead ECG is the most reliable. It uses twelve signals taken from nine different body sensors. This creates a large amount of data that must be recorded and filtered in parallel.⁷

Current heart monitoring systems are not able to provide high-level real-time computational analyses on the spot. Instead, data must be transferred to more powerful devices for processing.⁷ A heart monitoring system developed by CardioNet transfers ECG readings to a PDA where the readings are transmitted to a central server across a cellular phone connection.²

20.4.3 Tele-Ultrasonography with OTELO

Ultrasound scanning is a noninvasive method for providing clinical examinations.³ The growing acceptance of telemedicine has allowed tele-ultrasonography systems to emerge. These systems make it possible for medical professionals to work together across great distances. A major drawback of current tele-ultrasonography systems is that although the equipment is portable enough, an expert must be on-site to operate the machinery. Removing this restriction will bring more widespread use of this effective examination technique.³ A new medical robotic system, mobile tele-echography using an ultra-light robot (OTELO), is designed to allow a medical ultrasound expert to perform examinations on remote patients. OTELO is funded by the European

Information Society Technologies. Its portability is a product of the robot's increased degrees of freedom and its ability to use several different communication modes.³

The OTELO system consists of three parts: the expert station, the patient station, and the communication links.³ At the expert station, a medical worker operates a pseudohaptic fictive probe. The probe is used to control the positioning of the robot. It is designed to mimic the same type of ultrasound probes that medical experts are used to handling.³ The communication links create the bridge between the expert and patient stations. The robotic system at the patient station has six degrees of freedom.³ This can emulate human hand movement in the X, Y, Z, and diagonal directions.³ The robot is designed to be lightweight.³ The robot at the patient station operates the ultrasound probe based on the instructions transmitted across the communication links by the medical worker at the expert station.³

The ultrasound probe used with the OTELO system can send ultrasound still images back the operator at the expert station.³ The most important information being transmitted on the communication links are the robotic control data, the ultrasound still images, and the medical ultrasound streaming data.³ The medical ultrasound streaming data is the most demanding in terms of data-rate requirements.³ The remote robotic system and the expert station employ a force feedback system to aid in control of the fictive probe.³

Although most ultrasound scans do not require much communication between the patient and the examiner, the data stream must still give priority to vocal and even video communications.³ The amount of available bandwidth that is possible with 3G wireless technologies can support a simultaneous streaming of the data needed to perform the examination as well as any video and verbal communications between the patient and doctor.³ In some cases the available bandwidth cannot support this simultaneous streaming and so the system may need to establish a set interval for patient-examiner communications.³ At least 80% of the tests performed using remote communication methods have achieved results that are similar to those gathered from more conventional ultrasonography examinations.³

The 3G wireless communication platform can support either high or low mobility uses.³ The high-mobility platform can provide a data rate close to 150 kbps to a user traveling at around 120 km/h.³ Low mobility can provide more than double the data rate of high mobility but the user must travel at a speed under 5 km/h. Indoor 3G use has a 2 Mbps data rate.³ 3G wireless connections can provide a number of services including videoconferencing, video streaming, Internet browsing, and application sharing.³

Videoconferencing differs from video streaming in that there is two-way trafficking of the voice and video data between end users.³ Application sharing allows a process running on one terminal to make use of the processing resources provided by the server.³ The communication bottleneck is in the uploading at the patient station.³ The patient station must send the ultrasound still images, the ultrasound streaming data, the robot control data, and any other ambient sounds and video.³ OTELO traffic can be mapped to three of the quality-of-service classifications used by the Third Generation Partnership Project for Universal Mobile

Telecommunication Systems.³ The video would use the “streaming” class, which maintains time relations, and the “conversational” class can be used for the ultrasound images because of their real-time requirements.³

The European Information Society Technologies organization performed an experiment to measure the end-to-end system performance of OTELO over a 3G network.³ The experiment focused on average throughput and the end-to-end packet delay and jitter.³ The ultrasound scanner produced images at 13 fps. Video-conferencing frames were (320 × 240) pixels and frames using the Quarter Common Intermediate Format had (176 × 144) pixel resolution.³ Video compression was performed using the H263 codec, which is designed for video coding for bit rates as low as 20 kbps.³ The expert station sent 16 bytes of robot control data at 70-ms intervals.³ A laptop connected the patient station to the 3G terminal through a wireless network card.³ Tests were performed at different times and using different network load conditions.³ Using the Vodafone 3G network, both the patient and expert stations were in situated in London, England, but at separate universities.³

The 3G implementation of OTELO used the Real-Time Protocol and the User Datagram Protocol.³ The RTP is an Internet standard for real-time applications, and UDP is typically used when video streams are transmitted on wireless links.³ Robot control data was transmitted by UDP even though UDP does not provide reliable data transfer. Accurate robotic control at the patient station could be maintained as long as the minimal rate of packet loss was less than 0.5%.³ Although TCP provides reliable data transfer, the TCP retransmissions would have introduced unwanted delays.³

A complete ultrasound scan takes around four minutes.³ The tests showed that the expert station received 80% of its packets at 18.5 kbps and the other packets came in at over 50 kbps.³ Under good network conditions the expert side experienced no packet loss on the RTP stream, but loss reached 0.17% when the network was congested.³

The packets received at the expert station had a maximum delay of 0.3 sec; 50% of the received packets had 0.12-sec delay and 30% showed delays around 0.22 sec.³ As the delay increases video packets that are received beyond their play limit are counted as lost.³ Delay loss also ruins the ultrasound images by causing poor fps production.³

In networking, the round-trip time (RTT) is normally measured from the time a message is sent until the time an acknowledgment is received for that message. Delay is measured in a similar way for the OTELO system. RTT is seen between the time a robot control message is sent from the expert side and the time the patient side responds with ultrasound stream images.³ Packet sizes of 100, 200, 300, 500, 1000, and 1400 bytes can be generated by the H263 video codec and all sizes were tested for delay performance.³ Using the above RTT model, the maximum delay was found to be around 325 ms.³

20.4.4 Biomedical Sensor Networking Platforms

Many of the recent developments regarding the use of wireless sensor networks in pervasive patient monitoring systems have been for cardiac patients.² Commercial

systems like the CardioNet mentioned earlier can transmit signals to a PDA, where they are then routed to a central server for review.² Another wearable monitoring system called MIThril, developed by Pentland, uses a PDA to capture ECG data, GPS position, skin temperature, and galvanic skin response.² Still most of the wireless networks developed today are designed for tracking or environment monitoring applications. Such systems include Intel's iMote, UCC's DSYS25, and also Berkeley's Mica2 and Telos.² Context-aware sensing platforms like SmartITs and MITes exist but the integrated sensor design means that major platform changes are needed to support the inclusion of a physiological sensor.²

A flexible operating system is needed to allow for research extensions.² A biomedical sensor network that was designed and implemented to address these concerns is presented in Lo et al.² The BSN hardware platform has a stackable design and can accommodate the inclusion of many different types of sensors.²

Wireless communication in the system uses the 802.15.4 standard from IEEE and can handle continuous context and physiological sensing.² Patient status is monitored with wireless physiological sensors but relying solely on these devices can cause false detections.² Changes in motion and stress artifacts can suddenly change heart rate, and it is important to distinguish between changes due to things like exercise and issues like arrhythmia.² Sensors that are context aware are used in the BSN to identify the clinically relevant situations.² Like other wearable wireless monitoring systems the BSN uses a PDA to combine the data from the various sensors and relay the information to a central server.²

Nodes in the BSN use Texas Instruments' (TI) 16-bit, low-power Reduced Instruction Set processor with 60 kb of flash memory, 2 kb of RAM, 12-bit ADC, and 6 analog channels.² The wireless module has a 50-m range and a throughput around 250 kbps.²

It uses 512 kb of flash memory for data storage.² Nodes run the TinyOS developed by UC Berkeley.² TinyOS is an open source operating system that incorporated efficient energy usage and includes a modular set of software-building blocks to aid in code development.² Files built into TinyOS are as small as 200 bytes.² In this BSN the operating system takes sensor measurements, makes routing decisions, and controls power usage.² This leaves TinyOS in charge of both the hardware and wireless network operations.² Because the BSN used the low-power TI microcontroller, it only requires 0.01 mA of power in active mode and 1.3 mA when it is performing its most computationally demanding calculations.² Nodes are 26 mm, which makes them practical as wearable devices.² For this BSN architecture there are several wireless biosensors. This includes three-lead ECG, a two-lead ECG strip, and SpO₂ sensors.² Also included are accelerometers and temperature and humidity sensors, which are used to provide context information for the biosensor data.² Context detection is based on a TSOM and Bayesian framework.² A compact flash card is used with a PDA, so sensor signal can be gathered and reviewed with the PDA and routed to a central server if needed.² Transmitting the data would use a WiFi/GRPS network, and once it reaches the server it can be used for trend analysis and placed in long-term storage.²

20.4.5 *Next-Generation Network-on-Chip Protocol*

Another type of biomedical sensor network is being developed to discover if nano-single-chip solutions can be used effectively in computationally intensive biomedical applications.⁷ Like many of the other architectures mentioned, monitoring the activity of the human heart is the main purpose of this platform.⁷ The importance of this issue comes from not only the large number of deaths attributed to cardiovascular disease and stroke, but also the fact the death rate is growing every year.⁷ Cardiovascular disease is also indiscriminate and unrelated to either gender or ethnicity.⁷ The architecture presented here uses the electrocardiogram signal because it is common in current practice and it is one of the most reliable techniques for revealing heart malfunctions.⁷

One of the challenges in this BSN is to analyze the very large amounts of ECG recorded data in parallel while keeping up with the hard-time deadlines of the real-time requirements needed in life-critical systems.⁷ Another issue involves the mismatch between the analysis techniques and the relatively high sampling frequency capable with most state-of-the-art sensors.⁷ Modern sensors do lack the ability to provide high-computational, real-time analysis at the patient's location and this situation worsens for patients with mobility.⁷ Also difficult to achieve is a wireless connection that is always functional and always connected but such features are important to life-critical applications because the related data cannot afford to be lost.⁷

The solution presented here will parallel-process the large biomedical computation of the 12-lead ECG on a wearable nano-multiprocessor network-on-chip (NoC).⁷ The chip can transmit the full, raw ECG data when it is requested and the chip also is designed to send regular reports and secure alert signals when needed.⁷ Raw electrocardiogram data has sizes in the megabyte range but the other reports can be handled with just a few bytes.⁷ The implementation follows a biomedical requirement that sets the data chunks for the heart readings at 4-second recordings.⁷ Processors are commercial very-large-instruction-word DSPs and are connected to off-the-shelf biomedical sensors.⁷

State-of-the-art biomedical sensors are characterized by increasing energy efficiency allowing longer lifetimes and higher sampling frequencies.⁷ Many of the sensors available today can also connect and run wirelessly.⁷ The application presented in Khatib, Russo, and Naviev⁷ makes use of two types of networks. There is a sensor network located on the patient's body and an NoC, where the sensor signals are analyzed.⁷ The twelve-lead electrocardiogram produces data that can reach sizes up to hundreds of megabytes every hour.⁷ Physicians use the twelve-lead ECG not only because of its efficiency but also because this method allows the heart to be viewed in a three-dimensional form.⁷ With the heart viewable in three dimensions, it is possible to detect an abnormality and also pinpoint its location or position on the heart itself, giving increased insight towards treatment options.⁷

Each of the leads can sense data independently.⁷ In a 12-lead, each of the 12 signals can be assigned to a different nanotechnology processor of the NoC

multiprocessor.⁷ With this model the number of leads could be increased beyond 12 by adding more processing elements in the interconnected NoC platform.⁷ Reading data continuously every 4 seconds emulates a sensor sending continuous data to an intermediate buffer that holds 4 seconds of data sampled at 1 KHz.⁷ Experiments run on this NoC platform were tested using sampling frequencies at 0.25, 1, and 10 KHz.⁷ It can be noted that current noncommercial biomedical sensors reach their maximum sampling rate around 10 KHz.⁷ The application program module uses an autocorrelation function to calculate the period of the heartbeat. The function is used to show how much the heart signal resembles itself after a certain time lag.⁷

More than one network-on-chip is needed to analyze the signals and give reports. There is a need for an efficient set of algorithms and protocols to interconnect many different NoCs.⁷ Consider the network-on-chip a self-governing system with private internal policies that are designed and implemented by the manufacturer.⁷ This NoC autonomous system uses its own unique address that is defined to be the NoC hardware address.⁷ Addressing the NoC in this way allows for two types of communication: (1) intra-chip communication, where some specific core p wants to interact with some specific core q on the same NoC; and (2) inter-chip communication, where a core p on one NoC can interact with some application located on a different NoC autonomous system.⁷

Several different protocols exist for both the inter- and intra-communication mechanisms.⁷ One principle in intra-NoC communication is that different NoCs can be connected either physically or wirelessly but should always use a main and a backup connection.⁷ The interoperability protocol developed for this particular NoC platform was created with C++.⁷ Its convergence time for identifying issues goes as high as 13.69 ms.⁷

20.4.6 Wireless LANs and Patient Monitoring

The effective monitoring of a patient must perform periodic transmissions of vital signs and alert signals when reading across thresholds. Vital signs used by these systems should include (but not be limited to) heart rate, blood pressure, and body temperature.⁸ Patient monitoring requires comprehensive wireless networks that must be reliable, scalable, secure, and fast. These networking systems must be able to utilize the network infrastructure efficiently so that a mobile monitoring device can operate in public or private wireless networks and ad hoc modes as well.⁴ An effective patient-monitoring solution must include a location tracker, and the ability to deal with interference and message conflicts.⁴ A message priority scheme must be implemented to ensure that emergency alert messages experience minimal delay due to network traffic.⁸

Wireless systems like Bluetooth or IEEE 802.11 can provide data transfer rates from several hundred kbps up to a couple of dozen Mbps.⁸ Bluetooth networks would limit the number of users per piconet to eight. To allow for more

users/patients, multiple piconets may be used. Increasing the number of piconets also increases the amount of interference in ISM band.⁸

The ISM bands are reserved for unlicensed commercial use and are used in taxi services, police cars, CB radios, and many other areas. The most popular frequency is the 2.4-GHz range, which is also used by most wireless LANs. To limit interference and increase data transfer rates, patient-monitoring systems should consider using a wider bandwidth for their WLAN.⁴ One issue is the large number of patients that may be linked to a single provider. The amount of data generated by the frequent and necessary messages of a monitoring system may be overwhelming.⁴ Adding more machines helps to put off large computational loads on computers, but a different solution is needed for the physician who simultaneously receives the heart rate readings for dozens of patients.

The service area of most wireless access points is around 100 m and is affected by mobility and obstacles.⁴ Simple doors, walls, or windows can lower a wireless signal by 30 to 90% as it passes through.⁴ Link quality varies over time and links can and will fail occasionally so that absolute uninterrupted connectivity cannot be guaranteed and the system-monitoring resources will need to cope with this without issuing premature emergency alerts.⁴

The throughput for a wireless link is affected by things like distance and the number of users on a single access point. This makes it necessary to provide overlapping wireless coverage that will also be spread over a wide area. The number of users per access point can also be limited to provide for better access.⁴ Multiple access points in close proximity can improve connectivity but having several different wireless LANs in the same area can cause interference issues. There is currently no way to claim an area for a particular LAN.⁴

While a wireless signal is in transit it may be reflected off various objects and the receiver would get multiple copies of the same signal. The echoed signals are slightly shifted in time and overlap with the original, making decoding more difficult. Higher data-speed rates mean smaller bits, and as the bits become smaller the echo shift interference becomes more significant. Patient-monitoring wireless LANs must also contend with interference from microwave ovens and radar transmitters.⁴ Most wireless LANs use Collision Sense Multiple Access with Collision Avoidance. This protocol uses different types of messages and each message has to wait for the channel to be free before it can be transmitted. The wait time depends on the type of message being sent, and a lower wait time can be seen as a higher priority. This scheme can be used to minimize the delay on emergency alert messages.⁴

20.4.7 Un-cooperation of Routers in Wireless Patient Monitoring

All of the platforms introduced thus far have assumed that all of the routers are cooperative, functioning properly, trustworthy, and supplied with enough power to

find another device that can act as a router for forwarding messages to one or more health care institutions.⁹ Approaches to more realistic routing scenarios involving ad hoc networks designed for patient monitoring are presented in Varshney.⁹ In real life, routers can be trusted to work forever or always be cooperative.

The end-to-end delivery of messages carrying vital signs of patients will be negatively affected by the failure or uncooperation of routers.⁹ The impact of uncooperative nodes in wireless patient monitoring can be minimized by using more persistent routing schemes.⁹ It is important to identify ways to reduce or eliminate the potential for router uncooperation to ensure the safe monitoring of patients. Uncooperation between routers can be due to several different issues. Some patients may just turn off their devices that could act as routers for others or simply may not want their devices to act as routers for others.⁹ Some devices may experience failures related to hardware, batteries, and network access failures.⁹ Some devices may be programmed to save energy and may thus be unwilling to route packets for others in an ad hoc situation.⁹

Overcoming uncooperation can be handled in several different ways. We can use persistence in transmissions involving the devices, which may eventually transmit at a reduced power level if the device mobility results in a reduced distance to the next-hop.⁹ The routers can treat noncooperations as reduced device density for routing purposes.⁹ The networking protocols can use reliable multicast and reliable broadcast based routing schemes.⁹ The system could also be designed to detect and exclude noncooperative devices when planning routing paths as in the formation of a multicast tree.⁹

In multicast routing, patient information is sent to multiple but not all health care professionals. This requires creation of multicast tree or structure but the reliability of message delivery is significantly enhanced.⁹ Broadcast routing, where patient information is sent in all directions to all nodes, will lead to the best reliability of message delivery, but the resulting network traffic can be excessive. Reliable multicast and reliable broadcast utilize a persistent transmission of patient information, and thus these can lead to a significantly better reliability performance than multicast and broadcast and thus may be more suitable when failed or uncooperative devices exist in the network.⁹

The ad hoc networks are autonomous and decentralized networks where multiple different entities may be involved. Some routers could be selfish, or misbehaving, as far as routing of messages are concerned.⁹ These routers may not want to forward messages from other devices as this consumes some of their energy and processing resources.⁹ Some routers may be inclined to go to energy-conservation mode, and thus not be able to receive and route messages.⁹ Energy levels of different nodes may be used in deciding best routes, and some selfish nodes may under-report their current levels of available power to receive fewer messages for routing.⁹

Techniques for reducing noncooperation among routers include offering incentives, utilizing social considerations, and simply forcing routers to cooperate.⁹ It may be necessary to make proper routing a prerequisite for maintaining membership in

a patient-monitoring ad hoc network.⁹ Member devices that do not route a certain number of patient-monitoring messages would consequently be removed from the group.⁹ To encourage cooperation from routers, incentives can be offered.⁹ The incentives could range from payments for routing to higher priority for their future messages.⁹ Payments for routing cooperation could include processing a node's networking resources, which can be used later for transmission of their own packets.⁹ Priority-based forwarding, where messages from a device are forwarded based on the history of cooperation from that device⁹ and a router forwarding an emergency message will receive a considerable high priority for future transmission.⁹

The addition of incentives might lead to competition among routers for forwarding certain patient monitoring messages.⁹ Devices that recently joined an ad hoc network would be given a default priority and devices that fail and rejoin would have a lowered priority to act as a deterrent against devices that claim a false failed status to avoid routing patient monitoring messages.⁹ Also social incentives can be utilized. A device can be set up to route messages of your friends.⁹ Some patients may be more willing to have their devices act as routers for forwarding messages as an act of charity.⁹ Factors like reputation can also be used to obtain cooperation from routers.⁹ There has been work done on reputation-based systems, where nodes detect selfish nodes by using second-hand information and then work to isolate the selfish nodes from the network.⁹ Reputation-based schemes must be careful that rumors and false readings do not result in the isolation of a cooperating node.⁹

These are some of the assumptions that are used by this system in order to ensure the tractability and manageability of the model.⁹ The end-to-end path between patients and health care professionals is based on pure ad hoc networks. This is more complex than that in practice where a combination of wireless LANs and ad hoc networks could be used, and thus it is likely to underestimate the end-to-end performance.⁹ Also all devices have equal power budget, or range, and processing power. This simplifying assumption will overestimate the performance as devices are likely to be diverse in power budget.⁹ Furthermore, the devices are uniformly distributed in the service area.⁹ This is a simplifying assumption as in real-life more patients will be clustered.⁹ Finally, the uncooperative and failed routers are randomly distributed and lead to a reduced density of routers for message forwarding.⁹ In practice, failure and uncooperation may be co-related.⁹

The parameters used in deriving performance results for 4 routing schemes were 300 patients, 3 health care professionals, and 18 cooperating devices.⁹ For multicast based routing, the number of routes was 9 and for broadcast-based routing the number of routes was 45. The number of attempts used by reliable broadcast and reliable multicast was two. The reliability of message delivery was measured by varying the level of noncooperations and failures.⁹ Higher degrees of failures and noncooperation affect the reliability of message delivery for multicast routing.⁹ It was assumed that failures and noncooperation are mutually exclusive, where these two factors will have a cumulative effect on the probability of message delivery.⁹

When vital signs are in the normal range, 98% reliability is acceptable but the reliability must be improved for handling emergency situations.⁹ When the number of misbehaving routers is decreased, the reliability of the ad hoc system approaches 100%.⁹ The test showed that persistent routing techniques were able to overcome a high degree of router failures.⁹ Four schemes were used and it was observed that it may be useful to allow devices to switch between schemes to create a more balanced system.⁹

20.5 Conclusions

Recent advances in technology have opened up ways to improve upon the health care system. The use of telemedicine will provide higher-quality service and increased efficiency to the practice of medicine. Emergency and critical response professionals can be given immediate access to a wealth of vital information. Having accurate patient information on hand will save time and reduce the number of medical errors. Telemedicine will also make it easier to provide medical services to remote locations. There are several new services being developed for use in the field of telemedicine. CodeBlue is an infrastructure that can provide first responders with tools that will raise their success rate and make their jobs safer. Accurate heart monitoring services are being developed by CardioNet. This technology will help reduce the impact of the number one cause of death in the world. Ultrasound is used as a noninvasive and safe means of performing medical diagnoses. The OTELO robot will remove the need for ultrasonography to be performed by an on-site expert and such diagnoses can then be provided to a wider range of individuals. Patient monitoring is a necessity and keeping a watchful eye on a large number of patients can be physically overwhelming. Wireless networking devices that are readily available can make the process much more efficient.

With the development of new products from telemedicine, there is a growing need to develop standards for the field. Standards will improve reliability and allow the interoperability of the various different services being created. Current medical standards for things like prescription medicine are so complex and disjointed that it becomes even more difficult to develop telemedicine standards. New medical standards are being implemented by the U.S. government and will help to structure the use of information better across the health care system.

Health care services play an important role in the lives of most of the world's population. The need for constant, reliable service has created a situation where service costs rise at high rates while service quality tends to decline. The industry makes efforts to provide service to more individuals using fewer qualified professionals. Telemedicine is a means of lowering costs and expanding coverage without losing health care quality.

Acknowledgment

The work is partially supported by the U.S. National Science Foundation (NSF) under grants CNS-0716211 and CNS-0716455.

References

1. K. Lorincz, D.J. Malan, T.R.F. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnyder, G. Mainland, M. Welsh, and S. Moulton, Sensor networks for emergency response: Challenges and opportunities, *IEEE Transactions on Pervasive Computing*, 3, 4, 16–23, 2004.
2. B.P.L. Lo, S. Thiemjarus, R. King, and G. Yang, Body Sensor Network — A Wireless Sensor Platform for Pervasive Healthcare Monitoring, Dept. of Computing, Imperial College of Science, Technology and Medicine, London, U.K.
3. S. Garawi, R.S.H. Istepanian, and M.A. Abu-Rgheff, 3G Wireless Communications for Mobile Robotic Tele-Ultrasonography Systems, *IEEE Communications Magazine*, 44, 4, 91–96, 2006.
4. U. Varshney, Patient monitoring using infrastructure-oriented wireless LANs, *International Journal of Electronic Healthcare*, 2, 2, 149–163, 2006.
5. Y.B. Choi, J.S. Krause, H. Seo, K.E. Capitan, and K. Chung, Telemedicine in the USA: Standardization through Information Management and Technical Applications, *IEEE Communications Magazine*, 44, 4, 41–48, 2006.
6. U. Varshney, Pervasive Healthcare, *IEEE Communications Magazine*, 36, 12, 138–140, 2003.
7. I.A. Khatib, G. Russo, and R. Nabiev, Performance Analysis of Interoperability Protocols and Algorithms in Networks-on-Chip for the Next Generation Biomedical Sensor Networks: A Study on Human Heart ECG Monitoring and Analysis via Nano-Multiprocessor-Networks, San Marco Project Research Center.
8. U. Varshney, Transmission of Emergency Messages in Wireless Patient Monitoring: Routing and Performance Evaluation, presented at the 39th Hawaii International Conference on System Sciences, 2006.
9. U. Varshney, Addressing Un-cooperation of Routers in Wireless Patient Monitoring, presented at the 19th IEEE Symposium on Computer-Based Medical Systems, 2006.
10. A. Ganz, R.S.H. Istepanian, and O.K. Tonguz, Advanced mobile technologies for health care applications, *Journal of Mobile Multimedia*, 1, 4, 271–272, 2006.
11. U. Varshney, Using Wireless Networks for Enhanced Monitoring of Patients, presented at the 10th Americas Conference on Information Systems, New York, August 2004.
12. U. Varshney, Enhancing Wireless Patient Monitoring by Integrating Stored and Live Patient Information, presented at the 19th IEEE Symposium on Computer-Based Medical Systems, 2006.
13. U. Varshney and S. Sneha, Patient Monitoring Using Ad Hoc Wireless Networks: Reliability and Power Management, *IEEE Communications Magazine*, 44, 4, 49–55, April 2006.

Index

- 869 MHz European social alarm frequency, 18, 19
 - 3G, vii, 113, 176–178, 186, 271, 276, 297–315, 341–342, 375, 384–387, 395–396
 - Long-term evolution (LTE), 342
 - 3.5G High Speed Downlink Packet Access (HSDPA), 315, 342
 - 3.75G, *see* High Speed Uplink Packet Access
 - 4G wireless networks, 267–293, 298
 - adaptive resource allocation, 273
 - ehealth network architecture, 280
 - heterogeneous environment, 272
 - multiple types of services, 273
 - in weCare-Next-generation Integrated Wireless Tele-health system, 66
 - wideband CDMA, 68
- A**
- Abbot Freestyle Navigators, 124–125
 - Abdominal swelling, 101
 - in CHF, 93
 - Acceleration sensors, 22, 23
 - Access control list
 - in IEEE 802.15.4, 218
 - in design of security policies, 223–224
 - control of, 223
 - Accuracy of telemedicine equipment, 356
 - Active energy expenditure, 19
 - Activity sensors, 24–25
 - for elderly diabetics, 166
 - Ad hoc networks
 - in telecardiology, 67–68
 - in CodeBlue infrastructure, 393
 - in wireless patient monitoring, 401–403
 - Adapting data, 327–328
 - Adapting functionality, 326
 - Adaptability of mobile computing, 323
 - how to develop, 329
 - mechanism, 326
 - state-based approach, 330
 - using proxies, 331–333
 - Admission control, 274, 284, 285, 287, 290–291
 - Advanced encryption standard (AES), 183
 - Advanced Health and Disaster Aid Network (AID-N), 65
 - Adverse events, 30
 - Aggregation, in design of security policies, 224
 - Agility, 328–329
 - Alarming systems, 3–5. *See also* Personal supervision systems
 - Aldosterone antagonists, 88
 - Altitude sensors, in horizontal position determination, 16
 - “Always on” IP-based services, 108
 - diabetes care requirements, 150
 - AMON approach, 5, 20, 26
 - infrastructure requirements, 22
 - sensors and clinical results, 21–22
 - wrist devices, 20–21
 - Angiotensin-converting enzyme (ACE) inhibitors, 88
 - Angle of arrival sensors, 15
 - Animas insulin pumps, 132
 - Ankle swelling, 101
 - in CHF, 93
 - “Anywhere anytime” cardiac care, 66, 67
 - Application-aware adaptation, 325
 - Application transparent, 325
 - Arrhythmia detection, 108
 - Arrhythmia disorders, 106. *See also* Cardiac arrhythmias

Arrhythmia monitoring system (AMS), 107–109

Arterial distensibility, 8

Artificial intelligence, use with elderly diabetics, 166–167

Artificial pancreas devices, 128, 139
 always-on technology requirements, 150
 progress toward ambulatory, 148–149

Asymmetric key encryption, 187

Attacks on medical wireless sensor networks, 218–221
 sybil attack, 219
 sinkhole attack, 219
 black hole attack, 219
 gray hole attack, 219
 wormhole attack, 219
 sleep attack, 219
 fairness attack, 219
 denial of service (DOS) attack, 219
 attack on biomedical sensors, 220
 attack on PAN coordinators, 220

Atherosclerosis, detection by PWV, 8

Atrial fibrillation, 105

Attribution, in design of security policies, 224

Authentication page, mobile medical records system, 55

Automated alarm systems
 in diabetes care, 148
 for elderly person surveillance, 6

Automated insulin injection control, 131
 closed loop control, 134–138
 Diabetes Advisory System (DIAS), 134
 partially closed loop control, 131–134

Automated voice reminders, for elderly diabetics, 164

Autominder system, 166–167

Availability, 187–188

B

802.11b standard, 107

B-type natriuretic peptide (BNP), 91

Balance, and security, 191

Bandwidth allocation, 285, 287, 289, 291

Bandwidth issues, with motes, 111, 113

Bankruptcy game, 287
 characteristic function, 288
 payoff vectors, 288

Basal insulin levels, 131

Battery power, 24
 for blood pressure measurement, 11
 for mobile medical records system, 55

Bayesian learning
 in automated insulin injection systems, 137
 in DIABTel system, 153

Beacon-enabled clusters, 217–218

Bed/chair occupancy sensor, 19, 24

Beta blockers, 88

Bidirectional communication, for integrated alarm monitoring systems, 36–37

Biocompatibility issues, for wearable devices, 10

Biomedical sensor networks, 397–398
 hardware, 213–216

Biosensor shirts, 106, 109

Bit errors, in mote applications, 112–113

Block ciphers, 199
 choice, 77

Blood glucose concentration, relationship to insulin concentration, 135

Blood glucose management
 automated insulin injection control, 120, 131
 automated micro-sensor techniques, 119–121
 glucose microsensors for, 121–126
 injection methods, 127–131
 insulin pumps for, 126–127

Blood oxygen saturation (SpO₂), 9, 16, 21, 69, 70, 86
 sensors, 71
 transmissive *vs.* reflective PPG measurement for, 11

Blood pressure monitoring, 8, 16, 69, 89, 90, 105
 blood pressure change, 70
 OMRON home-use devices for, 18
 sensors for, 72

Blood values, 9

Bluetooth technology, 18, 108, 113, 180
 in diabetes care, 150, 155
 use in AMS prototype, 107–108

Body area networks (BANs), 109

Body sensor networks (BSNs), 196
 Bio-channels in, 197
 topology, 197–198
 security challenge, 198

BodyMedia lifestyle monitoring, 19

Bradycardia, 70, 105, 111

Breathing sensors, 72

C

- Caching, 332
- Call center service, in AMS prototype, 107
- Camera-based location tracking, 15
- Consent and notification, in design of security policies, 223
- Capacitive sensors, 12
- Card and pin authentication, 55
- Cardiac arrest, detection of, 7
- Cardiac arrhythmias, 105. *See also* Arrhythmia disorders
- Cardiac impedance, 91, 92
- Cardiac monitoring
 - central server, 109
 - first-generation prototype, 107–109
 - issues with motes, 112–113
 - packet size limits, 113
 - preliminary data collection results, 111–112
 - second-generation prototype, 109–111
 - with sensor networks, 103–105
 - telemedicine for cardiac health, 105–107
- Cardiac output, 92
- Cardiac packet attacks, 80
- Cardiology, ix, 61
 - CHF monitoring and management, 85–87
 - mobile tele-cardiology, 63–65
 - personal cardiac monitoring with sensor networks, 103–105
- CardioNet, 65, 394, 397, 403
 - biomedical sensor networking platform, 397–398
- Cardiovascular diseases, 64, 105
- Care Coordination approach, 169
- Caregiver's Assistant, 182
- CDMA200 standard, 18
- Cellular network, 271
- Cellular technologies, 67
 - integration with ad hoc networks, 68
 - in patient monitoring, 18
 - unreliability for telecardiology, 65
- Center for Aging Services Technologies (CAST), 180
- Central trusted security servers (CTSS), 218, 226
- Chen, Hui, xiii
- ChipOx module, 11
- Chronic diseases
 - diabetes as paradigm of, 144
 - increased rate of, 104
- Cipher block chaining (CBC), 199
 - mode of operation, 199
- Client-side intercept (CSI), *see* CSI
- Client-side proxy, *see* CSI
- Clinical confidentiality and integrity, 222
- Closed loop control solutions, 134, 150
 - in diabetes care, 148
 - model predictive control, 137
 - neural network control, 138
 - pole-assignment control, 134–136
 - self-tuning adaptive control, 136–137
- Clothing-integrated ECG, 17
- Cluster-based security, 78
- CODA, 327, 330
- CodeBlue project, 65, 73, 393, 403
- Cognitive impairments, 163
 - continuous glucose monitoring for, 165
 - in elderly diabetics, 162
- Cognitive orthotics, for elder diabetes care, 164–167
- Cognitive radio techniques, 275, 282–284
- Color coding, use in mobile medical records system, 56
- Commercial glucose sensors, 123, 126
 - specification comparisons, 124–125
- Commercial insulin pumps, 131
 - specification comparisons, 132–133
- Communication history, 45
 - in integrated alarm monitoring systems, 44, 46
- Communication protocol, converting from RS-232C to TCP/IP, 34
- Compliance, and security issues, 191
- Compression fidelity, 299
- Computer-supported collaborative work (CSCW), 148
- Confidentiality
 - and availability issues, 187–188
 - ethical issue, 356–357
 - in diabetes care systems, 156
 - in ECG transmission, 75
 - in mobile telemedicine, 183–187
 - legal issues, 356–357
 - with paper *vs.* digital patient records, 51
- Congestive heart failure (CHF), 87–88, 105
 - aggressive monitoring, 92
 - architecture for patient monitoring application, 93
 - Framingham criteria, 89
 - humidity variation monitoring, 98, 99
 - level 1 monitoring, 89
 - level 2 monitoring, 89–91

- level 3 monitoring, 92–93
- level 4 monitoring, 92
- medications for, 88
- monitoring and management, 85–86
- monitoring example, 99–101
- monitoring stages, 90
- and need for head raising pillow, 89
- preparation for intervention, 92
- RSSI variation for bedroom sensors, 95–96
- sensor networks for remote monitoring, 86–87
- smart monitoring system, 92–101
- stages of, 88–92
- symptoms diary, 93
- Connecting for Health (CFH) initiative, 52
- Context-aware pill bottles, 165
- Continuous blood pressure measurement, 8
- Continuous glucose monitoring, 148
 - in elderly diabetics, 165
- Controlled flooding, 68
- Conventional telemedicine, 369–371
 - problems of, 371
- Cooperative game-theory framework, 285
- Coordination, and security issues, 191
- COPD, remote monitoring of, 101
- Coronary artery disease, 64
- Cost containment, with mobile telecardiology, 64
- Cost-effective Portable Telemedicine kit, 211–212
- Cough, in CHF, 93, 101
- CRC ITU-T, 36, 37
- Crossbow Inc., 73
 - MIB 510 base station, 110
- Cross-layer design, 315
- Cricket, 394
- Cryptographic primitives, in body sensor networks, 198–202
- CSI, 332
- CS model, 327
 - Impact of mobility on, 327
- Cyclic redundancy check (CRC), mote
 - limitations, 112–113
- Cygnus Corp, 122
- Cywin, 111

D

- Daily activity diary, in CHF monitoring, 97
 - da Vinci™ robotic system, 355
- Data agility, 328–329
- Data collection, from medical devices, 39
- Data collection interface (DCI), 30
 - for integrated alarm monitoring systems, 34–36
- Data compression, 377
- Data fidelity, 327–328
- Data format, for integrated alarm monitoring systems, 36–37
- Data integrity
 - issues in mobile telemedicine, 183–188
 - in mobile telecardiology, 75
- Data link-level security, 185
- Data mining techniques, 86
- Database server, for mobile medical patient records, 52
- Dehydration monitoring, 8
- Deltec insulin pumps, 132
- Denial of service (DoS) attacks, 187
- Device design guidelines, 9–10
- Device errors, 31
- DexCom, 124–125
- Diabetes, ix, 117
 - automated blood glucose management in, 119–121
 - gestational, 146
 - improving glycemic control among elderly with, 161–163
 - incidence, 144
 - long-term complications of, 144, 146–147, 162
 - mobile multi-access telemedicine workspace, 146
 - mobile telemedicine for, 143–145
 - as paradigm of chronic disorders, 144
 - type 1 *vs.* type 2, 120, 146
- Diabetes Advisory System (DAS), 120, 121, 134
- Diabetes care
 - ambulatory artificial pancreas in, 148–149
 - Autominder system, 166–167
 - balancing food and insulin intake in, 164
 - Care Coordination approach, 169
 - challenges in, 145–149
 - cognitive orthotics in, 164–167
 - compliance problems, 163
 - concise content requirements, 149
 - current process recuperation requirements, 149
 - daily *vs.* weekly monitoring results, 169–170

- DIABTel distributed architecture, 151–152
 - handheld dietary device efficacy, 164–165
 - hospital admission rates, 170
 - interface requirements, 149
 - keystroke reduction requirements, 149
 - media adaptability requirements, 149
 - medical problem, 145–146
 - mobile applications in DIABTel system, 152–155
 - mobile telemedicine for, 143–145
 - mobile telemedicine system components, 150–155
 - modular design requirements, 149
 - multi-channel messaging services for, 148
 - navigation requirements, 149
 - PEARL system, 166–167
 - technical requirements of mobile telemedicine systems for, 149–150
 - technological reminders for, 163–164
 - telemedicine and shared care services in, 147–148
 - treatment options, 145–147
 - virtual medical offices, 167–171
 - Diabetes control and Complications Trial Research Group (DCCT), 162
 - DIABTel system, 150–151
 - application server agents, 151
 - communication server agents, 151
 - DIABTelMobile application, 155, 156
 - distributed architecture, 151–152
 - event messages in, 152
 - mobile applications in, 152–155
 - multi-access organizer, 152
 - patient electronic logbook, 153
 - PDA Smart Assistant application, 153–154
 - WebPDA component, 154–155
 - Diagnostic accuracy, 299
 - Diastolic heart failure, 88
 - Diastolic pressure, 92
 - Dielectric spectroscopy, 121, 122
 - Differencing, 332
 - Digital patient records, 50
 - Digital signature schemes, 228
 - Digoxin, 88
 - Dimensions of data fidelity, 328
 - of spatial data, 328
 - of telemetry data, 328
 - of video data, 328
 - Disetronic insulin pumps, 132
 - Distance measurement, 15
 - Diuretics, 88
 - Documenting errors, 32
 - Door opening sensors, 20, 24
 - Doppler flow pulse, 17
 - Dropped packet ratios, in cardiac monitoring prototypes, 112
 - Drug interactions, complications from, 105
 - Dynamic routing scheme, 68
 - Dyspnea, 88, 89, 91, 93, 101
- ## E
- E-health, 350
 - E-health services, 275
 - types of, 275–277
 - follow-up service, 275, 278
 - intra-hospital monitoring, 276, 279
 - mobile health care service, 276
 - medical information management service, 277
 - prehospital service, 276, 279
 - telehomecare service, 276, 279
 - ECG data collection, 71, 86
 - ECG Halters, 105, 107
 - ECG microsensors, 73–75
 - ECG sensors, 71, 215
 - three-lead, 73
 - ECG signal waveform, 75
 - ECG technologies, 10, 22, 64
 - secure transmission, 75–80
 - security analysis, 80
 - sensors for patient monitoring, 17
 - heart monitoring, 394
 - ECG throughput, *vs.* patient density, 69
 - ECG transmission, 341
 - ECG transmission delay, 69, 75
 - Echocardiograms, with WSNs, 182
 - Edema, 88
 - Ejection fraction, 92
 - Elderly person surveillance, 6, 23
 - and benefits of aging in place, 171
 - challenges with computer learning, 168
 - and changes in mental processing speeds, 167–168
 - cognitive orthotics, 164–167
 - PEARL system, 166–167
 - technological reminders and sensors for, 163–164
 - telemedicine for elderly diabetes, 161–163
 - Tunstall solution, 25–26
 - and virtual medical offices, 167–171

Electrical skin impedance spectroscopy, 13
 Electrocardiogram (ECG), 9, 105
 noncontact electrodes for, 107
 Electroencephalogram (EEG) sensors, 71–72
 Electrolyte balance, across cells, 122
 Electromyogram (EMG) sensors, 71
 Electronic reminders, 165
 Elliptic Curve Cryptography (ECC), 394
 EMERGE approach, 5, 22–23
 activity sensors, 24–25
 environmental sensors, 25
 infrastructure requirements, 25
 miscellaneous sensors and signs, 24
 weight sensors, 25
 wrist device and integrated sensors, 23–24
 Encryption standards, 187
 in WLANs, 183
 Energy efficiency, in medical security, 76
 Enuresis sensor, 19
 Environmental sensors, 25
 Epilepsy sensor, 19
 Error detection, with CRC, 37
 Error rates, *vs.* output power levels, 112
 Error reporting
 automation of, 33
 transparent, 32–33
 Error Reporting System (ERS), 30
 Ethical issues, ix, 355
 confidentiality 356–357
 Evaluation questionnaire, mobile medical records prototype, 58
 Expiratory tidal volume, 40

F

Fall detection sensors, 12, 19, 20, 23
 False alarms, 31
 False signals, 15
 Fatigue, in CHF monitoring, 101
 Fidelity, 299, 327–328
 Fine-grained medical data, motes for, 104
 Finger clips, 11
 Finger rings, 11
 Finger sticks, 121
 First responders
 data security issues, 180
 protecting communication among, 183
 use of sensor networks by, 181
 Flooding attacks, 187
 Fluid intake, monitoring of, 89

follow-up service, 275, 278
 Forrester Research, 64
 Framingham criteria, 89
 Frequency hopping, 108
 Fusion approach, 5
 to patient monitoring, 26

G

3G cellular technology, 177, 179, 341
 confidentiality and integrity safeguards, 184–186, 186
 in mobile telemedicine, 178–180
 in tele-ultrasonography with OTELO, 395–396
 Gastroesophageal reflux disease (GERD), 181
 Gateway attacks, 80
 General Package Radio Service (GRPS), 341
 Gestational diabetes, 146
 Global positioning systems (GPS), 87, 107, 182
 use in location information, 13
 Global system for mobile communication (GSM), 186
 Glucose microsensors, 121
 commercial sensors, 123, 126
 dielectric spectroscopy, 122
 finger sticks, 121
 GlucoWatch, 122–123
 intravenous monitoring, 121
 subcutaneous, 122
 GlucoWatch, 122–123
 Glycated hemoglobin (HbA1c) test, 147
 automated reminders based on, 164
 in context of tight glycemic control, 162
 Glycemic control
 benefits of tight, 162
 improving for elderly diabetics, 161–163
 GMP Wireless Medicine Corp., 65
 GSM standard, 18, 21, 341

H

Handheld dietary device, 164–165
 Hardware upgrades, for mobile medical records system, 53
 Hash Function, 242
 HbA1c. *See* Glycated hemoglobin (HbA1c) test

- Head raising pillow. *See also* Orthopnea and CHF, 89
 - increased need for, 90
 - Header reduction, 332
 - Health education, via Web, 162
 - Health inequality, 350–351
 - Health Insurance Portability and Accountability Act (HIPAA) of 1996, 177, 189. *See also* HIPAA privacy rules; HIPAA security rules
 - Healthcare costs, for chronic illness, 104
 - Heart disease, ECG monitoring for, 9
 - Heart failure. *See also* Congestive heart failure (CHF)
 - stages of, 88–92
 - Heart rate stability, 70
 - Heart rhythm, 7, 69
 - Hierarchical security architecture, 225, 227–228
 - High blood pressure, 8
 - High care unit (HCU), 30, 31
 - High Speed Uplink Packet Access, 342, 345
 - HIPAA privacy rules, 190
 - HIPAA security rules, 189–190
 - HIPPOCRATE, 300
 - Holter monitor, 71
 - Home telemedicine unit (HTU), 168
 - Hop-to-hop relay, 72
 - vs.* cellular network technology, 69
 - Horizontal position determination, 16, 24
 - Housing requirements, for wearable devices, 10
 - Human error, 30, 31
 - dangers in partially closed loop control systems, 134
 - Humidity variation, monitoring in CHF, 98, 99
 - Hyperglycemia, 145
- I**
- IEEE 802.15 ultra-wideband technology, 15
 - IEEE 802.11a, 178
 - IEEE 802.15.4, 238–240
 - security in, 240–249
 - power management in, 249–250
 - IGMP, 327
 - Indoor location sensors, 13
 - microwave motion detection sensors, 14
 - passive infrared (PIR) motion detection sensors, 13–14
 - RF angle of arrival, 15
 - RF received signal strength indication (RSSI), 14
 - Infinite impulse response (IIR) filters, 134
 - Informatics for Diabetes Education and Telemedicine Project (IDEAT), 168
 - Information flow, in design of security policies, 224
 - Information theft, 186
 - Informed consent, 358–359
 - Infrastructure requirements
 - AMON approach, 22
 - EMERGE approach, 25
 - Tunstall approach, 20
 - Infusion pumps
 - recorded values, 41
 - serial communication specifications, 36
 - Initialization vectors, 77, 200, 231–232
 - in mobile telecardiology, 77
 - Input-output connections, for wearable devices, 10
 - Inspiratory tidal volume, 40
 - Installation and maintenance, mobile medical records system, 52
 - Institutional Review Board (IRB), 354–355
 - Institute of Medicine (IOM), 31
 - Insulet insulin pumps, 132
 - Insulin absorption rate, 136
 - Insulin concentration, relationship to blood glucose concentration, 135
 - Insulin injection methods
 - automated injection control, 131–138
 - for blood glucose management, 127
 - commercial pumps, 131, 132–133
 - comparisons, 127–128
 - insulin selection for, 130–131
 - micro-needle array, 128–129
 - safety issues, 131
 - subcutaneous devices, 128
 - Insulin pumps, 126–127, 148
 - reliability issues, 131
 - Insulin selection, matching to injection method, 130–131
 - Insulin therapy, 147
 - Integrated alarm monitoring systems, 29–32
 - advantages of, 32
 - alarm information indicators, 40–44, 43, 44
 - communication history indicators, 44, 45, 46

communication results with medical devices, 39

data collection interface (DCI) specification, 34–36

data format and bidirectional communication, 36–37

for medical devices, 39–46

system screen appearance, 39–40, 42

system specifications, 34–39

system structure for information on medical devices, 37–39

and transparent error reporting, 32–33

Intelligent cups, 24

Intelligent light switches, 25

Intelligent medicine bottles, 24

Intelligent software-defined radio, 282, 284

Intelligent therapy advising, in DIABTel system, 151

Intelligent walking sticks, 24

Intensive care unit (ICU)

- integrated alarm monitoring systems for, 29–32
- transparent error reporting in, 32–33

Interference, 15, 111

- reducing via frequency hopping, 108

International Diabetes Federation, 144

Internet Group Management Protocol (IGMP), *see* IGMP

Interoperability, security and, 190

Interrogators, 180

Intrahospital monitoring, 276, 279

Intravenous dialysis, 121

Intravenous monitoring, for blood glucose, 121

Investigation of perception and satisfaction, 359

Iontophoresis, 123

iPAQ hp2210 PDA, 153, 154

Ischemic heart disease, monitoring of, 101

J

Johns Hopkins University, 65

Joint source and channel coding paradigm, 299

K

Key distribution, 202–206

- biometrics method-based, 203–206
- pre-distribution, 202–203

Key exchange, 239

Key generation enforcing security policies, 226

Keying model, 241

Key update, 239

- frequency, 239

L

Laissez faire, 325

Latent errors, 30

Legal issues, 356

- confidentiality 356–357
- licensing, 357–358
- informed consent, 358–359

License-free radio frequencies, 86, 92

Licensing, 357

Life cycle, for wearable devices, 10

Life support, and integrated alarm monitoring systems, 32

Lifelines, 20

LifeSync wireless ECG system, 65

Likert scale, for mobile medical records system, 56

LINCOS project, 212, 386–387

Link quality, difficulties in estimating, 110

Lispro fast-acting insulin, 130, 131, 135

Load balancing, 273–274, 284

Local access and transport areas (LATAs), 191

Location-based tracking, 392

Location sensors, 24, 25, 74, 87, 182

- additional sensors for, 16–18
- camera-based, 15
- for elderly diabetics, 166
- horizontal position determination, 16
- indoor location, 13–15
- microwave motion detection sensors, 14
- passive infrared (PIR) motion detection sensors, 13–14
- RF angle of arrival, 15
- RF received signal strength indication (RSSI), 14
- RF ultra-wide band sensors, 15
- ultrasonic, 15
- for wrist wearable devices, 13

LOGINAT project, the, 300

Longevity, *vs.* quality of life, 104

Low blood pressure, 8

M

- m-health, 298
- m-JSCC/D, 299
- Machine learning techniques, 86
- Man-in-the-middle attacks, 80
- Mandatory reporting, 31
- MANET security, 76, 77, 78, 178–180
 - confidentiality and integrity safeguards, 185
- Medical alerts, in cardiac monitoring, 69, 70
- Medical devices
 - classification of errors, 31
 - communication results, 39
 - data collection from, 39
 - display of device abnormalities, 43
 - human error relevant to, 30
 - information indicators, 37–39
 - integrated alarm monitoring systems for, 39
 - serial communication specifications, 36
- Medical error, 30
- Medical imaging, 299
- Medical information management service, 277
- Medical information server, 51
- Medical JSCC/D, *see* m-JSCC/D
- Medical sensor networks (MSNs), 70. *See also*
 - Sensor networks
 - hardware components, 70
 - low-power, small-size ECG microsensors, 73–75
 - for nursing homes, 72
 - telecardiology design based on, 73
- Medical super-sensor (MSS), 72
- Medical standards, 392
- Medical telemetry systems, 37, 104
- Medical video streaming, 299
 - Robust multi-layer controller structure for, 304
- Medication dispenser, 19
 - technological compliance aids, 163
- Medication Event Monitoring System caps (MEMs), 163
- Mesh networking, 274
- Message authentication code, 201
- Memory aids, 163. *See also* Technological reminders
- Memory Mirror, 182
- Metabolic equivalents, daily expenditure of, 89
- Mica2Dot motes, 110, 111
- Micro-needle arrays, 126
 - honeycomb formation, 129
 - for insulin injection, 128–129
- Microsensors, 181
 - automated blood glucose management with, 119–121
 - integration with RFID technologies, 182
 - in mobile telemedicine, 182
- Microwave motion detection sensors, 14
- Middleware for Mobile Systems, 322
- MIDSTEP European project, 300
- MITHri I, 397
- Miniaturization, 21
 - in telecardiology design, 73–75
 - in wrist devices, 20
- MiniMed Guardian Real-Time Systems, 124–125
- MiniMed insulin pumps, 132
- MiniMed Paradigm Real-Time Systems, 124–125
- Minimum necessary use, 190
- Mobile ad hoc networks (MANETs), 177. *See also* MANET security
 - confidentiality and integrity safeguards, 184
- Mobile cardiac outpatient telemetry (MCOT), 65
- Mobile computing
 - adaptability of, 323
 - application-aware adaptation, 325
 - constraints, 324
 - environments (MCE), 324
 - limitation, 322
 - transparency, 324
 - vision, 323
 - what is, 320
- Mobile health care networks, 342
 - security and authentication, 342
- Mobile health care service, 276
- Mobile medical data access system, 49. *See also* Remote wireless patients' data access system
- Mobile robotic tele-echography system, 298
- Mobile telecardiology
 - based on wireless and sensor networks, 63–65
 - cardiac monitoring software, 74
 - cardiac monitoring with wireless sensor networks, 69–72
 - ECG throughput *vs.* patient density, 69
 - ECG transmission delay *vs.* patient density, 69
 - MANET security, 76
 - microcontroller unit for, 73

- MSN-based design, 73–80
 - multipatient secure ECG transmission, 78–79
 - multiple sensors in, 71
 - performance analysis, 68–69
 - results based on integrated wireless networks, 67–69
 - routing in simplified heterogeneous wireless networks, 67–68
 - secure ECG transmission, 75–80
 - security analysis, 80
 - significance of next-generation wireless networks for, 65–67
 - single-patient secure ECG transmission, 75–78
 - three-lead ECG sensor for, 73
 - wireless network features for, 67
 - Mobile telemedicine, ix
 - applications, 178
 - availability issues, 187–188
 - confidentiality and integrity in, 183–187
 - for diabetes care, 143–145
 - disadvantages, 188
 - 3G cellular technologies in, 178–180
 - MANET in, 178–180
 - microsensors in, 182
 - RFID tags in, 180–181, 182
 - security and privacy in, 175–177
 - security and privacy regulations in, 189–190
 - systems, 145, 149–150, 196, 326, 384–385
 - technologies in, 178–182, 179
 - technology security implications, 183–188
 - wireless sensor networks in, 81–182
 - WLANs in, 178–180
 - WPANs in, 178–180
 - Mobile telephones, use in diabetes care, 150
 - Mobility behaviors, in nursing homes, 72
 - Mobility management, 274, 283
 - Mobility transparency, 324
 - Model predictive control, 120, 121, 137
 - MoreTrack algorithm, 74
 - Moteiv's Tmote Sky sensors, 94
 - Motes, 86, 113, 213–214
 - bandwidth issues, 111
 - issues in cardiac monitoring, 112–113
 - limitations with fine-grained medical data, 104
 - link quality issues, 110
 - packet size limits, 113
 - power resource constraints, 111
 - MoteTrack, 394
 - Motion detection sensors
 - for elderly diabetics, 166
 - microwave-based, 14
 - PIR, 13–14
 - in telecardiology, 71
 - Motion sensors, 215–216
 - Movement detection, 12, 19, 23
 - for elderly diabetics, 166
 - Movement supervision, 7
 - MPEG-4 video, 308
 - Multi-hop communications, 274, 284
 - Multi-hop security, 76
 - Multipath reflection, 14
 - Multiplatform portability, with DIABTel system, 154
 - Myocardial infarctions, 105
- ## N
- N-person cooperative game, 285
 - Core value, 289–290
 - Shapley value, 289–290
 - Near misses, 30
 - Network-on-Chip Protocol
 - Next-Generation, 398–399
 - Network selection, 273, 282
 - Network switching, 64
 - Networking technologies
 - for patient monitoring, 18
 - role in mobile telemedicine, ix
 - Neural network control, 120, 121, 138
 - Neural network nonlinear predictive controller, 138
 - Noise, in location sensors, 14
 - Noncontact electrodes, 107, 109
 - Nonlinear predictive control (NLPC), 137
 - NPH slow-acting insulin, 130, 131
 - Nursing homes
 - cardiac monitoring software for, 74
 - MSNs for, 72
 - smart, 86
- ## O
- Odyssey, 329, 334–336
 - application adaption model, 334
 - application interaction with, 334–335
 - asymmetric links, 336

- expressive links, 336
- rover, 335–336
- viceroy, 335
- wardens, 335
- OMRON medical home-use devices, 18
- One-hop security mechanism, 77
- Open-flow microperfusion, 122
- OPNET, 68
- Orthopnea, 93, 101
- Oscillometric blood pressure measurement, 11–12
- OTELO, 301–304, 394–396, 403
- Output power levels, *vs.* error rates, 112
- Over-the-counter (OTC) devices, 105

P

- Patient care records (PCR), 210, 213, 218, 221, 226
- Patient monitoring
 - wireless LANs, 399–401
 - bluetooth, 400
 - ad hoc networks, 401–403
- PANs, 177
- Paper records, disadvantages of, 50–51
- Partially closed loop control, 120, 121
 - insulin injection by, 131
 - physician-prescribed regimens, 131, 134
- Passive infrared (PIR) motion detection sensors, 13–14
- Patient area networks (PANs), 179. *See also* PANs
- Patient care and monitoring, ix
 - AMON approach, 20–22, 26
 - blood oxygen saturation (SpO₂), 9, 16
 - blood pressure, 8, 16
 - blood values, 9
 - BodyMedia lifestyle monitoring, 19
 - capacitive sensors, 12
 - commercially available devices, 18–20
 - device design guidelines, 9–10
 - ECG-type techniques, 11, 17
 - elderly person surveillance, 6
 - electrocardiogram (ECG), 9
 - EMERGE approach, 22–25
 - fusion approach, 26
 - integrated alarm monitoring systems, 29–30
 - location sensors, 13–16
 - movement and fall detection, 7, 12
 - networking and communication
 - technologies for, 18
 - OMRON medical home-use devices, 18
 - oscillometric blood pressure measurement, 11–12
 - personal health devices, 7
 - personal supervision and alarming systems, 3–5
 - photoplethysmography (PPG), 11
 - post-trauma care, 6
 - pressure sensors, 12
 - pulse and heart rhythm, 7
 - pulse sensors, 10–11
 - pulse wave velocity (PWV), 8–9, 17–18
 - remote wireless patient data access system, 49–50
 - sensors and signs, 19–20
 - sensors for wrist wearable devices, 10–13
 - skin humidity, 8, 13
 - skin temperature, 12
 - system examples, 18–25
 - temperature, 7–8
 - Tunstall supervision approach, 19, 25–26
 - wearable devices, 20
 - wrist wearable device technologies, 10–18
- Patient empowerment, increasing through
 - DIABTel system, 153
- PEARL system, for elderly diabetics, 166–167
- Performance analysis, mobile telecardiology, 68–69
- Persistence, in design of security policies, 224
- Personal digital assistants (PDAs), 177
 - in cardiac monitoring, 86, 93
 - for elderly diabetics, 165
 - use in DIABTel system, 152
- Personal health devices, 7
- Personal supervision systems, 3–5. *See also* Alarming systems
 - target groups, 5–7
 - vital parameters, 7–9
- Personalized healthcare, 104
- Pervasive healthcare, 389, 390
- pH monitoring, with WSNs, 181–182
- Photoplethysmography (PPG), in wrist wearable devices, 11
- PHP, 53
- Physical activity duration, 19
- Piezoelectric insulin pumps, 126
- Piezoelectric pressure sensors, 12
- Polar heart rate monitors, 11
- Pole-assignment control, 120, 121, 134–136

Political issues, in telemedicine, ix
 Portability issues, mobile medical records system, 53
 Post-trauma care, 6
 Power management, 249
 Power requirements
 constraints with mote devices, 111
 power consumption sources, sensor networks, 76
 and security, 78
 in telecardiology design, 73–75
 for wearable devices, 10
 Prehospital service, 276, 279
 Pressure mats, 19, 24
 Pressure sensors, 12
 Privacy, ix
 and availability issues, 187–188
 in diabetes care systems, 156
 in telemedicine, 173, 175–177, 188–189
 Problem solving, impairments in older adults, 164
 Processing speed, age-related changes in, 167–168
 Productivity gains, *vs.* security, 177
 Protocol reduction, 333
 Protocol specification consistency, 185
 Proxies, 331–333
 Pulse detection, 10, 23, 86
 Pulse oximeters, 105, 181, 214
 Pulse parameters, 7
 Pulse sensors, for wrist wearable devices, 10–11
 Pulse transit time (PTT), 8, 17
 Pulse wave velocity (PWV), 8–9
 sensors for, 17–18
 PZT film, insulin pumps with, 126

Q

QRS complex, 106
 detection of, 111
 Quality of data, *see* data fidelity
 Quality of life, 105
 vs. longevity, 104
 Quality of Service (QoS), 269, 285, 321, 341
 in mobile information access applications, 328
 management, 282
 provisioning, 274
 Queued remote procedure call (QRPC), 336

R

R-peak detection algorithm, 111
 Radio frequency identification (RFID).
 See also RFID sensors
 confidentiality and integrity safeguards, 184
 Random number generator, 201–202
 Re-usable identification module (RUIM), 186
 Reading comprehension, age-related changes in, 168
 Received signal strength indication (RSSI), 14, 24
 sensors for CHF monitoring, 95–96, 97–98
 vs. distance, 14
 Record opening, in design of security policies, 223
 Redundancies, protecting against hackers via, 187
 Reflective PPG, 11
 Relational database, in mobile medical records system, 54
 Relocatable dynamic object (RDO), 335
 Remembrance agents, 165
 Remote authentication dial-in user service (RADIUS), 183
 Remote monitoring, 105
 of CHF, 86–87
 in elderly diabetics, 162
 service requirements for artificial pancreas devices, 150
 Remote wireless patients' data access system, 49–50
 authentication page, 55
 database server, 52–53
 doctors' satisfaction ratings, 58
 evaluation questionnaire, 58
 general workflow structure, 51
 handheld device and interface, 53–56
 nurses' satisfaction ratings, 59
 patient selection and view, 56
 relational database structure, 54
 sample interface pages, 57
 system architecture, 51–56
 system evaluation, 56–59
 vital signs monitoring via, 50–51
 Replay attacks, 187
 Resource allocation, 274
 Resource reservation, 284
 Resource Reservation Protocol (RSVP),
 see RSVP

Reverse Iontophoresis, 123
 RF angle of arrival sensors, 15
 RF ultra-wide band sensors, 15
 RFID sensors, 177

- confidentiality and integrity safeguards, 184, 186–187
- for elderly diabetics, 165
- integration with microsensors, 182
- in mobile medical records system, 55
- in mobile telemedicine, 180–181

 RFID-based wireless telemedicine systems, 216–217
 Risk Management, 355
 Robotistics, 395

- da Vinci™ robotic system, 355
- HIPPOCRATE, 300
- OTELO, 395
- tele-operated robot, 300
- ultrasound telemedical robotic system, 299
- Vilchis, 300

 Routing

- in simplified heterogeneous wireless telecardiology networks, 67–69
- via controlled flooding approach, 68

 RS-232C, 34, 35
 RSVP, 327

S

Salt intake, monitoring of, 89
 SARS outbreaks, tracking with RFID systems, 180
 Satellite technology, features for cardiac monitoring, 67
 Seamless handoff, 269
 Secure packet forwarding, 185
 Security, ix

- algorithms, 225–230
- of body sensor networks, 195–198
- cluster-based, 78
- in diabetes telemedicine, 150
- in IEE 802.15.4 clusters, 237–239
- and interoperability, 190
- low-energy requirements, 76
- in mobile medical records system, 55
- multi-hop *vs.* single-hop, 76
- outstanding issues, 190–191
- with paper *vs.* digital patient records, 51, 53
- and power management, 237–239

- protocol for wireless medical networks with multi-hopping, 230
- in telemedicine, 173, 175–177
- with wireless sensor networks, 209–213

 Security energy consumption, 78
 Security policies, 221–225

- design, 222–225

 Security regulations, impact on mobile telemedicine, 189–190
 Self-tuning adaptive control, 120, 121, 136–137
 SenseWear armband, 19
 Sensors

- motion, 215–216

 Sensor networks, 64

- anatomy of, 381–382
- applications, 380
- cardiac monitoring based on, 69–72
- case studies, 384–387
- for CHF remote monitoring, 86–87
- for elderly diabetics, 163–164
- issues in personal cardiac monitoring with, 103–105
- security problems, 383–384
- telecardiology based on, 63–65
- what is, 381
- Sensor node, 393

 Sensorcon Inc., 73
 Sensors

- AMON approach, 21–22
- in commercially available devices, 19–20
- EMERGE approach, 23–24
- for wearable devices, 10
- for telemedicine, 377

 Separation theorem, 299
 Session key generation, 230–231

- Inter-cluster, 230–231

 Serial interface, for DCI, 35
 Server-side intercept (SSI), *see* SSI
 Server-side proxy, *see* SSI
 Service-level security, 185
 Shannon's theory, 299
 Shared care services, 145, 148

- in diabetes management, 147–148

 Shortness of breath, 88, 89
 Signal latency, 109
 Signal loss, 15, 111

- reducing with frequency hopping, 108

 Simple Object Access Protocol (SOAP), 94
 Simulation model, 250
 Single-hop security, 76
 Sinus arrest, 111

Size requirements, for wearable devices, 10
 SK attacks, among CHs, 80
 Skin humidity, 8
 Skin impedance, 24
 Skin temperature, 7, 24
 sensors for, 12
 Skipjack cryptography, 77, 79, 231–232
 in mobile telecardiology, 64
 Sleep behavior, 97
 in CHF monitoring, 96
 Sleep duration, 19
 SMART cardiac monitoring system, 65
 Smart houses, for elderly diabetics, 166
 Smart monitoring system
 bedroom behavior monitoring, 94–99
 for CHF, 92–94
 database component, 94
 knowledge component, 94
 monitoring example, 99–101
 Spoofing attacks, 187
 SSI, 332
 Staff lapse of memory, 32
 Store-and-forward, 393
 Subcutaneous glucose microsensors, 121, 122
 Subcutaneous injection devices, 127, 128
 Lispro with, 130, 135
 Subscriber identity module (SIM), 186
 Supervised autonomy, in diabetes care, 148
 Supraventricular tachycardia, with narrow
 QRS complex, 111
 Symmetric-key key establishment protocol
 (SKKE), 242–249
 Symptom diary, in CHF monitoring, 101
 SYRTECH system, 300
 Syringe pumps, 41
 Systolic heart failure, 87–88
 Systolic pressure, 92

T

Tachycardia, 70, 105, 1111
 Target groups
 elderly person surveillance, 6
 for patient care and monitoring devices, 6
 personal health devices, 7
 post-trauma care, 6
 TCP/IP, 34, 108
 Technological reminders
 for elderly diabetics, 163–164
 for memory impaired adults, 162
 Telcomed, 20
 Tele-cardiology. *See* Mobile telecardiology
 Tele-consultation, 299
 Tele-Ultrasonography, 394
 with OTELO, 301–304, 394
 Tele-ultrasound systems for remote diagnosis,
 300
 Telehealth, 350
 in-home applications, 356
 TeleInVivo European project, the, 300
 Telehomecare service, 276, 279
 Telemedicine, 103, 350, 390, 369, 391, 403
 applications in, 178
 current use of, 370
 current state of, 371
 conventional, 369–371
 for cardiac health, 105–107
 efficacy for elderly diabetics, 161–163
 evolution of, 64
 improving elderly glycemic control via,
 161–163
 increase in physicians' workload
 due to, 147
 history of, 369–370
 licensing, 357–358
 proactive *vs.* reactive models, 103
 research, 351, 360–363
 security and privacy in, 173, 175–177
 wireless, 372–379
 wireless IP for, 386
 future research, 387
 pervasive healthcare, 390–392
 medical standards and, 392–393
 popular forms of, 393
 TeleMedMail, 321
 Telemedicine program
 longevity, 361–362
 Teletrauma system, 180
 Temporal key integrity protocol (TKIP), 183
 Thoracic fluid content, 92
 Tight glycemic control, 162
 Time difference of arrival (TDOA), 15
 Time utilization issues, paper-based patient
 records, 50–51
 TinyOS, 111
 Total energy expenditure, 19
 Traditional wireless health care systems, 277
 limitations, 277–278
 Transchamber flow gradient, 92
 Transfusion pumps, 41
 Transmissive PPG, 10

Transparent error reporting, 32–33, 33
 Transponders, 180
 Treadmill test, 88
 Trusted computing base, in design of security policies, 224
 TSM500 PDA, 154
 Tunstall supervision approach, 5, 19, 25–26
 Two-way interactive, 393
 TinyOS, 397
 Type 1 diabetes, 120, 146
 Type 2 diabetes, 120, 146

U

Ultra-wide band sensors, 15, 18
 Ultrasonography, 300, 395
 Ultrasonic location sensors, 15
 Ultrasound scan, 395
 ultrasound telemedical robotic system, 299
 Universal Mobile Telecommunication Systems (UMTS), 18, 21, 341
 in mobile telecardiology, 65
 use in mobile medical records system, 52
 Un-cooperation of routers
 in wireless patient monitoring, 401–403
 Underestimation of reporting, 32
 Unequal error protection (UEP), 299
 United Kingdom NHS database server, 53
 Universal Mobile Telecommunication Systems (UMTS), 18, 52, 274, 341–343
 University of Alabama, xiii
 Use-relevant errors, 31
 User-friendliness, and error reporting, 33

V

Vascular pressures, 92
 Vascular resistance, 92
 Ventilator, 41
 output values, 39, 40
 recorded values, 41
 serial communication specifications, 36
 Ventricular tachycardia, with broad QRS complex, 111
 Vertical handoff, 274
 Vibration alarms, 20
 Videoconference, 358–359
 H.323 standard, 385
 synchronized, 162

 in tele-ultrasonography with OTELO, 395–396
 two-way, 369, 371
 Vilchis, 300
 Virginia State University, xiii
 Virtual Private Network (VPN)
 in mobile health care network, 340
 Vital monitors, 41, 44
 serial communication specifications, 36
 Vital parameters, 7, AMON approach
 measurement of, 26
 blood oxygen saturation (SpO₂), 9
 blood pressure, 8
 blood values, 9
 in cardiac monitoring, 69
 collecting in real time, 182
 electrocardiogram (ECG), 9
 movement and fall detection, 7
 pulse and heart rhythm, 7
 pulse wave velocity (PWV), 8–9
 skin humidity, 8
 temperature, 7–8
 Vital sign, 341, 403
 Voluntary error reporting, 31

W

Wearable devices
 AMON approach, 20–21
 cardiac monitoring shirt/vest, 106
 commercially available, 20
 for continuous blood glucose monitoring, 122
 Wearable medical devices (WMD), 9
 Wearable server system, 107, 108
 Web documents, in DCI, 37, 38
 WebExpress, 332
 caching, 332
 differencing, 332
 protocol reduction, 333
 header reduction, 333
 Web graphical interface, in DIABTel system, 152
 Weight gain, in CHF, 89
 Weight sensors, 25, 86, 89
 WiFi protected access (WPA), 183, 185
 WiFi standard, 107
 WiiSARD cardiac monitoring system, 65
 WiMAX, 67, 270, 342
 Wired equivalent privacy (WEP), 183

- Wireless alarm buttons, 20
- Wireless body area network (WRAN), 72
- Wireless communication board, 73
- Wireless garments, 166
- Wireless local area networks (WLANs), 67, 177, 271
 - confidentiality and data integrity issues, 183, 184, 185
 - limitations for telecardiology, 66
 - in mobile telemedicine, 178–180
 - patient monitoring, 399–491
- Wireless metropolitan area network (WMAN), 269, 270
 - IEEE 802.16/WiMax, 270
 - IEEE 802.20/MobileFi, 270
- Wireless networks
 - in medical telemetry, 104–105
 - potential features for cardiac monitoring, 67
 - results in telecardiology, 67–69
 - significance for telecardiology, 65–67
 - telecardiology based on, 63–65
- Wireless patient area networks (PANs), 177
 - un-cooperation of routers, 400–403
- Wireless personal area networks (WPAN), 178–179, 184, 196, 271
 - low rate (LR-WPAN), 239, *see* IEEE 802.15.4
- Wireless QoS support, lack of, 66
- Wireless sensor networks (WSNs), 86, 177. *See also* Sensor networks
 - confidentiality and integrity safeguards, 184, 187
 - echochardiograms via, 182
 - in mobile telemedicine, 181–182
 - pH monitoring with, 181–182
 - pulse oximetry with, 181
- Wireless telemedicine, vii, 210, 372–379
 - advantages of, 372
 - applications, 65
 - disadvantages of, 378–379
 - hardware and software, 374–378
 - kit for nursing home and retirement center, 385
 - models, 212–213
 - RFID-based systems, 216–217
- WMTS band, 86
- WPAN technology, 178
 - confidentiality and integrity safeguards, 184, 185–186
 - for first responder applications, 180
 - in mobile telemedicine, 178–180
- Wrist wearable devices
 - capacitive or pressure-based sensors in, 12
 - design guidelines and challenges, 9–10
 - ECG-type techniques for, 11
 - EMERGE approach, 23–24
 - integrated sensors for, 10–13
 - location sensors for, 13–16
 - movement and fall detection in, 12
 - oscillometric blood pressure measurement in, 11–12
 - photoplethysmography (PPG) in, 11
 - pulse sensors, 10–11
 - skin temperature sensors for, 12
 - technologies for, 10

X

Xiao, Yang, xiii

Z

ZigBee alliance, 218, 229, 241–242
 ZigBee security application programming interfaces (APIs), 229–230
 ZigBee system, 18, 24

Mobile Telemedicine

A Computing and Networking Perspective

Remote-Access Medical Care

The concept of medical treatment from a distance (in absentia care) is actually quite ancient, dating back to tribal days where smoke signals were used to warn of serious disease in a community. Nowadays, telemedicine is used to facilitate treatment in rural areas, where the nearest doctor is miles away, through various forms of information technology, including videoconferencing and digital imaging. It can also be used to conveniently monitor chronically ill patients through electronic devices so that they can enjoy a better quality of life.

In **Mobile Telemedicine: A Computing and Network Perspective**, noted computer scientists Yang Xiao and Hui Chen examine computing and networking dilemmas arising from wireless and mobile telemedicine. Comprised of the contributions of many prominent international researchers, the book discusses the relative merits and limitations of the existing technology and sheds light on future developments. It begins with a discussion of patient care and monitoring through items such as personal alarm systems. It then reviews the current methods available to monitor cardiac and diabetic patients, analyzes the security and privacy considerations that arise with respect to the transmission of sensitive information, and examines issues relating to networking support. Finally, it concludes with a section on the opportunities and challenges faced by those involved at this intersection of healthcare and communications. By bridging the fields of medicine and information technology, this volume serves as a useful springboard for those pioneering IT researchers looking for a comprehensive reference guide.