Wireless Networks and Mobile Communications Series

BROADBAND MOBILE MULTIMEDIA

Techniques and Applications

Edited by
Yan Zhang • Shiwen Mao • Laurence T. Yang • Thomas M. Chen

CRC Press
Taylor & Francis Group
AN AUERBACH BOOK

BROADBAND MOBILE MULTIMEDIA

WIRELESS NETWORKS AND MOBILE COMMUNICATIONS

Dr. Yan Zhang, Series Editor
Simula Research Laboratory, Norway
E-mail: yanzhang@ieee.org

Unlicensed Mobile Access Technology: Protocols, Architectures, Security, Standards and Applications
Yan Zhang, Laurence T. Yang and Jianhua Ma
ISBN: 1-4200-5537-2

Wireless Quality-of-Service: Techniques, Standards and Applications
Maode Ma, Mieso K. Denko and Yan Zhang
ISBN: 1-4200-5130-X

Broadband Mobile Multimedia: Techniques and Applications
Yan Zhang, Shiwen Mao, Laurence T. Yang and Thomas M Chen
ISBN: 1-4200-5184-9

The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems
Lu Yan, Yan Zhang, Laurence T. Yang and Huansheng Ning
ISBN: 1-4200-5281-0

Millimeter Wave Technology in Wireless PAN, LAN, and MAN
Shao-Qiu Xiao, Ming-Tuo Zhou and Yan Zhang
ISBN: 0-8493-8227-0

Security in Wireless Mesh Networks
Yan Zhang, Jun Zheng and Honglin Hu
ISBN: 0-8493-8250-5

Resource, Mobility and Security Management in Wireless Networks and Mobile Communications
Yan Zhang, Honglin Hu, and Masayuki Fujise
ISBN: 0-8493-8036-7

Wireless Mesh Networking: Architectures, Protocols and Standards
Yan Zhang, Jijun Luo and Honglin Hu
ISBN: 0-8493-7399-9

Mobile WIMAX: Toward Broadband Wireless Metropolitan Area Networks
Yan Zhang and Hsiao-Hwa Chen
ISBN: 0-8493-2624-9

Distributed Antenna Systems: Open Architecture for Future Wireless Communications
Honglin Hu, Yan Zhang and Jijun Luo
ISBN: 1-4200-4288-2

AUERBACH PUBLICATIONS
www.auerbach-publications.com
To Order Call: 1-800-272-7737 • Fax: 1-800-374-3401
E-mail: orders@crcpress.com

BROADBAND MOBILE MULTIMEDIA

Techniques and Applications

Edited by
Yan Zhang • Shiwen Mao
Laurence T. Yang • Thomas M. 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

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: 978-1-4200-5184-1 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Broadband mobile multimedia : techniques and applications / edited by Yan Zhang ... [et al.].
p. cm. -- (Wireless networks and mobile communications ; v, 9)
Includes bibliographical references and index.
ISBN-13: 978-1-4200-5184-1
ISBN-10: 1-4200-5184-9
1. Mobile computing. 2. Multimedia systems. 3. Broadband communication systems. I. Zhang, Yan-Qing.

QA76.59.B76 2008
004.6'6--dc22 2007050624

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

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

Table of Contents

Preface

"Follow the money" was W. Mark Felt's (a.k.a. *Deep Throat*) advice to Bob Woodward and Carl Bernstein for unraveling the Watergate scandal. After years of research and development, there have been significant advances in signal processing, networking and delivery technologies, network infrastructure and deployment, as well as successful business models. Multimedia is now ready to hit the market. Users are not satisfied with simple forms of communications anymore. A wide range of multimedia applications are emerging, such as Voice-over-IP, online chatting, video on demand, Internet Protocol Television (IPTV) or cellvision/mobile TV, and interactive gaming, among others. Service providers are making great efforts to move toward "triple play" or "quad play." This trend is highlighted by the recent multi-billion-dollar eBay/Skype deal and the Google/YouTube deal.

An equally important advice for engineers, researchers, and for all of us is to "find the bottleneck." Thus, where is the performance bottleneck in the big picture as multimedia is becoming widely available in the Internet? The answer, we believe, is wireless access networks, such as third-generation (3G) and beyond wireless networks, Wi-Fi, WiMAX, and Bluetooth wireless local area networks (WLAN), *ad hoc*/mesh networks, and wireless sensor networks. Despite considerable advances, current wireless networks are not able to offer comparable data rates as do their wired counterparts. Although it frees users from a socket and cable, mobility brings about a new level of challenge, including time-varying wireless channels and dynamic topology and connectivity. The situation is even more serious in the case of multihop wireless networks, where end-to-end throughput quickly decreases as hop count increases, largely due to carrier sensing and spatial reuse issues. As the increasing demand for multimedia communications continues to drive the expansion of consumer and enterprise markets as well as the evolution of wireless technologies, multimedia service provisioning is believed to be one of the prerequisites to guarantee the success of the next-generation wireless networks.

This book aims to meet this compelling need by providing a collection of the latest advances in this important problem area. Given the considerable research effort being made and the vast literature that exists, it is not possible to provide a complete coverage of all the related issues. However, we aim to provide a big picture of state-of-the-art research, a representative sampling of important research outcomes, and in-depth treatments of selected topics in the area of broadband wireless multimedia communications. Overall, this book is a useful technical guide covering introductory concepts, fundamental techniques, latest advances, and open issues in broadband wireless multimedia. A large number of illustrative figures, cross-references, as well as comprehensive references for readers interested in more details are provided.

This book consists of 15 chapters, which are organized into four parts as follows:

- Multimedia systems
- Multimedia over *ad hoc* and sensor networks
- Multimedia over wireless local area networks
- QoS and enabling technologies

Part I introduces various broadband wireless multimedia systems and surveys related work. Part II focuses on the routing and cross-layer design issue of multimedia communication over multihop wireless *ad hoc*/sensor networks, where video is used as a reference application. Part III explores various issues related to multimedia communications over WLANs, which constitute a dominant part of today's broadband wireless access networks. Part IV presents latest advances in QoS provisioning mechanisms and other enabling technologies, including end-to-end QoS provisioning, middleware, mobility management, scheduling, and power control.

The salient features of the book are as follows:

- Identifies the basic concepts, key technologies, and cutting-edge research outcomes, as well as open problems and future research directions in the important problem area of broadband mobile multimedia
- Provides comprehensive references on state-of-the-art technologies for broadband wireless multimedia
- Contains a sufficient number of illustrative figures for easy reading and understanding of the materials
- Allows complete cross-referencing through a broad coverage on layers of the protocol architecture
- In-depth treatment of selected problems/technologies for enabling wireless multimedia service

The book represents a useful reference for techniques and applications of broadband wireless multimedia. Target readers include students, educators, telecommunication service providers, research strategies, scientists, researchers, and engineers working in the areas of wireless communications, wireless networking, and multimedia communications. It can also be used as a textbook for an advanced selected topic course on broadband wireless multimedia for graduate students.

This book would not have been possible without the efforts and the time invested by all the contributors. They were extremely professional and cooperative, and did a great job in the production of this book. Our reviewers provided valuable comments/feedbacks, which, we believe, greatly helped improve the quality of this book. Special thanks go to Rich O'Hanley, Jessica Vakili, and Karen Schober of Taylor & Francis Group for their continuous support, patience, and professionalism from the beginning to the final stage. Last but not least, we thank our families and friends for their constant encouragement, patience, and understanding throughout this project, which was a pleasant and rewarding experience.

Yan Zhang
Shiwen Mao
Laurence T. Yang
and **Thomas M. Chen**

Editors

Yan Zhang holds a PhD from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Since August 2006, he has been working with the Simula Research Laboratory, Norway (http://www.simula.no/). He is an associate editor of *Security and Communication Networks* (Wiley) and is also on the editorial board of three international journals. He has served as coeditor for a number of special journal issues and a score of books. Dr. Zhang has been active as a chair, cochair, or organizer for many workshops since 2006. He received the Best Paper Award and Outstanding Service Award in the IEEE 21st International Conference on Advanced Information Networking and Applications (AINA-07). His research interests include resource, mobility, spectrum, energy, and security management in wireless networks and mobile computing. He is also a member of IEEE and IEEE ComSoc.

Shiwen Mao is an assistant professor in the Department of Electrical and Computer Engineering, Auburn University, Auburn, Alabama. He received his BS and MS in electrical engineering from Tsinghua University, Beijing, People's Republic of China in 1994 and 1997, respectively. He also received his BS in business management from Tsinghua University in 1994; an MS in system engineering and a PhD in electrical and computer engineering from Polytechnic University, Brooklyn, New York, in 2000 and 2004, respectively. He was a research member at IBM China Research Lab, Beijing from 1997 to 1998, and a research intern at Avaya Labs-Research,

Holmdel, New Jersey, in the summer of 2001. He was a research scientist in the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, Virginia, from December 2003 to April 2006. He is an associate editor of two respected journals, and he served as cochair of the Wireless Communications and Networking Symposium of CHINACOM 2008. Dr. Mao coauthored two textbooks: *TCP/IP Essentials: A Lab-Based Approach* (Cambridge University Press, 2004) and *Principles of Network Performance Analysis* (Cambridge University Press, 2008). He is a member of the IEEE and received a certificate of appreciation from the IEEE Computer Society in 2007. Dr. Mao's research interests include cross-layer design and optimization in multihop wireless networks, cognitive networks, and multimedia communications.

Laurence T. Yang is a professor at St Francis Xavier University, Antigonish, Nova Scotia, Canada. His research focuses on high-performance computing and networking, embedded systems, ubiquitous/pervasive computing, and intelligence. He has published approximately 280 papers in refereed journals, conference proceedings, and book chapters in these areas. He has been involved in more than 100 conferences and workshops as a program/general conference chair, and in more than 200 conference and workshops as a program committee member. He has served on numerous IEEE Technical Committees and on the IFIP Working Group 10.2 on Embedded Systems. In addition, he is the editor-in-chief of nine international journals and a few book series. He serves as an editor for international journals. He has contributed to, edited, or coedited 30 books. He has won numerous awards for his research, including several IEEE Best Paper Awards.

Thomas M. Chen is an associate professor in the Department of Electrical Engineering at Southern Methodist University (SMU). Before joining SMU, he was a senior member of the technical staff at GTE Laboratories (now Verizon) in Waltham, Massachusetts. He is currently the editor-in-chief of *IEEE Communications Magazine*. He also serves as a senior technical editor for *IEEE Network*, and was the founding editor of *IEEE Communications Surveys*. He serves as a member-at-large on the IEEE Communications Society Board of Governors. He is a coauthor of *ATM Switching Systems* (Artech House 1995). He received the IEEE Communications Society's Fred W. Ellersick best paper award in 1996.

List of Contributors

P. Bellavista
University of Bologna
Bologna, Italy

Frank Brouwer
Twente Institute for Wireless and Mobile Communications
Enschede, The Netherlands

Irene de Bruin
Twente Institute for Wireless and Mobile Communications
Enschede, The Netherlands

Min Chen
University of British Columbia
Vancouver, British Columbia, Canada

Gene Cheung
HP Research Laboratories
Tokyo, Japan

Kaushik R. Chowdhury
Georgia Institute of Technology
Atlanta, Georgia

Chen-Nee Chuah
University of California
Davis, California

A. Corradi
University of Bologna
Bologna, Italy

Nikki Cranley
Dublin Institute of Technology
Dublin, Ireland

Igor D.D. Curcio
Nokia Research Center
Tampere, Finland

L. Foschini
University of Bologna
Bologna, Italy

Yusun Fu
Huawei Technologies Co., Ltd
Shanghai, China

Jianhua He
Swansea University
Swansea, United Kingdom

Y. Thomas Hou
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Tao Jiang
Huazhong University of Science and Technology
Wuhan, China

Thanasis Korakis
Polytechnic University
Brooklyn, New York

Victor C.M. Leung
University of British Columbia
Vancouver, British Columbia, Canada

Danjue Li
Cisco System Ltd
San Jose, California

Shiwen Mao
Auburn University
Auburn, Alabama

Shunliang Mei
Tsinghua University
Beijing, China

Tommaso Melodia
State University of New York at Buffalo
Buffalo, New York

Scott F. Midkiff
Virginia Polytechnic Institute and State University
Blacksburg, Virginia
and
National Science Foundation
Arlington, Virginia

Gabriel-Miro Muntean
Dublin City University
Dublin, Ireland

R. Muralishankar
PES Institute of Technology
Bangalore, India

Shivendra Panwar
Polytechnic University
Brooklyn, New York

R. Venkatesha Prasad
Delft University of Technology
Delft, The Netherlands

Vijay S. Rao
Delft University of Technology
Delft, The Netherlands

H.N. Shankar
PES Institute of Technology
Bangalore, India

Hanif D. Sherali
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Zhi Sun
Georgia Institute of Technology
Atlanta, Georgia

Zuoyin Tang
Swansea University
Swansea, United Kingdom

Zhifeng Tao
Mitsubishi Electric Research Laboratories
Cambridge, Massachusetts

Leandros Tassiulas
University of Thessaly
Volos, Greece

Ramakrishna Vedantham
Nokia Research Center
Palo Alto, California

Wei Wei
University of California
Berkeley, California

Neill Whillans
Twente Institute for Wireless and Mobile Communications
Enschede, The Netherlands

Youqian Xiao
Huawei Technologies Co., Ltd
Shanghai, China

Zongkai Yang
Huazhong University of Science and Technology
Wuhan, China

Laurence T. Yang
St. Francis Xavier University
Antigonish, Nova Scotia, Canada

S.J. Ben Yoo
University of California
Davis, California

Rong Yu
South China University of Technology
Guangzhou, China

Yong Yuan
Huazhong University of Science and Technology
Wuhan, China

Avideh Zakhor
University of California
Berkeley, California

Qian Zhang
Hong Kong University of Science and Technology
Hong Kong, China

Yan Zhang
Simula Research Laboratory
Lysaker, Norway

Ya-Qin Zhang
Microsoft Corporation
Beijing, China

Wenwu Zhu
Microsoft Corporation
Beijing, China

MULTIMEDIA SYSTEMS

I

Chapter 1

Design Challenges for Wireless Multimedia Sensor Networks

Tommaso Melodia and Kaushik R. Chowdhury

Contents

Low-cost CMOS cameras, microphones, and sensors have recently become ubiquitously available. This has fostered the development of the so-called wireless multimedia sensor networks (WMSNs), that is, distributed systems composed of wirelessly interconnected devices that can ubiquitously retrieve multimedia content such as video and audio streams, still images, and scalar sensor data from the environment. This chapter discusses the state-of-the-art in algorithms, protocols, and hardware for WMSNs; open research issues are discussed in detail. Architectures for WMSNs are explored, along with their advantages and drawbacks. Existing solutions and open research issues at the application, transport, network, link, and physical layers of the communication stack are investigated, along with possible cross-layer synergies and optimizations.

1.1 Introduction

Wireless sensor networks (WSNs) [1] have significantly attracted the attention of the research community in the past few years. Significant results in this area have resulted in the increased interest in potential civil and military applications. In fact, sensor networks can consistently enhance our ability to gather physical information from the environment. However, as of today, most deployed WSNs measure scalar physical phenomena such as temperature, pressure, humidity, or location of objects. Sensor network communication protocols are usually designed for applications that have low bandwidth demands, and are delay-tolerant.

The recent availability of inexpensive hardware such as CMOS cameras and microphones, able to ubiquitously capture multimedia content from the environment, is paving the way for the development of the so-called wireless multimedia sensor networks (WMSN) [2,3], that is, networks of wirelessly interconnected devices that allow retrieving video and audio streams, still images, and scalar sensor data. With rapid improvements and miniaturization in hardware, a single device can be equipped with audio and visual information-collection modules. As an example, the Cyclops image capturing and inference module [4] is designed for extremely light-weight imaging and can be interfaced with a host mote such as Crossbow's MICA2 [5] or MICAz [6]. In addition, WMSNs will be able to store, process in real-time, correlate, and fuse multimedia data originating from heterogeneous sources.

The challenges associated with the delivery and in-network processing of multimedia content require the sensor network paradigm to be rethought to provide applications with a predefined level of quality of service (QoS). Because the goal of minimizing the energy consumption has driven most of the research in sensor networks so far, mechanisms to efficiently deliver application-level QoS, and to map these requirements to network-layer metrics such as latency and jitter, have not been primary concerns for sensor network researchers.

Conversely, algorithms, protocols, and techniques to deliver multimedia content over large-scale networks have been the focus of intensive research in the past 20 years, especially in Asynchronous Transfer Mode (ATM) wired and wireless networks. Later, many of the results derived for ATM networks have been readapted, and architectures such as Diffserv and Intserv for Internet QoS delivery have been developed. However, there are several peculiarities that make the QoS delivery of multimedia content in sensor networks an even more challenging, and largely unexplored, task.

This chapter provides a detailed description of the main research challenges in algorithms and protocols for the development of WMSNs. In particular, Section 1.2 describes the applications of WMSNs, whereas Section 1.3 describes the main characteristics of WMSNs, including the major factors influencing their design, reference architecture, and internal organization. Section 1.4 suggests possible architectures for WMSNs and describes their characterizing features. Section 1.5 discusses the functionalities handled at the application layer of a WMSN, including possible advantages and challenges of multimedia in-network processing. Section 1.6 discusses the existing solutions and open research issues at the transport, network, link, and physical layers of the communication stack, respectively. Finally, Section 1.8 concludes the chapter.

1.2 Applications of Wireless Multimedia Sensor Networks

WMSNs will enhance existing sensor network applications such as tracking, home automation, and environmental monitoring. Moreover, new applications will be enabled by the possibility of retrieving multimedia content such as the following:

- *Multimedia surveillance sensor networks*. Wireless video sensor networks will be composed of interconnected, battery-powered miniature video cameras, each packaged with a low-power wireless transceiver that is capable of processing, sending, and receiving data. Video and audio sensors will be used to enhance and complement existing surveillance systems against crime and terrorist attacks. Large-scale networks of video sensors can extend the ability of law enforcement agencies to monitor areas, public events, private properties, and borders.
- *Storage of potentially relevant activities*. Multimedia sensors could infer and record potentially relevant activities (thefts, car accidents, traffic violations), and make video/audio streams or reports available for future query.
- *Traffic avoidance, enforcement, and control systems*. It will be possible to monitor car traffic in big cities or highways and deploy services that offer traffic routing advice to avoid congestion. In addition, smart parking advice systems based on WMSNs [7] will allow monitoring available parking spaces and provide drivers with automated parking advice, thus improving mobility in urban areas. Moreover, multimedia sensors may monitor the flow of vehicular traffic on highways and retrieve aggregate information such as average speed and number of cars. Sensors could also detect violations and transmit video streams to law enforcement agencies to identify the violator, or buffer images and streams in case of accidents for subsequent accident scene analysis.
- *Advanced healthcare delivery*. Telemedicine sensor networks [8] can be integrated with third-generation (3G) multimedia networks to provide ubiquitous healthcare services. Patients will carry medical sensors to monitor parameters such as body temperature, blood pressure, pulse oximetry, electrocardiogram (ECG), and breathing activity. Furthermore, remote medical centers will perform advanced remote

monitoring of their patients via video and audio sensors, location sensors, and motion or activity sensors, which can also be embedded in wrist devices [8].

- *Automated assistance for the elderly and family monitors*. Multimedia sensor networks can be used to monitor and study the behavior of elderly people as a means to identify the causes of illnesses that affect them such as dementia [9]. Networks of wearable or video and audio sensors can infer emergency situations and immediately connect elderly patients with remote assistance services or with relatives.
- *Environmental monitoring*. Several projects on habitat monitoring that use acoustic and video feeds are being envisaged, in which information has to be conveyed in a time-critical fashion. For example, arrays of video sensors are already being used by oceanographers to determine the evolution of sandbars via image-processing techniques [10].
- *Structural health monitoring*. Multimedia streams can be used to monitor the structural health of bridges [11] or other structures.
- *Person locator services*. Multimedia content such as video streams and still images, along with advanced signal-processing techniques, can be used to locate missing persons, or identify criminals or terrorists.
- *Industrial process control*. Multimedia content such as imaging, temperature, or pressure, among others, may be used for time-critical industrial process control. "Machine vision" is the application of computer vision techniques to industry and manufacturing, where information can be extracted and analyzed by WMSNs to support a manufacturing process such as those used in semiconductor chips, automobiles, food, or pharmaceutical products. For example, in the quality control of manufacturing processes, details or final products are automatically inspected to find defects. In addition, machine vision systems can detect the position and orientation of parts of the product to be picked up by a robotic arm. The integration of machine vision systems with WMSNs can simplify and add flexibility to systems for visual inspections and automated actions that require high-speed, high-magnification, and continuous operation.

1.3 Design and Characteristics of Wireless Multimedia Sensor Networks

There are several factors that mainly influence the design of a WMSN, which are outlined in this section.

- *Application-specific QoS requirements*. The wide variety of applications envisaged on WMSNs will have different requirements. In addition to data delivery modes typical of scalar sensor networks, multimedia data includes "snapshot" and "streaming multimedia" content. Snapshot-type multimedia data contains event-triggered observations obtained in a short time period. Streaming multimedia content is generated over longer time periods and requires sustained information delivery. Hence, a strong foundation is needed in terms of hardware and supporting high-level algorithms to deliver QoS and consider application-specific requirements. These requirements may pertain to multiple domains and can be expressed, among others, in terms of a combination of bounds on energy consumption, delay, reliability, distortion, or network lifetime.

- *High bandwidth demand*. Multimedia content, especially video streams, require transmission bandwidth that is orders of magnitude higher than that supported by currently available sensors. For example, the nominal transmission rate of state-of-the-art IEEE 802.15.4 compliant components such as Crossbow's [6] MICAz or TelosB [12] motes is 250 kbit/s. Data rates at least one order of magnitude higher may be required for high-end multimedia sensors, with comparable power consumption. Hence, high data rate and low power consumption transmission techniques need to be leveraged. In this respect, the ultra wideband (UWB) transmission technique seems particularly promising for WMSNs, and its applicability is discussed in Section 1.7.
- *Cross-layer coupling of functionalities*. In multihop wireless networks, there is a strict interdependence among functions handled at all layers of the communication stack. Functionalities handled at different layers are inherently and strictly coupled due to the shared nature of the wireless communication channel. Hence, the various functionalities aimed at QoS provisioning should not be treated separately when efficient solutions are sought.
- *Multimedia source coding techniques*. Uncompressed raw video streams require excessive bandwidth for a multihop wireless environment. For example, a single monochrome frame in the NTSC-based Quarter Common Intermediate Format (QCIF; 176×120) requires approximately 21 kB, and at 30 frames per second (fps), a video stream requires over 5 Mbit/s. Hence, it is apparent that efficient processing techniques for lossy compression are necessary for multimedia sensor networks. Traditional video coding techniques used for wireline and wireless communications are based on the idea of reducing the bit rate generated by the source encoder by exploiting source statistics. To this aim, encoders rely on "intraframe" compression techniques to reduce redundancy within one frame, whereas they leverage "interframe" compression (also known as "predictive encoding" or "motion estimation") to exploit redundancy among

subsequent frames to reduce the amount of data to be transmitted and stored, thus achieving good rate-distortion performance. Because predictive encoding requires complex encoders, powerful processing algorithms, and entails high energy consumption, it may not be suited for low-cost multimedia sensors. However, it has recently been shown [13] that the traditional balance of complex encoder and simple decoder can be reversed within the framework of the so-called distributed source coding, which exploits the source statistics at the decoder, and by shifting the complexity at this end allows the use of simple encoders. Clearly, such algorithms are very promising for WMSNs and especially for networks of video sensors, where it may not be feasible to use existing video encoders at the source node due to processing and energy constraints.

- *Multimedia in-network processing*. Processing of multimedia content has mostly been approached as a problem isolated from the network-design problem, with a few exceptions such as joint source-channel coding [14] and channel-adaptive streaming [15]. Hence, research that addressed the content delivery aspects has typically not considered the characteristics of the source content and has primarily studied cross-layer interactions among lower layers of the protocol stack. However, the processing and delivery of multimedia content are not independent and their interaction has a major impact on the levels of QoS that can be delivered. WMSNs will allow performing multimedia in-network processing algorithms on the raw data. Hence, the QoS required at the application level will be delivered by means of a combination of both cross-layer optimization of the communication process and in-network processing of raw data streams that describe the phenomenon of interest from multiple views, with different media, and on multiple resolutions. Hence, it is necessary to develop application-independent and self-organizing architectures to flexibly perform in-network processing of multimedia contents.
- *Power consumption*. This is a fundamental concern in WMSNs, even more than in traditional WSNs. In fact, sensors are battery-constrained devices, whereas multimedia applications produce high volumes of data, which require high transmission rates and extensive processing. The energy consumption of traditional sensor nodes is known to be dominated by the communication functionalities, whereas this may not necessarily be true in WMSNs. Therefore, protocols, algorithms, and architectures to maximize the network lifetime while providing the QoS required by the application are a critical issue.
- *Flexible architecture to support heterogeneous applications*. WMSN architectures will support several heterogeneous and independent applications with different requirements. It is necessary to develop

flexible, hierarchical architectures that can accommodate the requirements of all these applications in the same infrastructure.

- *Multimedia coverage.* Some multimedia sensors, in particular video sensors, have larger sensing radii and are sensitive to direction of acquisition (directivity). Furthermore, video sensors can capture images only when there is unobstructed line of sight between the event and the sensor. Hence, coverage models developed for traditional WSNs are not sufficient for predeployment planning of a multimedia sensor network.
- *Integration with Internet Protocol (IP) architecture.* It is of fundamental importance for the commercial development of sensor networks to provide services that allow querying the network to retrieve useful information from anywhere and at any time. For this reason, future WMSNs will be remotely accessible from the Internet, and will therefore need to be integrated with the IP architecture. The characteristics of WSNs rule out the possibility of all IP sensor networks and recommend the use of application-level gateways or overlay IP networks as the best approach for integration between WSNs and the Internet [16].
- *Integration with other wireless technologies.* Large-scale sensor networks may be created by interconnecting local "islands" of sensors through other wireless technologies. This needs to be achieved without sacrificing on the efficiency of the operation within each individual technology.

1.4 Network Architecture

The problem of designing a "scalable network architecture" is of primary importance. Most proposals for WSNs are based on a flat, homogeneous architecture in which every sensor has the same physical capabilities and can only interact with neighboring sensors. Traditionally, the research on algorithms and protocols for sensor networks has focused on "scalability," that is, how to design solutions whose applicability would not be limited by the growing size of the network. Flat topologies may not always be suited to handle the amount of traffic generated by multimedia applications including audio and video. Similarly, the processing power required for data processing and communications, and the power required to operate it, may not be available on each node.

1.4.1 Reference Architecture

In Figure 1.1, we introduce a reference architecture for WMSNs, where three sensor networks with different characteristics are shown, possibly deployed

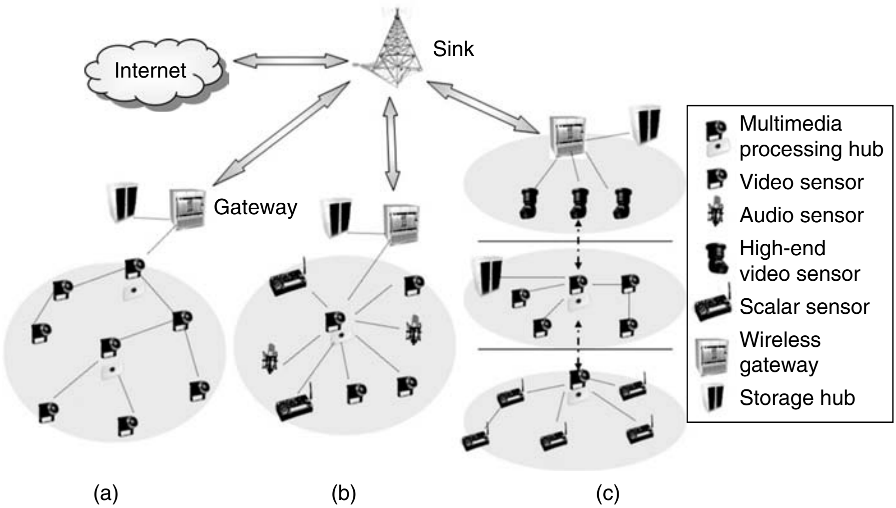

Figure 1.1 Reference architecture of a WMSN. (a) Single-tier flat, homogeneous sensors, distributed processing, centralized storage; (b) single-tier clustered, heterogeneous sensors, centralized processing, centralized storage; (c) multitier, heterogeneous sensors, distributed processing, distributed storage.

in different physical locations. The first cloud on the left shows a single-tier network of homogeneous video sensors. A subset of the deployed sensors have higher processing capabilities, and are thus referred to as "processing hubs." The union of the processing hubs constitutes a distributed processing architecture. The multimedia content gathered is relayed to a "wireless gateway" through a multihop path. The gateway is interconnected to a "storage hub," which is in charge of storing multimedia content locally for subsequent retrieval. Clearly, more complex architectures for distributed storage can be implemented when allowed by the environment and the application needs, which may result in energy savings because by storing it locally the multimedia content does not need to be wirelessly relayed to remote locations. The wireless gateway is also connected to a central "sink," which implements the software front end for network querying and tasking. The second cloud represents a single-tiered clustered architecture of heterogeneous sensors (only one cluster is depicted). Video, audio, and scalar sensors relay data to a central clusterhead (CH), which is also in charge of performing intensive multimedia processing on the data (processing hub). The CH relays the gathered content to the wireless gateway and storage hub. The last cloud on the right represents a multitiered network, with heterogeneous sensors. Each tier is in charge of a subset of the functionalities. Resource-constrained, low-power scalar sensors are in charge of performing simpler tasks, such as detecting scalar physical measurements, whereas resource-rich, high-power devices are responsible for more complex tasks. Data processing and storage can be performed in a distributed fashion at each different tier.

1.4.2 Single- versus Multitier Sensor Deployment

One possible approach for designing a multimedia sensor application is to deploy homogeneous sensors, and program each sensor to perform all possible application tasks. Such an approach yields a flat, single-tier network of homogeneous sensor nodes. An alternative, multitier approach is to use heterogeneous elements (Figure 1.2) [17]. In this approach, resource-constrained, low-power elements are in charge of performing simpler tasks, such as detecting scalar physical measurements, whereas resource-rich, high-power devices take on more complex tasks. For instance, a surveillance application can rely on low-fidelity cameras or scalar acoustic sensors to perform motion or intrusion detection, whereas high-fidelity cameras can be woken up on demand for object recognition and tracking. In Ref. 18, a multitier architecture is advocated for video sensor networks for surveillance applications. The architecture is based on multiple tiers of cameras with different functionalities, with the lower tier composed of

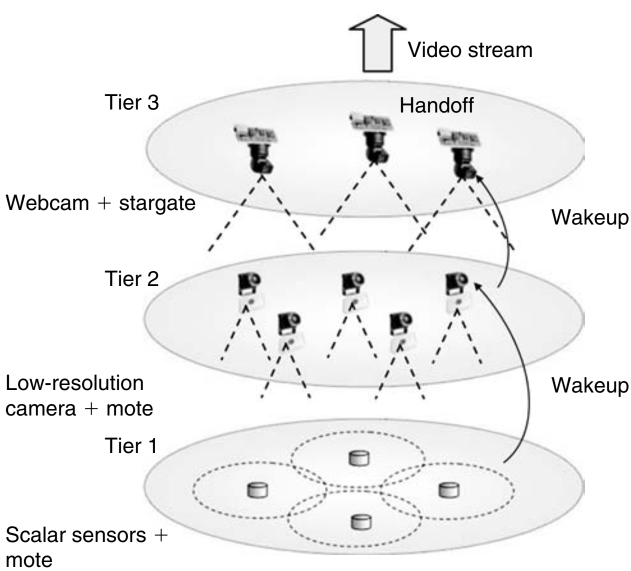

Figure 1.2 The multitier architecture of SensEye. (From Kulkarni, P., Ganesan, D., Shenoy, P., and Lu, Q. SensEye: A Multi-Tier Camera Sensor Network. *Proc. of ACM Multimedia*. Singapore, November 2005.)

low-resolution imaging sensors, and the higher tier composed of high-end pan-tilt-zoom cameras. It is argued, and shown by means of experiments, that such an architecture offers considerable advantages with respect to a single-tier architecture in terms of scalability, lower cost, better coverage, higher functionality, and better reliability.

1.4.3 Coverage

In traditional WSNs, sensor nodes collect information from the environment within a predefined "sensing range," that is, a roughly circular area defined by the type of sensor being used.

Multimedia sensors generally have larger sensing radii and are also sensitive to the direction of data acquisition. In particular, cameras can capture images of objects or parts of regions that are not necessarily close to the camera itself. However, the image can obviously be captured only when there is an unobstructed line of sight between the event and the sensor. Furthermore, each multimedia sensor/camera perceives the environment or the observed object from a different and unique viewpoint, given the different orientations and positions of the cameras relative to the observed event or region. In Ref. 19, a preliminary investigation of the coverage problem for video sensor networks is conducted. The concept of sensing range is replaced with the camera's "field of view," that is, the maximum volume visible from the camera. It is also shown how an algorithm designed for traditional sensor networks does not perform well with video sensors in terms of coverage preservation of the monitored area.

1.4.4 Internal Organization of a Multimedia Sensor

A sensor node is composed of several basic components, as shown in Figure 1.3: a sensing unit, a processing unit (central processing unit [CPU]), a communication subsystem, a coordination subsystem, a storage unit (memory), and an optional mobility/actuation unit. Sensing units are usually composed of two subunits: sensors (cameras, audio, or scalar sensors) and analog to digital converters (ADCs). The analog signals produced by the sensors based on the observed phenomenon are converted to digital signals by the ADC, and then fed into the processing unit. The processing unit, which is generally associated with a storage unit (memory), manages the procedures that make the sensor node collaborate with the other nodes to carry out the assigned sensing tasks. A communication subsystem connects the node to the network, and is composed of a transceiver unit and of communication software. The latter includes a communication protocol stack and system software such as middleware, operating systems,

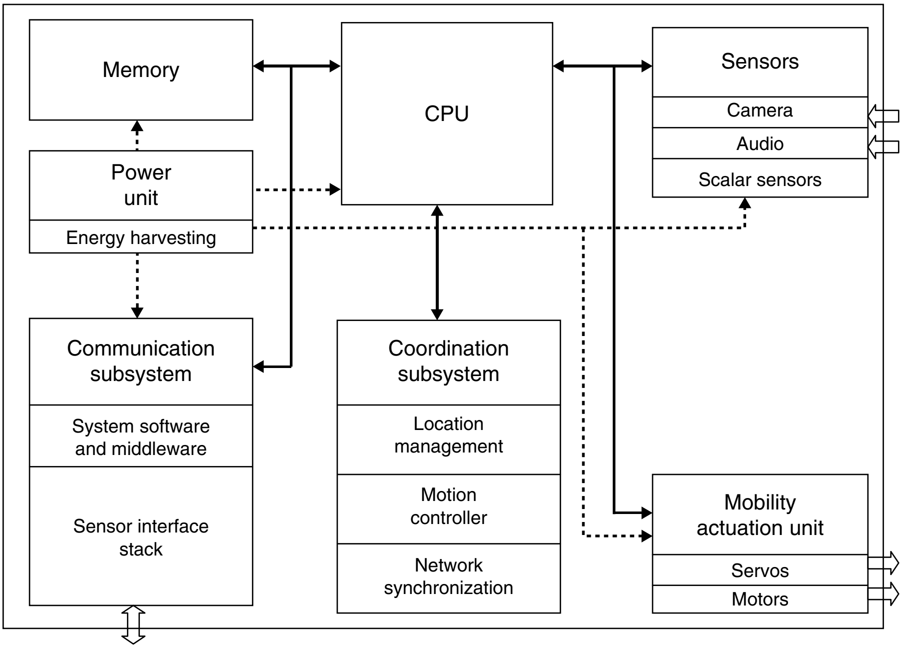

Figure 1.3 Internal organization of a multimedia sensor.

and virtual machines. A coordination subsystem is in charge of coordinating the operation of different network nodes by performing operations such as network synchronization and location management. An optional mobility/actuation unit can also be controlled by the coordination subsystem. Finally, the whole system is powered by a power unit that may be supported by an energy scavenging unit such as solar cells.

1.5 Application Layer

The functionalities handled at the application layer of a WMSN are characterized by high heterogeneity, and encompass traditional communication problems as well as more general system challenges. The services offered by the application layer include (i) providing traffic management and admission control functionalities, that is, prevent applications from establishing data flows when the network resources needed are not available; (ii) performing source coding according to application requirements and hardware constraints by leveraging advanced multimedia encoding techniques; (iii) providing flexible and efficient system software, that is, operating systems and middleware to export services for higher-layer applications to build upon; and (iv) providing primitives for applications to leverage collaborative, advanced in-network multimedia processing techniques. This section provides an overview of these challenges.

1.5.1 Traffic Classes

Admission control has to be based on QoS requirements of the overlying application. We envision that WMSNs will need to provide support and differentiated service for several different classes of applications. In particular, they will need to provide differentiated service between real-time and delay-tolerant applications, and loss-tolerant and loss-intolerant applications. Moreover, some applications may require a continuous stream of multimedia data for a prolonged period of time (multimedia streaming), whereas some other applications may require event-triggered observations obtained in a short time period (snapshot multimedia content). The main traffic classes that need to be supported are

- *Real-time, loss-tolerant, multimedia streams*. This class includes video and audio streams, or multilevel streams composed of video/audio and other scalar data (e.g., temperature readings), as well as metadata associated with the stream, that need to reach a human or automated operator in real-time, that is, within strict delay bounds, and that are however relatively loss-tolerant (e.g., video streams can be within a certain level of distortion). Traffic in this class usually has a high bandwidth demand.
- *Delay-tolerant, loss-tolerant, multimedia streams*. This class includes multimedia streams that, being intended for storage or subsequent offline processing, does not need to be delivered within strict delay bounds. However, due to the typically high bandwidth demand of multimedia streams and limited buffers of multimedia sensors, data in this traffic class needs to be transmitted almost in real-time to avoid excessive losses.
- *Real-time, loss-tolerant, data*. This class may include monitoring data from densely deployed scalar sensors such as light sensors whose monitored phenomenon is characterized by spatial correlation or loss-tolerant snapshot multimedia data (e.g., images of a phenomenon taken from several multiple viewpoints at the same time). Hence, sensor data has to be received timely, but the application is moderately loss-tolerant. The bandwidth demand is usually between low and moderate.
- *Real-time, loss-intolerant, data*. This may include data from time-critical monitoring processes such as distributed control applications. The bandwidth demand varies between low and moderate.

- *Delay-tolerant, loss-intolerant, data*. This may include data from critical monitoring processes, with low or moderate bandwidth demand, that require some form of offline postprocessing.
- *Delay-tolerant, loss-tolerant, data*. This may include environmental data from scalar sensor networks, or non-time-critical snapshot multimedia content, with low or moderate bandwidth demand.

Table 1.1 presents a possible mapping of the applications, presented in Section 1.5, into the traffic classes described earlier.

1.5.2 Multimedia Encoding Techniques

There exists a vast literature on multimedia encoding techniques. The captured multimedia content should ideally be represented in such a way as to allow reliable transmission over lossy channels (error-resilient coding) using algorithms that minimize processing power and the amount of information to be transmitted. The main design objectives of a coder for multimedia sensor networks are thus

- *High compression efficiency*. Uncompressed raw video streams require high data rates and thus consume excessive bandwidth and energy. It is necessary to achieve a high ratio of compression to effectively limit bandwidth and energy consumption.
- *Low complexity*. Multimedia encoders are embedded in sensor devices. Hence, they need to be of low complexity to reduce cost and form factors, and of low power to prolong the lifetime of sensor nodes.
- *Error resiliency*. The source coder should provide robust and error-resilient coding of source data.

To achieve a high compression efficiency, the traditional broadcasting paradigm for wireline and wireless communications, where video is compressed once at the encoder and decoded several times, has been dominated by predictive encoding techniques. These, used in the widely spread International Standards Organization (ISO) Motion Picture Experts Group (MPEG) schemes, or the International Telecommunication Union – Telecommunication Standardization Sector (ITU-T) recommendations H.263 [20] and H.264 [21] (also known as AVC or MPEG-4 part 10), are based on the idea of reducing the bit rate generated by the source encoder by exploiting source statistics. Hence, intraframe compression techniques are used to reduce redundancy within one frame, whereas interframe compression exploits the correlation among subsequent frames to reduce the amount of data to be transmitted and stored, thus achieving good rate-distortion performance. Because the computational complexity is

Table 1.1 Classification of WMSN Applications and QoS Requirements

Application	*Classification (Data Type)*	*Level QoS*	*Mapped Bandwidth*	*Description*
Surveillance networks, monitoring elderly patients, person locator services	Multimedia	Real-time, loss-tolerant	High	Multilevel streams composed of video/audio and scalar data, as well as stream metadata, that need to reach the user in real time
Storage of potent relevant activities	Multimedia	Delay-tolerant, loss-tolerant	High	Streams intended for the storage of subsequent offline processing, but need to be delivered quickly owing to the limited buffers of sensors
Traffic avoidance, enforcement, and control systems	Data	Real-time, loss-tolerant	Moderate	Monitoring data from densely deployed spatially correlated scalar sensors or snapshot multimedia data (e.g., images of phenomena taken from simultaneous multiple viewpoints)
Industrial process control	Data	Real-time, loss-intolerant	Moderate	Data from time-critical monitoring processes such as distributed control applications
Structural health monitoring	Data	Delay-tolerant, loss-tolerant	Moderate	Data from monitoring processes that require some form of offline post-processing
Environmental monitoring	Data	Delay-tolerant, loss-tolerant	Low	Environmental data from scalar sensor networks, or non-time-critical snapshot multimedia content

dominated by the motion estimation functionality, these techniques require complex encoders, powerful processing algorithms, and entail high energy consumption, whereas decoders are simpler and loaded with lower processing burden. For typical implementations of state-of-the-art video compression standards, such as MPEG or H.263 and H.264, the encoder is five to ten times more complex than the decoder [13]. Hence, to realize low-cost, low-energy-consumption multimedia sensors it is necessary to develop simpler encoders and still retain the advantages of high compression efficiency.

However, it is known from information-theoretic bounds established by Slepian and Wolf for lossless coding [22] and by Wyner and Ziv [23] for lossy coding with decoder side information that efficient compression can be achieved by leveraging knowledge of the source statistics at the decoder only. This way, the traditional balance of complex encoder and simple decoder can be reversed [13]. Techniques that build upon these results are usually referred to as "distributed source coding." Distributed source coding refers to the compression of multiple correlated sensor outputs that do not communicate with one another [24]. Joint decoding is performed by a central entity that receives data independently compressed by different sensors. However, practical solutions have not been developed until recently. Clearly, such techniques are very promising for WMSNs and especially for networks of video sensors. The encoder can be simple and of low power, whereas the decoder at the sink will be complex and loaded with most of the processing and energy burden. For excellent surveys on the state-of-the-art of distributed source coding in sensor networks and in distributed video coding, see Refs 24 and 13, respectively. Other encoding and compression schemes that may be considered for source coding of multimedia streams, including JPEG with differential encoding, distributed coding of images taken by cameras having overlapping fields of view, or multilayer coding with wavelet compression, are discussed in Ref. 3. Here, we focus on recent advances on low-complexity encoders based on Wyner–Ziv coding [23], which are promising solutions for distributed networks of video sensors that are likely to have a major impact in future design of protocols for WMSNs.

The objective of a Wyner–Ziv video coder is to achieve lossy compression of video streams and achieve performance comparable to that of interframe encoding (e.g., MPEG), with complexity at the encoder comparable to that of intraframe coders (e.g., motion-JPEG).

1.5.2.1 Pixel-Domain Wyner–Ziv Encoder

In Refs 25 and 26, a practical Wyner–Ziv encoder is proposed as a combination of a pixel-domain intraframe encoder and interframe decoder system for video compression. A block diagram of the system is reported in Figure 1.4. A regularly spaced subset of frames are coded using a conventional intraframe coding technique, such as JPEG, as shown at the bottom

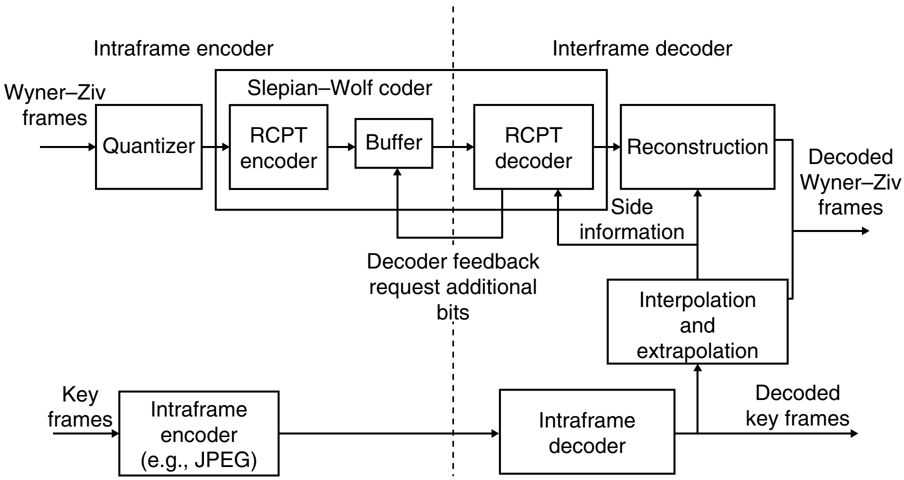

Figure 1.4 Block diagram of a pixel-domain Wyner–Ziv encoder. (From Aaron, A., Rane, S., Zhang, R., and Girod, B. Wyner–Ziv Coding for Video: Applications to Compression and Error Resilience. *Proc. of IEEE Data Compression Conf. (DCC)*, pp. 93–102. Snowbird, UT, March 2003. With permission.)

of the Figure 1.4. These are referred to as “key frames.” All frames between the key frames are referred to as “Wyner–Ziv frames” and are intraframe encoded but interframe decoded. The intraframe encoder for Wyner–Ziv frames (shown on top) is composed of a quantizer followed by a Slepian–Wolf coder. Each Wyner–Ziv frame is quantized and blocks of symbols are sent to the Slepian–Wolf coder, which is implemented through rate-compatible punctured turbo codes (RCPT). The parity bits generated by the RCPT coder are stored in a buffer. A subset of these bits is then transmitted on request from the decoder. This allows adapting the rate based on the temporally varying statistics between the Wyner–Ziv frame and the side information. The parity bits generated by the RCPT coder are in fact used to “correct” the frame interpolated at the decoder. For each Wyner–Ziv frame, the decoder generates the side information frame by interpolation or extrapolation of previously decoded key frames and Wyner–Ziv frames. The side information is leveraged by assuming a Laplacian distribution of the difference between the individual pixels of the original frame and the side information. The parameter defining the Laplacian distribution is estimated online. The turbo decoder combines the side information and the parity bits to reconstruct the original sequence of symbols. If reliable decoding of the original symbols is impossible, the turbo decoder requests additional parity bits from the encoder buffer.

Compared to predictive coding such as MPEG or H.26X, pixel-domain Wyner–Ziv encoding is much simpler. The Slepian–Wolf encoder only requires two feedback shift registers and an interleaver. Its performance, in terms of peak signal-to-noise ratio (PSNR), is 2–5 dB better than the conventional motion-JPEG intraframe coding. The main drawback of this scheme is that it relies on online feedback from the receiver. Hence it may not be suitable for applications where video is encoded and stored for subsequent use. Moreover, the feedback may introduce excessive latency for video decoding in a multihop network.

1.5.2.2 Transform-Domain Wyner–Ziv Encoder

In conventional source coding, a source vector is typically decomposed into spectral coefficients by using orthonormal transforms such as the discrete cosine transform (DCT). These coefficients are then individually coded with scalar quantizers and entropy coders. In Ref. 27, a transform-domain Wyner–Ziv encoder is proposed. A blockwise DCT of each Wyner–Ziv frame is performed. The transform coefficients are independently quantized, grouped into coefficient bands, and then compressed by a Slepian–Wolf turbo coder. As in the pixel-domain encoder described in Section 1.5.2.1, the decoder generates a side information frame based on previously reconstructed frames. Based on the side information, a bank of turbo decoders reconstructs the quantized coefficient bands independently. The rate-distortion performance is between conventional intraframe transform coding and conventional motion-compensated transform coding.

A different approach consists of allowing some simple temporal dependence estimation at the encoder to perform rate control without the need for feedback from the receiver. In the PRISM scheme [28], the encoder selects the coding mode based on the frame difference energy between the current frame and a previous frame. If the energy of the difference is very small, the block is not encoded. If the block difference is large, the block is intracoded. Between these two situations, one of the different encoding modes with different rates is selected. The rate estimation does not involve motion compensation and hence is necessarily inaccurate if motion compensation is used at the decoder. Further, the flexibility of the decoder is restricted.

1.5.3 System Software and Middleware

The development of efficient and flexible system software to make functional abstractions and information gathered by scalar and multimedia sensors available to higher-layer applications is one of the most important challenges faced by researchers to manage complexity and heterogeneity of sensor systems. As in Ref. 29, the term "system software" is used here to refer to operating systems, virtual machines, and middleware, which export services to higher-layer applications. Different multimedia sensor network applications are extremely diverse in their requirements and in the way they interact with the components of a sensor system. Hence, the main desired characteristics of a system software for WMSNs can be identified as follows:

- Provides a high-level interface to specify the behavior of the sensor system. This includes semantically rich querying languages that allow specifying what kind of data is requested from the sensor network, the quality of the required data, and how it should be presented to the user.
- Allows the user to specify application-specific algorithms to perform in-network processing on the multimedia content [30]. For example, the user should be able to specify particular image-processing algorithms or multimedia coding format.
- Long-lived, that is, needs to smoothly support evolutions of the underlying hardware and software.
- Shared among multiple heterogeneous applications.
- Shared among heterogeneous sensors and platforms. Scalar and multimedia sensor networks should coexist in the same architecture, without compromising on performance.
- Scalable.

There is an inherent trade-off between degrees of flexibility and network performance. Platform independence is usually achieved through layers of abstraction, which usually introduces redundancy and prevents the developer from accessing low-level details and functionalities. However, WMSNs are characterized by the contrasting objectives of optimizing the use of the scarce network resources and not compromising on performance. The principal design objective of existing operating systems for sensor networks such as TinyOS is high performance. However, their flexibility, interoperability, and reprogrammability are very limited. There is a need for research on systems that allow for this integration.

We believe that it is of paramount importance to develop efficient, high-level abstractions that will enable easy and fast development of sensor network applications. An abstraction similar to the famous Berkeley Transmission Control Protocol (TCP) sockets, that fostered the development of Internet applications, is needed for sensor systems. However, different from the Berkeley sockets, it is necessary to retain control on the efficiency of the low-level operations performed on battery-limited and resource-constrained sensor nodes.

As a first step toward this direction, Chu et al. [31] recently proposed the Sensor data Library (Sdlib), a sensor network data and communications library built upon the nesC language [32] for applications that require best-effort collection of large-size data such as video monitoring applications. The objective of the effort is to identify common functionalities shared by several sensor network applications and to develop a library of thoroughly tested, reusable, and efficient nesC components that abstract high-level operations common to most applications, although leaving differences among them to adjustable parameters. The library is called Sdlib as an analogy to the traditional C++ Standard Template Library. Sdlib provides an abstraction for common operations in sensor networks although the developer is still able to access low-level operations, which are implemented as a collection of nesC components, when desired. Moreover, to retain the efficiency of operations that are so critical for sensor networks battery lifetime and resource constraints, Sdlib exposes policy decisions such as resource allocation and rate of operation to the developer, although hiding the mechanisms of policy enforcement.

1.5.4 Collaborative In-Network Processing

As discussed earlier, collaborative in-network multimedia processing techniques are of great interest in the context of a WMSN. It is necessary to develop architectures and algorithms to flexibly perform these functionalities in-network with minimum energy consumption and limited execution time. The objective is usually to avoid transmitting large amounts of raw streams to the sink by processing the data in the network to reduce the communication volume.

Given a source of data (e.g., a video stream), different applications may require diverse information (e.g., raw video stream versus simple scalar or binary information inferred by processing the video stream). This is referred to as application-specific querying and processing. Hence, it is necessary to develop expressive and efficient querying languages, and to develop distributed filtering and in-network processing architectures, to allow real-time retrieval of useful information.

Similarly, it is necessary to develop architectures that efficiently allow performing data fusion or other complex processing operations in-network. Algorithms for both inter- and intramedia data aggregation and fusion need to be developed, as simple distributed processing schemes developed for existing scalar sensors are not suitable for computation-intensive processing required by multimedia contents. Multimedia sensor networks may require computation-intensive processing algorithms (e.g., to detect the presence of suspicious activity from a video stream). This may require considerable processing to extract meaningful information and to perform compression. A fundamental question to be answered is whether this processing can be done on sensor nodes (i.e., a flat architecture of multifunctional sensors that can perform any task), or if the need for specialized devices, for example, "computation hubs," arises.

1.5.5 Open Research Issues

- Although theoretical results on Slepian–Wolf and Wyner–Ziv codings exist since 30 years, there is still a lack of practical solutions. The net benefits and the practicality of these techniques still need to be demonstrated.
- It is necessary to completely explore the trade-offs between the achieved fidelity in the description of the phenomenon observed and the resulting energy consumption. As an example, the video distortion perceived by the final user depends on source coding (frame rate, quantization) and channel coding strength. For example, in a surveillance application, the objective of maximizing the event detection probability is in contrast to the objective of minimizing the power consumption.
- As discussed earlier, there is a need for higher-layer abstractions that will allow the fast development of sensor applications. However, due to the resource-constrained nature of sensor systems, it is necessary to control the efficiency of the low-level operations performed on battery-limited and resource-constrained sensor nodes.
- There is a need for simple, yet expressive high-level primitives for applications to leverage collaborative, advanced in-network multimedia processing techniques.

1.6 Protocols

In applications involving high-rate data, the transport, network, and link layers assume special importance by providing congestion control, delay-bounded routing, fair and efficient scheduling, among other functionalities. Although a large body of work exists in this area, there is a growing trend for newer design approaches in the context of WMSNs. As an example, the User Datagram Protocol (UDP), which has been traditionally used for transporting multimedia content may not be suited for WMSNs, given its blind packet drop policy. There is a need to decouple network-layer reliability and congestion control, often weighting one over the other, to meet performance requirements of real-time traffic. Similarly, at the link layer, protocols allow higher bandwidth utilization through multiple-transceiver sensors, multiple input multiple output (MIMO) antennas may be devised.

Recent research efforts have stressed on cross-layer design principles, wherein information gathered at a given layer can be used for making better decisions in the protocols that operate at entirely different layers. As an example, high packet delays and low bandwidth can force the routing layer to change its route decisions. Different routing decisions alter the set of links to be scheduled, thereby influencing the performance of the Media

Access Control (MAC) layer. Furthermore, congestion control and power control are also inherently coupled [33], as the capacity available on each link depends on the transmission power. Moreover, specifically to multimedia transmissions, the application layer does not require full insulation from lower layers, but instead needs to perform source coding based on information from the lower layers to maximize the multimedia performance.

This section discusses the considerations in optimizing the performance of the protocols operating at each of the layers, as well as exploring their interdependencies.

1.6.1 Transport-Layer Protocols

To explore the functionalities and support provided by the transport layer, the following discussion is classified into (1) TCP/UDP and TCP-friendly schemes for WMSNs and (2) application-specific and nonstandardized protocols. Figure 1.5 summarizes the discussion in this section.

1.6.1.1 TCP/UDP and TCP-Friendly Schemes for WMSNs

For real-time applications such as streaming media, UDP is preferred over TCP as "timeliness" is of greater concern than "reliability." However, in WMSNs, it is expected that packets are significantly compressed at the source, and redundancy is reduced as far as possible owing to the high transmission overhead in the energy-constrained nodes. Simply dropping packets during congestion conditions, as undertaken in UDP, may introduce discernable disruptions if they contain important original content not captured by interframe interpolation, such as the region of interest (ROI) feature used in JPEG2000 [34] or the I-frame used in the MPEG family.

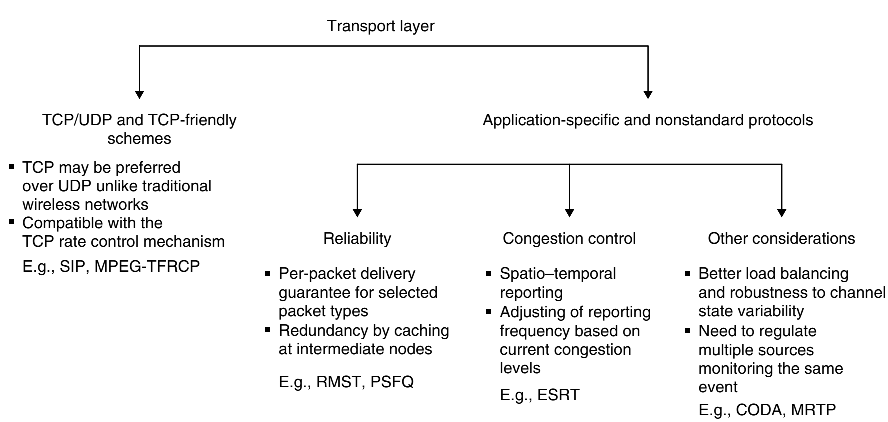

Figure 1.5 Classification of existing transport-layer protocols.

Unlike UDP, the TCP header can be modified to carry data-specific information including minimum resource requirements of different service classes in case of differentiated services.

We thus believe that TCP with appropriate modifications is preferable over UDP for WMSNs, if standardized protocols are to be used. With respect to sensor networks, several problems and their likely solutions such as large TCP header size, data versus address-centric routing, energy efficiency, among others, are identified and solutions are proposed in Ref. 35. We next indicate the recent work in this direction that evaluates the case for using TCP in WMSNs.

- *Effect of jitter induced by TCP*. A key factor that limits multimedia transport based on TCP and TCP-like rate control schemes is the jitter introduced by the congestion control mechanism. This can, however, be mitigated to a large extent by playout buffers at the sink, which is typically assumed to be rich in resources. As an example, the MPEG-TFRCP (TCP-Friendly Rate Control Protocol for MPEG-2 Video Transfer) [36] is an equation-based rate control scheme designed for transporting MPEG video in a TCP-friendly manner.
- *Overhead of the reliability mechanism in TCP*. The reliability mechanism provided by TCP introduces an end-to-end message passing overhead. Distributed TCP caching (DTC) [37] overcomes these problems by caching TCP segments inside the sensor network and by local retransmission of TCP segments. The nodes closest to the sink are the last-hop forwarders on most of the high-rate data paths and thus run out of energy first. DTC shifts the burden of the energy consumption from nodes close to the sink to the network, apart from reducing networkwide retransmissions.
- *Regulating streaming through multiple TCP connections*. The availability of multiple paths between the source and sink can be exploited by opening multiple TCP connections for multimedia traffic [38]. Here, the desired streaming rate and the allowed throughput reduction in presence of bursty data, such as video, is communicated to the receiver by the sender. This information is used by the receiver, which then measures the actual throughput and controls the rate within the allowed bounds by using multiple TCP connections and dynamically changing its TCP window size for each connection.

Although the recent TCP-based protocols for sensor networks, such as the Sensor Internet Protocol (SIP) [39] and the open source uIP [35] look promising, the inability to distinguish between bad channel conditions and network congestion is a major problem in TCP. This has motivated a new family of specialized transport-layer protocols where the design practices followed are entirely opposite to that of TCP [40], or stress on a particular functionality of the transport layer such as reliability or congestion control.

1.6.1.2 Application-Specific and Nonstandard Protocols

Depending on the application, both reliability and congestion control may be equally important functionalities or one may be preferred over the other. As an example, a video capture of a moving target enjoys a permissible level of loss tolerance and should be prevented from consuming a large proportion of network resources. The presence or absence of an intruder, however, may require a single data field but needs to be communicated without any loss of fidelity or information. We next list the important characteristics of such TCP-incompatible protocols in the context of WMSNs.

- *Reliability*. As discussed earlier in this section, certain packets in a flow may contain important information that is not replicated elsewhere, or cannot be regenerated through interpolation. Thus, if a prior recorded video is being sent to the sink, all the I-frames could be separated and the transport protocol should guarantee that each of these reach the sink. Reliable Multisegment Transport (RMST) [41] or the Pump Slowly Fetch Quickly (PSFQ) protocol [42] can be used for this purpose as they buffer packets at intermediate nodes, allowing for faster retransmission in case of packet loss. However, there is an overhead of using the limited buffer space at a given sensor node for caching packets destined for other nodes, as well as performing timely storage and flushing operations on the buffer.
- *Congestion control*. Multimedia data rates are typically in the order of 64 kbit/s for constant bit rate voice traffic, whereas video traffic may be bursty and approximately 500 kbit/s [43]. Although these data generation rates are high for a single node, multiple sensors in overlapped regions may inject similar traffic on sensing the same phenomenon, leading to network congestion. The Event-to-Sink Reliable Transport (ESRT) protocol [44] leverages the fact that spatial and temporal correlation exists among the individual sensor readings [45]. The ESRT protocol regulates the frequency of event reporting in a remote neighborhood to avoid congestions in the network. However, this approach may not be viable for all sensor applications as nodes transmit data only when they detect an event, which may be a short duration burst, as in the case of a video monitoring application. The feedback from the base station may hence not reach in time to prevent a sudden congestion due to this burst.
- *Other considerations*. As in TCP implementations, the use of multiple paths improves the end-to-end data transfer rate. The COngestion

Detection and Avoidance (CODA) protocol [46] allows a sink to regulate multiple sources associated with a single event in case of persistent network congestion. However, as the congestion inference in CODA is based on queue length at intermediate nodes, any action taken by the source occurs only after a considerable time delay. Other solutions include the Multiflow Real-time Transport Protocol (MRTP) [47] that is suited for real-time streaming of multimedia content by splitting packets over different flows, but does not consider energy efficiency or retransmissions for scalar data, both important considerations for a heterogeneous WMSN.

The success in energy-efficient and reliable delivery of multimedia information extracted from the phenomenon directly depends on selecting the appropriate coding rate, number of sensor nodes, and data rate for a given event [45]. For this purpose, new reliability metrics coupled with the application-layer coding techniques should be investigated.

1.6.2 Network Layer

Data collected by the sensor nodes may consist of different traffic classes, each having its own QoS requirement. We focus our discussion on the primary network-layer functionality of routing, although stressing on the applicability of the existing protocols for delay-sensitive and high bandwidth needs.

The concerns of routing, in general, differ significantly from the specialized service requirements of multimedia streaming applications. As an example, multiple routes may be necessary to satisfy the desired data rate at the destination node. Also, different paths exhibiting varying channel conditions may be preferred depending on the type of traffic and its resilience to packet loss. Primarily, routing schemes can be classified based on (i) network conditions that leverage channel and link statistics, (ii) traffic classes that decide paths based on packet priorities, and (iii) specialized protocols for real-time streaming that use spatio–temporal forwarding. Figure 1.6 provides a classification of the existing routing protocols and summarizes the discussion in this section.

1.6.2.1 QoS Routing Based on Network Conditions

Network conditions include interference seen at intermediate hops, the number of backlogged flows along a path, residual energy of the nodes, among others. A routing decision based on these metrics can avoid paths that may not support high bandwidth applications or introduce retransmission owing to bad channel conditions.

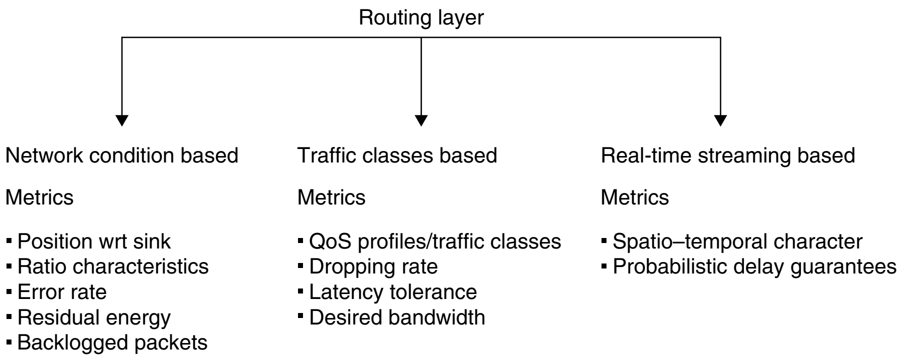

Figure 1.6 Classification of existing routing protocols.

The use of image sensors for gathering topology information [48] and generating QoS metrics [49] have been recently explored. Both construct a weighted cost function that takes into account different factors such as position, backlogged packets in the queue, and residual energy and communication parameters such as error rate to derive energy-efficient paths bounded by maximum allowed delays.

1.6.2.2 QoS Routing Based on Traffic Classes

As an example that highlights the need for network-level QoS, consider the task of bandwidth assignment for multimedia mobile medical calls, which include patients' sensing data, voice, pictures, and video data [8]. Unlike the typical source-to-sink multihop communication used by classical sensor networks, the proposed architecture uses a 3G cellular system in which individual nodes forward the sensed data to a cellular phone or a specialized information-collecting entity taking into account handoff effects. Different priorities are assigned to video data originating from sensors on ambulances, audio traffic from elderly people, and images returned by sensors placed on the body. It is important that some flows be preferred over the others given their time-critical nature and the level of importance to the patient's treatment.

1.6.2.3 Routing Protocols with Support for Streaming

The SPEED protocol [50] provides three types of real-time communication services, namely, real-time unicast, real-time area-multicast, and real-time area-anycast. It uses geographical location for routing and a key difference with other schemes of this genre is its spatio–temporal character, that is, it takes into account the timely delivery of the packets. As it works satisfactorily under scarce resource conditions and can provide service differentiation, SPEED takes the first step in addressing the concerns of real-time routing in WMSNs.

A significant extension over SPEED, the MMSPEED protocol [51] can efficiently differentiate between flows with different delay and reliability requirements. MMSPEED is based on a cross-layer approach between the network and the MAC layers in which a judicious choice is made over the reliability and timeliness of packet arrival. It is argued that the differentiation in reliability is an effective way of channeling resources from flows with relaxed requirements to flows with tighter requirements. Although current research directions make an effort to provide real-time streaming, they are still best-effort services.

Giving firm delay guarantees and addressing QoS concerns in a dynamically changing network is a difficult problem and yet is important for seamless viewing of the multimedia frames. Although probabilistic approaches such as MMSPEED take the first step toward this end, clearly, further work is needed in the networking layer.

1.6.3 MAC Layer

Research efforts to provide MAC-layer QoS can be mainly classified into (i) channel access policies, (ii) scheduling and buffer management, and (iii) error control. We next provide a brief description of each and highlight their support to multimedia traffic. The scope of this chapter is limited to the challenges posed by multimedia traffic in sensor networks and the efforts at the MAC layer to address them. A detailed survey of MAC protocols for classical sensor networks using scalar data can be found in Ref. 52. Figure 1.7 provides a classification of relevant MAC-layer functionalities and summarizes the discussion in this section.

1.6.3.1 Channel Access Policies

The main causes of energy loss in sensor networks are attributed to "packet collisions" and subsequent retransmissions, "overhearing packets" destined for other nodes, and "idle listening," a state in which the transceiver circuits remain active even in the absence of data transfer. Thus, regulating access to the channel assumes primary importance and several solutions have been proposed in the literature.

1.6.3.2 Contention-Based Protocols

Most contention-based protocols such as S-MAC [53], and protocols inspired by it [54], have a single-radio architecture. They alternate between sleep cycles (low power modes with transceiver switched off) and listen cycles (for channel contention and data transmission). However, we believe

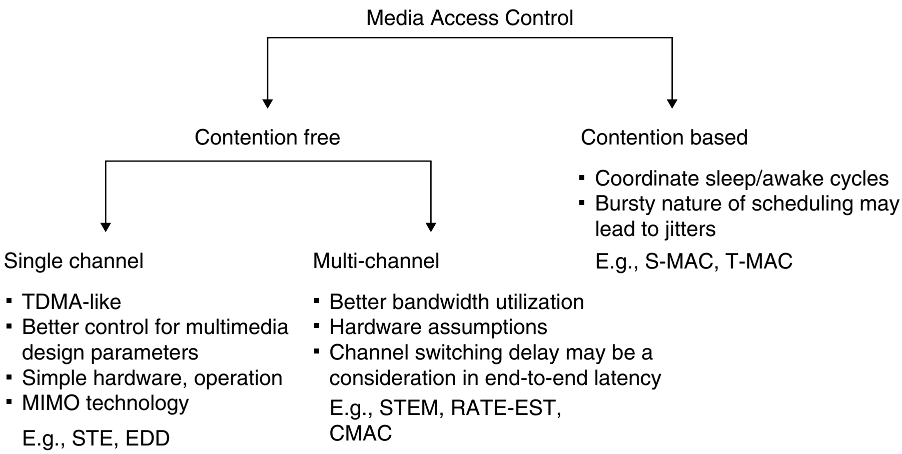

Figure 1.7 Classification of protocols at the data link layer.

that their applicability to multimedia transmission is limited owing to the following reasons:

- The primary concern in the protocols of this class is saving energy, and this is accomplished at the cost of latency and by allowing throughput degradation. A sophisticated duty cycle calculation based on permissible end-to-end delay needs to be implemented and coordinating overlapping "listen" period with neighbors based on this calculation is a difficult research challenge.
- Video traffic exhibits an inherent bursty nature and can lead to sudden buffer overflow at the receiver. This problem is further aggravated by the transmission policy adopted in T-MAC [54]. By choosing to send a burst of data during the listen cycle, T-MAC shows performance improvement over S-MAC, but at the cost of monopolizing a bottleneck node. Such an operation could well lead to strong jitters and result in discontinuous real-time playback.

1.6.3.3 Contention-Free Single-Channel Protocols

Time division multiple access (TDMA) is a representative protocol of this class in which the CH or sink helps in slot assignment, querying particular sensors, and maintaining time schedules. We believe that such protocols can be easily adapted for multimedia transmission and highlight the likely design considerations.

- TDMA schemes designed exclusively for sensor networks [55; and references therein] have a small reservation period (RP) that is generally contention-based, followed by a contention-free period that spans the rest of the frame. This RP could occur in each frame or at predecided intervals to assign slots to active nodes taking into consideration the QoS requirement of their data streams. The length of the TDMA frames and the frequency of the RP interval are some of the design parameters that can be exploited when designing a multimedia system. However, clock drift and synchronization issues must be accounted for they have a pronounced effect on small TDMA slot sizes.

- For real-time streaming video, packets are time-constrained and scheduling policies such as shortest time to extinction (STE) [56] or earliest due date (EDD) [57] can be adopted. Both of these are similar in principle as packets are sent in the increasing order of their respective delay tolerance, but differ in respect that EDD may still forward a packet that has crossed its allowed delay bound. Based on the allowed packet loss of the multimedia stream, the dependencies between packet dropping rate, arrival rate, and delay tolerance [56] can be used to decide the TDMA frame structure and thus ensure the smooth replay of data. This allows greater design choices as against Ref. 57, where the frame lengths and slot duration are considered constant.
- As sensor nodes are often limited by their maximum data transmission rate, depending on their multimedia traffic class, the duration of transmission could be made variable. Thus, variable TDMA (V-TDMA) schemes should be preferred when heterogeneous traffic is present in the network. Tools for calculating the minimum worst-case per-hop delay provided in Ref. 58 could be used in conjunction with the allowed end-to-end delay when V-TDMA schemes are used.
- The high data rate required by multimedia applications can be addressed by spatial multiplexing in MIMO systems that use a single channel but employ interference cancellation techniques. Recently, “virtual” MIMO schemes have been proposed for sensor networks [59], where nodes in proximity form a cluster. Each sensor functions as a single antenna element, sharing information and thus simulating the operation of a multiple antenna array. A distributed compression scheme for correlated sensor data, which specially addresses multimedia requirements, is integrated with the MIMO framework in Ref. 60. However, a key consideration in the MIMO-based systems is the number of sensor transmissions and the required signal energy per transmission. As the complexity is shifted from hardware to sensor coordination, further research is needed at the MAC layer to ensure that the required MIMO parameters such as channel state and desired diversity/processing gain are known to both the sender and receiver at an acceptable energy cost.

1.6.3.4 Contention-Free Multi-Channel Protocols

As discussed earlier in this section, existing bandwidth can be efficiently utilized by using multiple channels in a spatially overlapped manner. We observe that Berkeley's 3G MICA2 Mote has an 868/916 GHz multi-channel transceiver [5]. In Rockwell's WINS nodes, the radio operates on one of the 40 channels in the Industrial, Scientific and Medical (ISM) frequency band, selectable by the controller [61]. We next outline the design parameters that could influence MAC design in multi-channel WMSNs.

- Recent research has focused on a two-transceiver paradigm in which the main radio (MR) is supplemented by the presence of a wake-up radio (LR) having similar characteristics [62–64] or a simple low-energy design [65] that emits a series of short pulses or a busy tone. The LR is used for exchanging control messages and is assigned a dedicated channel. In high bandwidth applications, such as streaming video, the use of a separate channel for channel arbitration alone does not allow best utilization of the network resources.
- We propose that WMSNs use in-band signaling, where the same channel is used for both data and channel arbitration [66]. Although such protocols undoubtedly improve bandwidth efficiency, they introduce the problem of distinct channel assignment and need to account for the delay to switch to a different channel [67], as its cumulative nature at each hop affects real-time media. This work leaves an open question on whether switching delay can be successfully hidden with only one interface per node. If this is possible, it may greatly simplify sensor design, while performing, as well as a multi-channel, multi-interface solution.
- Multi-channel protocols are not completely collision-free as seen in the case of control packet collision [62,63,66], and the available channels cannot be assumed to be perfectly nonoverlapping. This may necessitate dynamic channel assignment, taking into account the effect of adjacent channel interference, to maintain the network QoS.

The use of multi-channel communication paradigm can also be used to highlight the need for cross-layer design strategies, described in detail in Section 1.6.4. Consider the scenario in Figure 1.8 in which a multiradio architecture is assumed. Existing data rates of approximately 40 and 250 kbit/s supported by the MICA2 and MICAz motes are not geared to support multimedia traffic and hence the use of multiple channels for the same flow greatly increases the data rate. The network layer must preferentially route packets to nodes that can support reception on more than one channel. The link layer undertakes sender–receiver negotiations and decides on the particular channels to be used for the impending data transfer.

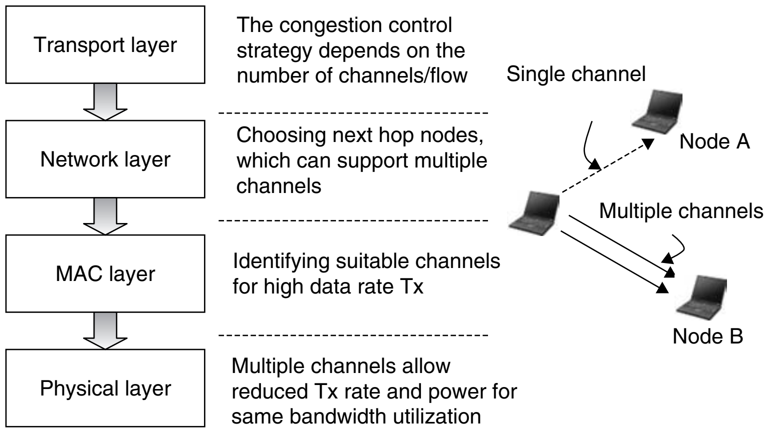

Figure 1.8 Interdependencies in the protocol stack leading to a cross-layer architecture in WMSNs.

The number of channels to be used may decide the congestion control algorithm adopted at the transport layer. Thus, although these complex interdependencies help in realizing high data rates on the resource-constrained sensor nodes, they also introduce significant challenges in protocol design and must be investigated further.

1.6.3.5 Scheduling

MAC-layer scheduling in the context of WMSNs differs from the traditional networking model in the sense that apart from choosing the queueing discipline that accounts for latency bounds, rate/power control, and consideration of high channel error conditions needs to be incorporated.

To generate optimal schedules that minimize both power consumption and the probability of missing deadlines for real-time messages, PARM [68] integrates the EDD metric described in Section 1.6.3 into an energy consumption function. Although significant performance improvements are demonstrated, this work needs to be extended for large-scale networks that are typically envisaged for WMSNs.

Queueing at the MAC layer has been extensively researched and several schemes with varying levels of complexity exist. Of interest to multimedia applications is the development of schemes that allow a delay bound and thus assure smooth streaming of multimedia content. E^2WFQ [69], a variant of the established weighted fair queueing (WFQ) discipline, allows adjustments to be made to the energy–latency–fidelity trade-off space. Extending WFQ, the wireless packet scheduling (WPS) presented in Ref. 70, addresses the concerns of delay and rate-sensitive packet flows, thus making it suitable for multimedia traffic. WPS, however, assumes that the channel error is completely predictable at any time and its practical implementation shows marked deviations from the idealized case in terms of worst-case complexity. This work is suitable for single-hop sensor–sink communication and multihop forwarding issues are not explored.

Network calculus [71,72] is a theory for deterministic queueing systems that allows the assignment of service guarantees by traffic regulation and deterministic scheduling. Through tools provided by network calculus, bounds on various performance measures, such as delay and queue length, at each element of the network can be derived and thus QoS of a flow can be specified. Arrival, Departure, and Service curves reflect the constraints that flows are subjected to within a network. The calculus relies on Min-plus algebra, in which addition and multiplication are replaced by minimum and addition, respectively, to operate on these curves. Current network calculus results have been mostly derived for wired networks, and assume static topologies and fixed link capacity, indicating that further work needs to be undertaken in dynamic scenarios typically envisaged for WMSNs.

1.6.3.6 Link-Layer Error Control

The inherent unreliability of the wireless channel, coupled with a low-frame loss rate requirement of the order of 10^{-2} for good-quality video, poses a challenge in WMSNs. Two main classes of mechanisms are traditionally employed to combat the unreliability of the wireless channel at the physical (PHY) and data link layer, namely forward error correction (FEC) and automatic repeat request (ARQ), along with hybrid schemes. ARQ mechanisms use bandwidth efficiently at the cost of additional latency. Hence, although carefully designed selective repeat schemes may be of some interest, naive use of ARQ techniques is clearly infeasible for applications requiring real-time delivery of multimedia content.

An important characteristic of multimedia content is "unequal importance," that is, not all packets have the same importance, as seen in the case of the I-frames in the MPEG family. Applying different degrees of FEC to different parts of the video stream, depending on their relative importance (unequal protection) allows a varying overhead on the transmitted packets. For example, this idea can be applied to layered coded streams to provide graceful degradation in the observed image quality in the presence of error losses, thus avoiding the so-called "cliff" effects [2].

In general, delivering error-resilient multimedia content and minimizing energy consumption are contradicting objectives. For this reason, and due to the time-varying characteristics of the wireless channel, several joint source and channel coding schemes have been developed (e.g., Ref. 14), which try to reduce the energy consumption of the whole process. Some recent papers [73,74] even try to jointly reduce the energy consumption of the whole process of multimedia content delivery, that is, jointly optimize source coding, channel coding, and transmission power control. However, most of these efforts have originated from the multimedia or coding communities, and thus do not jointly consider other important networking aspects of content delivery over a multihop wireless networks of memory-, processing-, and battery-constrained devices.

In Ref. 75, a cross-layer analysis of error control schemes for WSNs is presented. The effects of multihop routing and of the broadcast nature of wireless communications are investigated to model the energy consumption, latency, and packet error rate performance of error control schemes. As a result, error control schemes are studied through a cross-layer analysis that considers the effects of routing, medium access, and physical layer. This analysis enables a comprehensive comparison of FEC and ARQ schemes in WSNs. FEC schemes are shown to improve the error resiliency compared to ARQ. In a multihop network, this improvement can be exploited by reducing the transmit power (transmit power control) or by constructing longer hops (hop length extension) through channel-aware routing protocols. The analysis reveals that, for certain FEC codes, hop length extension decreases both the energy consumption and the end-to-end latency subject to a target packet error rate compared to ARQ. Thus, FEC codes are an important candidate for delay-sensitive traffic in WSNs. However, transmit power control results in significant savings in energy consumption at the cost of increased latency. There is also a need to integrate source and channel coding schemes in existing cross-layer optimization frameworks. The existing schemes mostly consider point-to-point wireless links, and neglect interference from neighboring devices and multihop routes.

1.6.4 Cross-Layer Protocols

Although a consistent amount of recent papers have focused on cross-layer design and improvement of protocols for WSNs, a systematic methodology to accurately model and leverage cross-layer interactions is still largely missing. Most of the existing studies decompose the resource allocation problem at different layers, and consider allocation of the resources at each layer separately. In most cases, resource allocation problems are treated either heuristically, or without considering cross-layer interdependencies, or by considering pairwise interactions between isolated pairs of layers.

In Ref. 76, the cross-layer transmission of multimedia content over wireless networks is formalized as an optimization problem. Several different approaches for cross-layer design of multimedia communications are discussed, including “bottom–up approach,” where the lower layers try to insulate the higher layers from losses and channel capacity variations, and “top–down,” where the higher layer protocols optimize their parameters at the next lower layer. However, only single-hop networks are considered.

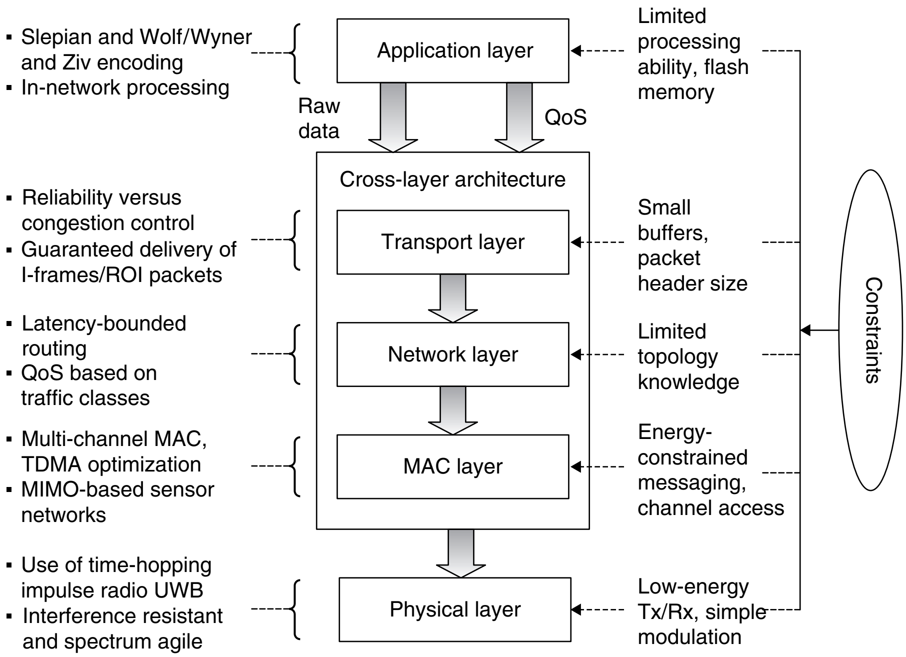

Figure 1.9 Cross-layer communication architecture of a multimedia sensor.

In Ref. 77, several techniques that provide significant performance gains through cross-layer optimizations are surveyed. In particular, the improvements of adaptive link-layer techniques such as adaptive modulation and packet-size optimization, joint allocation of capacity and flows (i.e., MAC and routing), joint scheduling, and rate allocation are discussed. Although still maintaining a strict layered architecture, it is shown how these cross-layer optimizations help to improve the spectral efficiency at the physical layer, and the PSNR of the video stream perceived by the user. Clearly, energy-constrained multimedia sensors may need to leverage cross-layer interactions one step further. At the same time, optimization metrics in the energy domain need to be considered as well.

Figure 1.9 summarizes our discussion on protocol design showing how application-layer QoS requirements and design constraints of WMSN can be addressed by carefully designing a cross-layer communication architecture.

1.7 Physical Layer

This section discusses the applicability of the UWB transmission technique, which we advocate as a potential technology for multimedia sensors.

1.7.1 Ultra Wideband Communications

The UWB* technology has the potential to enable low power consumption and high data rate communications within tens of meters, characteristics that make it an ideal choice for WMSNs.

UWB signals have been used for several decades in the radar community. Recently, the U.S. Federal Communications Commission (FCC) Notice of Inquiry in 1998 and the First Report and Order in 2002 [79] inspired a renewed flourish of research and development efforts in both academy and industry due to the characteristics of UWB that make it a viable candidate for wireless communications in dense multipath environments.

Although UWB signals, as per the specifications of the FCC, use the spectrum from 3.1 to 10.6 GHz, with appropriate interference limitation, UWB devices can operate using spectrum occupied by existing radio services without causing interference, thereby permitting scarce spectrum resources to be used more efficiently. Instead of dividing the spectrum into distinct bands that are then allocated to specific services, UWB devices are allowed to operate overlaid and thus interfere with existing services, at a low enough power level that existing services would not experience performance degradation. The First Report and Order by the FCC includes standards designed to ensure that existing and planned radio services, particularly safety services, are adequately protected.

There exist two main variants of UWB. The first, known as time-hopping impulse radio UWB (TH-IR-UWB) [78], and mainly developed by Win and Scholtz [80], is based on sending very short duration pulses (in the order of hundreds of picoseconds) to convey information. Time is divided into frames, each of which is composed of several chips of very short duration. Each sender transmits one pulse in a chip per frame only, and multi-user access is provided by pseudorandom time-hopping sequences (THS) that determine in which chip each user should transmit. A different approach, known as multicarrier UWB (MC-UWB), uses multiple simultaneous carriers, and is usually based on Orthogonal Frequency Division Multiplexing (OFDM) [81].

MC-UWB is particularly well suited for avoiding interference because its carrier frequencies can be precisely chosen to avoid narrowband interference to or from narrowband systems. However, implementing an MC-UWB front-end power amplifier can be challenging due to the continuous variations in power over a very wide bandwidth. Moreover, when OFDM is

* The U.S. Federal Communications Commission (FCC) defines UWB as a signal with either a fractional bandwidth of 20 percent of the center frequency or 500 MHz (when the center frequency is above 6 GHz). The FCC calculates the fractional bandwidth as $2(f_H - f_L)/(f_H + f_L)$, where f_H represents the upper frequency of the -10 dB emission limit and f_L represents the lower frequency limit of the -10 dB emission limit [78].

used, high-speed FFT processing is necessary, which requires significant processing power and leads to complex transceivers.

TH-IR-UWB signals require fast switching times for the transmitter and receiver and highly precise synchronization. Transient properties become important in the design of the radio and antenna. The high instantaneous power during the brief interval of the pulse helps to overcome interference to UWB systems, but increases the possibility of interference from UWB to narrowband systems. The RF front end of a TH-IR-UWB system may resemble a digital circuit, thus circumventing many of the problems associated with mixed-signal integrated circuits. Simple TH-IR-UWB systems can be very inexpensive to construct.

Although no sound analytical or experimental comparison between the two technologies is available to our knowledge, we believe that TH-IR-UWB is particularly appealing for WMSNs for the following reasons:

- It enables high data rate, very low-power wireless communications, on simple-design, low-cost radios (carrierless, baseband communications) [80].
- Its fine delay resolution properties are appropriate for wireless communications in dense multipath environment, by exploiting more resolvable paths [82].
- It provides large processing gain in the presence of interference.
- It provides flexibility, as data rate can be traded for power spectral density and multipath performance.
- Finding suitable codes for THS is trivial (as opposed to CDMA codes), and no assignment protocol is necessary.
- It naturally allows for integrated MAC/PHY solutions [83]. Moreover, interference mitigation techniques [83] allow realizing MAC protocols that do not require mutual temporal exclusion between different transmitters. In other words, simultaneous communications of neighboring devices are feasible without complex receivers as required by CDMA.
- Its large instantaneous bandwidth enables fine time resolution for accurate position estimation [84] and for network time distribution (synchronization).
- UWB signals have extremely low-power spectral density, with low probability of intercept/detection (LPI/D), which is particularly appealing for military covert operations.

1.7.1.1 Ranging Capabilities of UWB

Particularly appealing for WMSNs are UWB high data rate with low power consumption, and its positioning capabilities. Positioning capabilities are needed in sensor networks to associate physical meaning to the information gathered by sensors. Moreover, knowledge of the position of each network device allows for scalable routing solutions [85]. Angle-of-arrival techniques and signal strength-based techniques do not provide considerable advantages with respect to other transmission techniques, whereas time-based approaches in UWB allow ranging accuracy in the order of centimeters [84]. This can be intuitively explained by expression 1.1, which gives a lower bound on the best achievable accuracy of a distance estimate $\hat{d}$ [84]:

$$\sqrt{\mathrm{Var}(\hat{d})} \geq \frac{c}{2\sqrt{2}\pi\sqrt{\mathrm{SNR}}\beta'} \tag{1.1}$$

where c is the speed of light, SNR the signal-to-noise ratio, and β the effective signal bandwidth. As can be seen, the accuracy of the time-based localization technique can be improved by increasing either the effective bandwidth or the SNR. For this reason, the large bandwidth of UWB systems allows extremely accurate location estimations, for example, within 1 in. at $\mathrm{SNR} = 0$ dB and with a pulse of 1.5 GHz bandwidth. Excellent comprehensive surveys of the UWB transmission technique, and of the localization techniques for UWB systems, are provided in Refs 86 and 84, respectively.

1.7.1.2 Standards Based on UWB

The IEEE 802.15.3a task group has been discussing for three years an alternate physical layer for its high data rate wireless personal area networks (WPAN) standard. However, in early 2005, the group has been disbanded after not being able to reach a consensus on a single UWB-based standard between two competing proposals from two leading industry groups—the UWB Forum and the WiMedia Alliance. The UWB Forum proposal was based on a direct sequence (DS)-UWB technology, whereas the WiMedia alliance was proposing a multiband orthogonal frequency division multiplexing (MB-OFDM). The IEEE 802.15.4a task group is developing an alternate physical layer for low data rate, very low power consumption sensors, based on impulse radio UWB.

1.7.1.3 Open Research Issues

- Although the UWB transmission technology is advancing rapidly, many challenges need to be solved to enable multihop networks of UWB devices. In particular, although some recent efforts have been undertaken in this direction [83,87], how to efficiently share the medium in UWB multihop networks is still an open issue.
- As a step ahead, research is needed aimed at designing a cross-layer communication architecture based on UWB with the objective of

reliably and flexibly delivering QoS to heterogeneous applications in WMSNs, by carefully leveraging and controlling interactions among layers according to the application's requirements.

- It is necessary to determine how to provide provable latency and throughput bounds to multimedia flows in an UWB environment.
- It is needed to develop analytical models to quantitatively compare different variants of UWB to determine trade-offs in their applicability to high data rate and low power consumption devices such as multimedia sensors.
- A promising research direction may also be to integrate UWB with advanced cognitive radio [88] techniques to increase the spectrum utilization. For example, UWB pulses could be adaptively shaped to occupy portions of the spectrum that are subject to lower interference.

1.8 Conclusions

We discussed the state-of-the-art of research on WMSNs and outlined the main research challenges. Algorithms and protocols for the development of WMSNs were surveyed, and open research issues were discussed in detail. We discussed existing solutions and open research issues at the application, transport, network, link, and physical layers of the communication stack, along with possible cross-layer synergies and optimizations. We pointed out how recent work undertaken in Wyner–Ziv coding at the application layer, specialized spatio–temporal transport-layer solutions, delay-bounded routing, multi-channel MAC protocols, and UWB technology, among others, seem to be the most promising research directions in developing practical WMSNs. We believe that this research area will attract the attention of many researchers and that it will push one step further our ability to observe the physical environment and interact with it.

Acknowledgment

The authors would like to thank Prof. Ian F. Akyildiz for his valuable comments and continuous guidance and encouragement.

References

1. Akyildiz, I. F., W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. *Computer Networks (Elsevier)*, 38(4), 393–422, 2002.
2. Gurses, E. and O. B. Akan. Multimedia communication in wireless sensor networks. *Annals of Telecommunications*, 60(7–8), 799–827, 2005.