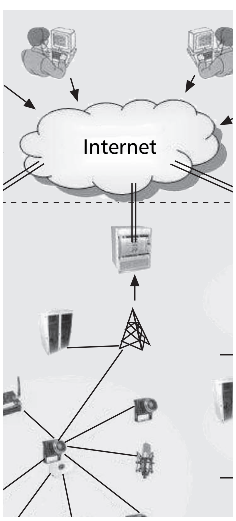
WIRELESS MULTIMEDIA SENSOR NETWORKS: A SURVEY

IAN F. AKYILDIZ, GEORGIA INSTITUTE OF TECHNOLOGY
TOMMASO MELODIA, STATE UNIVERSITY OF NEW YORK AT BUFFALO
KAUSHIK R. CHOWDURY, GEORGIA INSTITUTE OF TECHNOLOGY



The uthors discusse the state of the art and the major research challenges in architectures, algorithms, and protocols for wireless multimedia sensor networks.

ABSTRACT

In recent years, the growing interest in the wireless sensor network (WSN) has resulted in thousands of peer-reviewed publications. Most of this research is concerned with scalar sensor networks that measure physical phenomena, such as temperature, pressure, humidity, or location of objects that can be conveyed through low-bandwidth and delay-tolerant data streams. Recently, the focus is shifting toward research aimed at revisiting the sensor network paradigm to enable delivery of multimedia content, such as audio and video streams and still images, as well as scalar data. This effort will result in distributed, networked systems, referred to in this paper as wireless multimedia sensor networks (WMSNs). This article discusses the state of the art and the major research challenges in architectures, algorithms, and protocols for wireless multimedia sensor networks. Existing solutions at the physical, link, network, transport, and application layers of the communication protocol stack are investigated. Finally, fundamental open research issues are discussed, and future research trends in this area are outlined.

INTRODUCTION

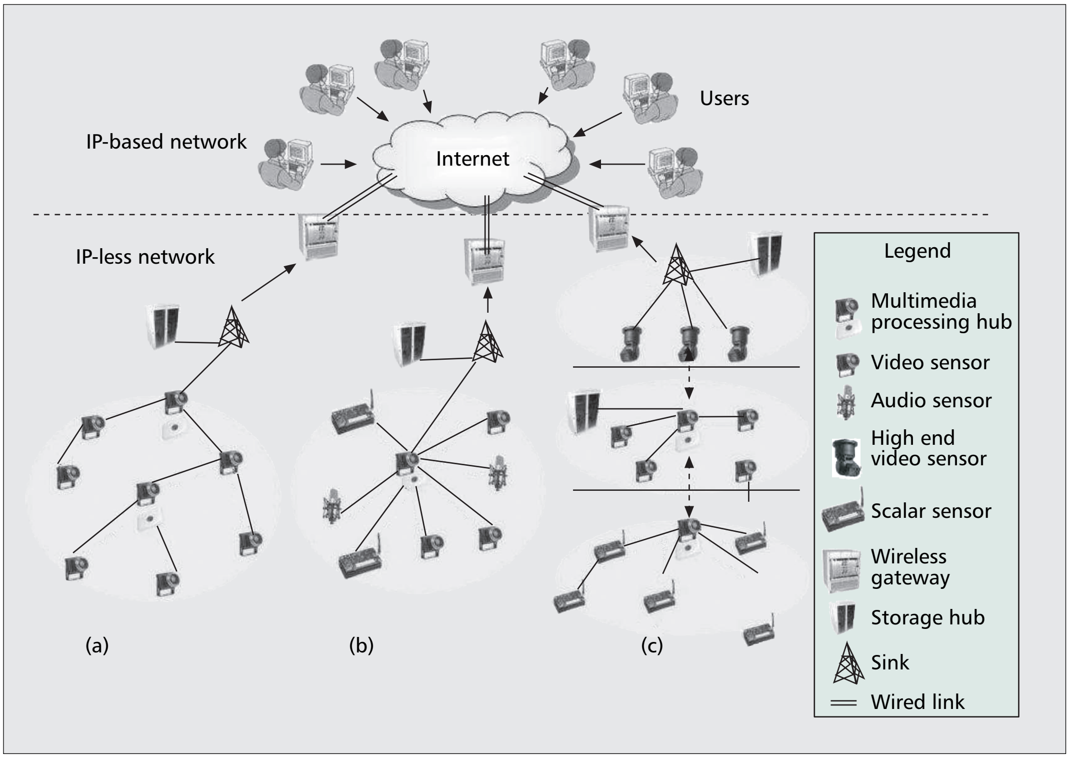
In recent years, the growing interest in the wireless sensor network (WSN) has resulted in thousands of peer-reviewed publications. Significant results in this area have enabled many civilian and military applications, and several start-up companies and large corporations are investing considerable amounts of resources in this technology. Most deployed wireless sensor networks measure scalar physical phenomena, such as temperature, pressure, humidity, or location of objects. In general, sensor networks are designed for data-only delay-tolerant applications with low bandwidth demands.

The integration of low-power wireless networking technologies with inexpensive hardware such as complementary metal–oxide semiconductor (CMOS) cameras and microphones is now enabling the development of distributed, networked systems that we refer to as wireless multimedia sensor networks (WMSNs), that is, networks of wireless, interconnected smart devices that enable retrieving video and audio streams, still images, and scalar sensor data. As an example, the Cyclops image-capturing and inference module [1], designed for extremely lightweight imaging, can be interfaced with a host mote such as Crossbow's MICA2 or MICAz, thus realizing an imaging device with processing and transmission capabilities. WMSNs will enable the retrieval of multimedia streams and will store, process in real-time, correlate, and fuse multimedia content captured by heterogeneous sources. We envision that users will be able to gather information about the physical environment by issuing simple textual queries, thus accessing multiple remote WMSNs connected to the Internet through application level gateways.

The characteristics of a WMSN diverge consistently from traditional network paradigms, such as the Internet and even from scalar sensor networks. Most potential applications of a WMSN require the sensor network paradigm to be rethought to provide mechanisms to deliver multimedia content with a predetermined level of quality of service (QoS). Whereas minimizing energy consumption has been the main objective in sensor network research, 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. Delivery of multimedia content in sensor networks presents new, specific system design challenges, which are the object of this article.

We discuss the state of the art and the main research challenges for the development of WMSNs. We begin the discussion by describing the main applications enabled by WMSNs and by introducing a reference architecture. Then, we point out the major factors influencing the design of WMSNs. Next, we discuss existing solutions and open research issues at the application, transport, network, link, and physical layers of the communication stack, respectively. Finally, we conclude the article.

1536-1284/07/$20.00 © 2007 IEEE



■ **Figure 1.** *Reference architecture of a wireless multimedia sensor network: 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 automated manufacturing processes, the integration of machine vision systems with WMSNs can simplify and add flexibility to systems for visual inspections and automated actions.

APPLICATIONS OF WIRELESS MULTIMEDIA SENSOR NETWORKS

Wireless multimedia sensor networks have the potential to enable many new applications. These can be classified as follows:

Multimedia Surveillance Sensor Networks. Surveillance sensor networks will be used to enhance and complement existing surveillance systems to prevent crime and terrorist attacks. Multimedia content, such as video streams and still images, as well as computer vision techniques, can be used to locate missing persons, identify criminals or terrorists, or infer and record other potentially relevant activities (thefts, car accidents, traffic violations).

Traffic Avoidance, Enforcement, and Control Systems. It will be possible to monitor car traffic in big cities or on highways and deploy services that offer traffic routing advice to avoid congestion or identify violations. In addition, smart-parking advice systems based on WMSNs will detect available parking spaces and provide drivers with automated parking advice.

Advanced Health Care Delivery. Telemedicine sensor networks can be integrated with third and fourth generation (3G/4G) cellular networks to provide ubiquitous health care services. Patients will carry medical sensors to monitor parameters such as body temperature, blood pressure, pulse oximetry, ECG, and breathing activity. Remote medical centers will monitor the condition of their patients to infer emergency situations.

Environmental and Structural Monitoring. Arrays of video sensors already are used by oceanographers to determine the evolution of sandbars using image processing techniques. Video and imaging sensors also are used to monitor the structural health of bridges or other civil structures.

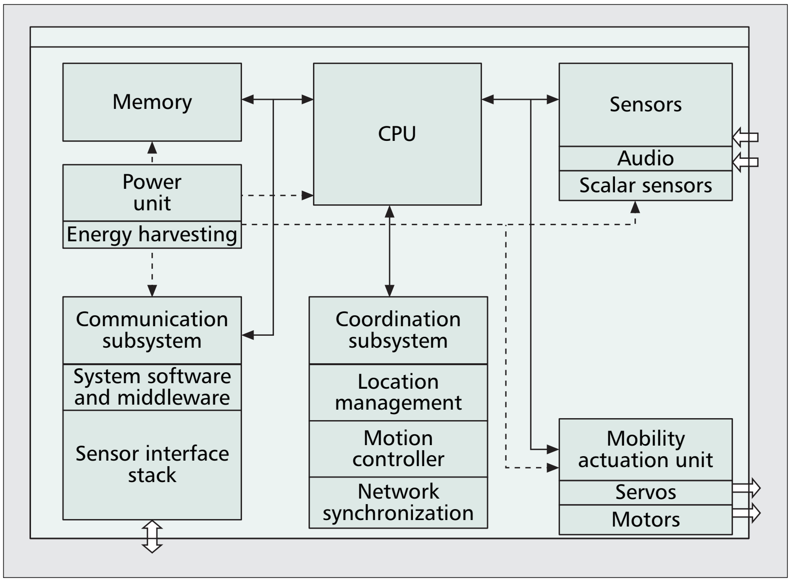
Industrial Process Control. Multimedia content such as imaging, temperature, or pressure, can be used for time-critical, industrial, process control. In automated manufacturing processes, the integration of machine vision systems with WMSNs can simplify and add flexibility to systems for visual inspections and automated actions.

NETWORK ARCHITECTURE

We introduce a reference architecture for WMSNs in Fig. 1, where users connect through the Internet and issue queries to a deployed sensor network. The functionality of the various network components are summarized in a bottom-up manner in the following list:

Standard Video and Audio Sensors. These sensors capture sound, still, or moving images of the sensed event and are typically of low resolution (in terms of pixel/inch for the video sensors and in dB for the audio sensors). They can be arranged in a single-tier network, as shown in the first cloud (Fig. 1), or in a hierarchical manner, as shown in the third cloud.

Scalar Sensors. These sensors sense scalar data and physical attributes, such as temperature, pressure, and humidity and report measured values to their clusterhead. They are typically resource-constrained devices in terms of energy supply, storage capacity, and processing capability.



■ **Figure 2.** *Internal organization of a multimedia sensor.*

Multimedia Processing Hubs. These devices have comparatively large computational resources and are suitable for aggregating multimedia streams from the individual sensor nodes. They are integral to reducing both the dimensionality and the volume of data conveyed to the sink and storage devices.

Storage Hubs. Depending upon the application, the multimedia stream is desired in real time or after further processing. These storage hubs allow data-mining and feature-extraction algorithms to identify the important characteristics of the event, even before the data is sent to the end user.

Sink. The sink is responsible for packaging high level user queries to network specific directives and returning filtered portions of the multimedia stream back to the user. Multiple sinks may be required in a large or heterogeneous network.

Gateway. This serves as the last mile connectivity by bridging the sink to the Internet and is also the only IP-addressable component of the WMSN. It maintains a geographical estimate of the area covered under its sensing framework to allocate tasks to the appropriate sinks that forward sensed data through it.

Users. Users are the highest end of the hierarchy and issue monitoring tasks to the WMSN based on geographical regions of interest. They are typically identified through their IP addresses and run application-level software that assigns queries and displays results obtained from the WMSN.

Factors Influencing the Design of Multimedia Sensor Networks

A multimedia sensor device may be composed of several basic components, as shown in Fig. 2: a sensing unit, a processing unit (CPU), a communication subsystem, a coordination subsystem, a storage unit (memory), and an optional mobility/actuation unit. Sensing units usually are composed of two subunits: sensors (cameras, microphones, and/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 executes the system software in charge of coordinating sensing and communication tasks and is interfaced with a storage unit. A communication subsystem interfaces the device 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, virtual machines, and so on. A coordination subsystem is in charge of coordinating the operation of different network devices, by performing operations such as network synchronization and location management. An optional mobility/actuation unit can enable movement or manipulation of objects. Finally, the whole system is powered by a power unit that may be supported by an energy scavenging unit, such as solar cells.

The following are several factors that influence the design of a WMSN:

Resource Constraints. Sensor devices are constrained in terms of battery, memory, processing capability, and achievable data rate.

Variable Channel Capacity. In multihop wireless networks, the capacity of each wireless link depends on the interference level perceived at the receiver. This, in turn, depends on the interaction of several functions that are distributively handled by all network devices such as power control, routing, and rate policies. Hence, the capacity and the delay attainable on each link are location dependent, vary continuously, and may be bursty in nature, thus making QoS provisioning a challenging task.

Cross-layer Coupling of Functionality. Because of the shared nature of the wireless communication channel, in multihop wireless networks, there is a strict interdependence among functions handled at all layers of the communication stack. This interdependence must be explicitly considered when designing communication protocols aimed at QoS provisioning.

Application-specific QoS Requirements. In addition to data delivery modes that are typical of scalar sensor networks, multimedia data include *snapshot* and *streaming multimedia* content. Snapshot-type multimedia data contain event-triggered observations obtained in a short time period (e.g., a still image). Streaming multimedia content is generated over longer time periods and requires sustained information delivery.

High Bandwidth Demand. Multimedia contents, especially video streams, require transmission bandwidth that is orders of magnitude higher than that supported by current off-the-shelf sensors. Hence, high data rate and low-power, consumption-transmission techniques must be leveraged. In this respect, the ultra-wide-band (UWB) transmission technology seems particularly promising for WMSNs, as discussed later.

Multimedia Source Coding Techniques.

State-of-the-art video encoders rely on *intra-frame* compression techniques to reduce redundancy within one frame and on *inter-frame* compression (also *predictive encoding* or *motion estimation*), to exploit redundancy among subsequent frames to reduce the amount of data to be transmitted and stored. Because predictive encoding requires complex encoders, powerful processing algorithms, and also entails high energy consumption, it may not be suited for low-cost multimedia sensors. However, it recently was shown in [2] that the traditional balance of complex encoder and simple decoder can be reversed within the framework of so-called *distributed source coding*. These techniques exploit the source statistics at the decoder and by shifting the complexity at this end, enable the design of simple encoders. Clearly, such algorithms are very promising for WMSNs, 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 been approached mainly as a problem isolated from the network-design problem, with a few exceptions, such as joint source-channel coding [3] and channel-adaptive streaming [4]. Similarly, 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, processing and delivery of multimedia content are not independent, and their interaction has a major impact on the achievable QoS. The QoS required by the application will be provided by means of a combination of *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.

In the following sections, the research challenges at different layers of the communication protocol stack are outlined in detail. These are summarized in Fig. 3.

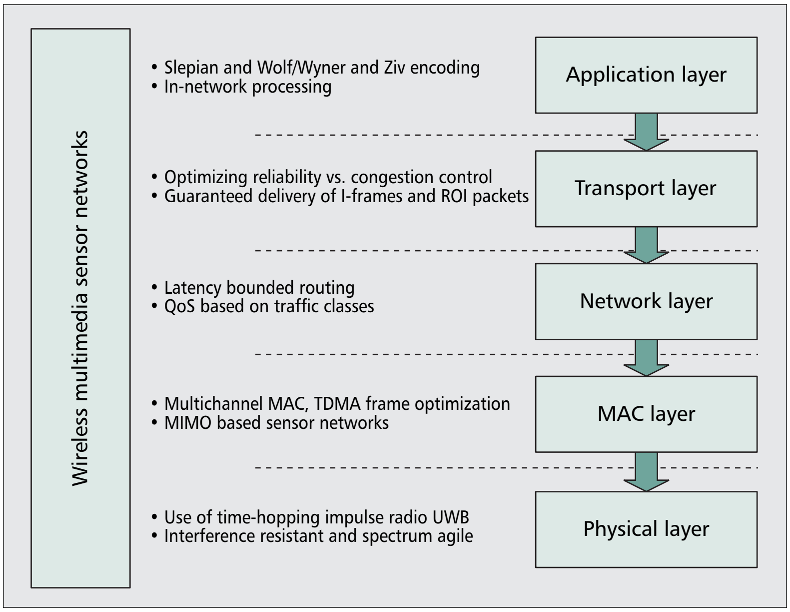


Figure 3. *Research challenges at different layers of the protocol stack.*

Physical Layer

Among other promising technologies, the UWB technology [5] has the potential to enable low power consumption, high, data-rate communication within tens of meters. There exist several variants of UWB. Time-hopping impulse radio UWB (TH-IR-UWB) is based on sending pulses of very short duration (on 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 pseudo-random time hopping sequences (THS) that determine in which chip each user should transmit. Simple TH-IR-UWB systems can be very inexpensive to construct. TH-IR-UWB is particularly appealing for WMSNs for several reasons. First, TH-IR-UWB enables high data rate, very low-power, carrierless communication on simple-design, low-cost radios. Moreover, it provides a large processing gain in the presence of interference, and it is flexible, because data rate can be traded for power spectral density and multipath performance. Importantly, the impulse radio technology naturally allows for integrated medium access control/physical (MAC/PHY) layer solutions, because interference mitigation techniques allow realizing MAC protocols that do not require mutual, temporal exclusion between different transmitters [6]. Hence, simultaneous communications of neighboring devices are feasible without complex receivers. Furthermore, the large instantaneous bandwidth enables fine-time resolution for accurate position estimation and for network synchronization. Finally, UWB signals have extremely low-power spectral density, with low probability of intercept/detection (LPI/D), which is particularly appealing for covert military operations.

Although the UWB transmission technology is advancing rapidly, many challenges must be solved to enable multihop networks of UWB devices. Although some recent efforts have been undertaken in this direction [6], the way to efficiently share the medium in UWB multihop networks is still an open issue. Research is required aimed at designing a cross-layer communication architecture based on UWB to support QoS in WMSNs and at guaranteeing provable latency and throughput bounds to multimedia flows in a UWB environment.

MAC Layer

The two main functions of the MAC layer are arbitration of the channel and providing error control and recovery schemes. There are several approaches for regulating the channel access based on contention, and we advocate the use of contention-free protocols for WMSNs. We also delve into the factors influencing the choice of

Class type	Data type	Bandwidth	Description
Real-time, loss-tolerant	Multimedia	High	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 the user in real time
Delay-tolerant, loss-tolerant	Multimedia	High	Streams intended for storage or subsequent offline processing that need to be delivered quickly due to the limited buffers of multimedia sensors
Real-time, loss-tolerant	Data	Moderate	Monitoring data from densely deployed scalar sensors characterized by spatial correlation or loss-tolerant snapshot multimedia data (e.g., images of a phenomenon taken from multiple viewpoints at the same time)
Real-time, loss-tolerant	Data	Moderate	Data from time-critical monitoring processes such as distributed control applications
Delay-tolerant, loss-intolerant	Data	Moderate	Data from monitoring processes that require some form of offline post processing
Delay-tolerant, loss-tolerant	Data	Low	Environmental data from scalar sensor networks or non-time-critical snapshot multimedia content

■ **Table 1.** *Traffic classes.*

forward error correction (FEC) schemes as against automatic repeat request (ARQ) in this section.

Channel Access Policies

Based on the nature of channel access, some MAC protocols are geared to provide high link-level throughput, reduce delays, or guarantee QoS for a given packet type. The main categories of these protocols are listed in the following, and their key features that can be useful for WMSNs are discussed.

Contention-Based Protocols — Existing schemes are based mostly on variants of the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) MAC protocol. When a device is receiving data, transmissions from all the devices in its transmission range are impeded. However, this is achieved by the use of random timers and a carrier sense mechanism, which in turn, results in uncontrolled delay and idle energy consumption. Moreover, WMSNs have different traffic classes, as shown in Table 1. Approaches similar to the hybrid coordination function-controlled channel access (HCCA) component present in IEEE 802.11e must be used. However, the sensor protocol stack, resident on about 4 KB of flash memory in current off-the-shelf motes, such as Crossbow's MICAz, based on the Chipcon 2420 chipset, must be simple and lightweight. Although S-MAC [7] and protocols inspired by it meet this requirement, they also introduce sleep periods to save on energy consumption but at the expense of latency and coordination complexity.

Channel contention can be significantly reduced by using multiple channels in a spatially overlapped manner. This is commonly accomplished by using two radios, in which one is delegated the task of channel monitoring. This is often a low-power transceiver and is responsible for waking up the main radio for data communication on a need basis. However, such multichannel schemes introduce the problem of distinct channel assignment and add hardware complexity.

Contention-free Single Channel Protocols — Time-division multiple access (TDMA) is a representative protocol of this class. Usually, the frame is organized with 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 pre-decided intervals to assign slots to active nodes, taking into consideration the QoS requirement of their data streams. Variable length of the TDMA frames (V-TDMA) and the frequency of the RP interval are some of the design parameters that can be exploited when designing a multimedia system. However, TDMA-based sensor networks usually present limited scalability and complex network-wide scheduling, apart from the problems of clock drift and synchronization issues.

Unlike TDMA, which aims exclusively to reserve the channel on a time basis, multiple input multiple output (MIMO) antenna systems employ interference cancellation techniques. Each sensor may function as a single antenna element, sharing information and thus simulating the operation of a multiple antenna array. A distributed MIMO-based compression scheme for correlated sensor data that especially addresses multimedia requirements, is proposed in [8]. However, with the increasing complexity associated with MIMO systems, further research is required at the MAC layer to ensure that the required parameters, such as channel state and desired diversity/processing gain are known to both the sender and the receiver at an acceptable energy cost in a WMSN.

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 and data link layer, namely forward error correction (FEC) and automatic repeat request (ARQ), along with hybrid schemes. 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. ARQ mechanisms, on the other hand, use bandwidth efficiently at the cost of additional latency involved with the re-transmission process. Recent comparisons made between ARQ and FEC reveal that for certain FEC block codes (BCH), longer routes decrease both the energy consumption and the end-to-end latency, subject to a target packet error rate compared to ARQ [9]. Thus, FEC codes are an important candidate for delay-sensitive traffic in WSNs.

Network Layer

Several design considerations of traditional WSN routing, such as energy optimization, link quality, and multipath and fault tolerance, among others also are applicable for WMSNs [10]. However, we focus our discussion on the primary network layer functionality of multimedia routing. We classify this further based on:

- Architectural and spatial attributes
- Real time support

Routing with Hierarchical Architectures and Correlation

From Fig. 1, we observe that hierarchical WMSNs can be deployed that have different types of sensors with varying capabilities. Thus, there may be two sets of routes with, for example, low-granularity image sensors forming one set, and the overlaying high resolution video sensors constituting the other. Both these routing schemes may require close interaction, as they carry packets describing the same event. Also, the feed from varying sensor types may need to be fused periodically, thus requiring common nodes along both paths. Hence, we believe that there must be a cooperative approach among routing algorithms operating at different layers of a hierarchical environment.

The effects of correlation also decide the choice of routes in a WMSN. Video sensors, in particular, may have their cameras oriented in different directions. It may be desirable that the routing path that is followed include those particular sensors that observe the same event in their restricted range and field of vision. Such a routing scheme also would facilitate in-network processing and remove redundancies in the data sent to the sink.

Real Time Routing Protocols

Meeting strict time deadlines, as required for streaming applications, and maintaining reliability constraints are often contradictory goals. Reducing the delay at each link and routing based on local channel conditions may help alleviate this problem. The Multi-Path and Multi-SPEED Routing Protocol (MMSPEED) is one such approach that attempts to balance between these two goals and spans over the network and MAC layers [11]. It differentiates between flows with different delay and reliability requirements to channel resources to flows. Although each node selects its next hop based on link-layer delay measurements, a feedback mechanism along the path helps correct local estimation inaccuracies. However, this research direction is still a best-effort practice and does not propose energy saving techniques or give firm guarantees in bursty multimedia traffic.

Transport Layer Protocols

Classical transport layer functionalities, such as providing end-to-end congestion control, become especially important in real-time delay-bounded applications, such as streaming media. We study these based on their underlying dependence on:

- User Datagram Protocol (UDP)
- TCP and TCP-compatibility

UDP Based Protocols

UDP is usually preferred over TCP in typical multimedia applications as *timeliness* is of greater concern than *reliability*. Selected features of existing standards for the Internet, such as Real-time Transport Protocol (RTP) may be adopted in context of WMSNs. RTP uses a separate control stream called Real-time Transport Control Protocol (RTCP) that allows dynamic adaptation to the network conditions. RTP runs over UDP, but provides support for a host of functions, such as bandwidth scaling and integration of different images into a single composite. In addition, the application level framing (ALF) allows manipulation of the header to suit application specific requirements. Through ALF, specific instructions can be encoded in the header that are typical for the WMSN application, while ensuring compatibility with the external IP-based data storage and monitoring network.

TCP and TCP Friendly Schemes for WMSNs

Typically, packets sent to the sink are highly compressed at the source with only a subset of the nodes transmitting that have non-redundant data. Compression standards such as the JPEG2000 and the MPEG introduce features such as the region of interest (ROI) and the I-frame respectively. These special packets carry original content that cannot be retrieved through interpolation. Hence, dropping packets indiscriminately, as in the case of UDP, may cause discernible disruptions in the multimedia content. Thus, we argue that some form of *selective reliability*, such as that provided by TCP, must be introduced for these packets in a WMSN.

Two key factors that limit multimedia transport based on TCP-like rate control schemes are the jitter introduced by the congestion control mechanism and the control message overhead. Existing solutions for transporting MPEG video in a TCP-friendly manner overcome this problem of jitter by assuming playout buffers at the sink. Distributed approaches, especially addressing the concerns of sensor networks, often cache TCP segments within the network and through local retransmissions, they reduce the message-

Meeting strict time deadlines, as required for streaming applications, and maintaining reliability constraints are often contradictory goals. Reducing the delay at each link and routing based on local channel conditions may help alleviate this problem.

Clearly, no single transport layer solution exists that addresses the diverse concerns of WMSNs. As an example, defining reliability metrics, based on the packet content, and coupling application layer coding techniques to reduce congestion may be promising directions in this area.

passing overhead. The use of TCP also can be argued from the point of bandwidth utilization. Multiple streams can be opened between source and sink, each of which may follow a different path. Thus by splitting the multimedia traffic into several smaller data-rate paths, and by dynamically changing the TCP window size for each connection, a fine-grained control on sensor traffic is possible that may extend network lifetime and enhance performance.

Being unable to distinguish between bad channel conditions and network congestion is a major problem in TCP.

Clearly, no single transport layer solution exists that addresses the diverse concerns of WMSNs. As an example, defining reliability metrics, based on the packet content, and coupling application layer coding techniques to reduce congestion may be promising directions in this area.

Application Layer

In this section, we overview challenges and functionality at the application layer with respect to the different traffic classes that may be seen in a typical WMSN application, as shown in Table 1.

Multimedia Encoding Techniques

The main design objectives of a coder for WMSNs are:

- *High compression efficiency*. It is mandatory 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 must 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.

The traditional broadcasting paradigm, 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 ISO MPEG schemes or the International Telecommunication Union-Telecommunication (ITU-T) recommendations H.263 and H.264 (also known as AVC or MPEG-4 part 10), reduce the bit rate generated by the source encoder by exploiting source statistics. Because the computational complexity is 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 a 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 [2]. Conversely, to realize low-cost, low-energy-consumption multimedia sensors, it is mandatory to develop simpler encoders but still retain the advantages of high compression efficiency.

Fortunately, it is known from information-theoretic bounds established by Slepian and Wolf for lossless coding and by Wyner and Ziv 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 [2]. 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 each other [12]. 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 low-power, and the decoder at the sink will be complex and loaded with most of the processing and energy burden. The reader is referred to [12] and [4] for excellent surveys on the state of the art of distributed source coding in sensor networks and in distributed video coding, 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 multi-layer coding with wavelet compression — are discussed in [13].

Collaborative In-network Processing

Given a source of data (e.g., a video stream), different applications may require diverse information (e.g., raw video stream vs. 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 distributed filtering and in-network processing architectures, to enable real-time retrieval of useful information.

Similarly, it is necessary to develop architectures to perform data fusion or other complex processing operations *in-network*. Algorithms for both inter-media and intra-media data aggregation and fusion must be developed, because simple distributed processing schemes developed for existing scalar sensors are not designed for multimedia contents.

Conclusions

We discussed the state of the art of research on WMSNs and outlined the main research challenges. We discussed existing solutions and open research issues at the physical, link, network, transport, and application layers of the communication stack. In particular, we believe that recent work undertaken in Wyner-Ziv coding at the application layer, the leveraging of spatial-temporal aspects of multimedia sensing in designing routing and transport layer solutions, MAC protocols that provide link latency bounds, and UWB technology, among others, seem to be the most promising research directions in developing practical WMSNs.

Acknowledgments

This material is based on work supported by the U.S. National Science Foundation under contract no. ECCS-0701559.

References

[1] M. Rahimi *et al.*, "Cyclops: In Situ Image Sensing and Interpretation in Wireless Sensor Networks," *Proc. ACM Conf. Embedded Networked Sensor Sys.*, San Diego, CA, Nov. 2005.

[2] B. Girod *et al.*," Distributed Video Coding," *Proc. IEEE*, vol. 93, no. 1, Jan. 2005, pp. 71–83.

[3] Y. Eisenberg *et al.*," Joint Source Coding and Transmission Power Management for Energy Efficient Wireless Video Communications," *IEEE Trans. Circuits and Sys. for Video Tech.*, vol. 12, no. 6, June 2002, pp. 411–24.

[4] B. Girod *et al.*," Advances in Channel-Adaptive Video Streaming," *Wireless Commun. and Mobile Comp.*, vol. 2, no. 6, Sept. 2002, pp. 549–52.

[5] L. Yang and G. B. Giannakis, "Ultra-WideBand Communications: An Idea Whose Time Has Come," *IEEE Sig. Processing*, vol. 21, no. 6, Nov. 2004, pp. 26–54.

[6] R. Merz *et al.*, "A Joint PHY/MAC Architecture for Low-Radiated Power TH-UWB Wireless Ad-Hoc Networks," *Wireless Commun. and Mobile Comp. J.*, vol. 5, no. 5, July 2005, pp. 567–80.

[7] W. Ye, J. Heidemann, and D. Estrin, "Medium Access Control with Coordinated, Adaptive Sleeping for Wireless Sensor Networks," *IEEE Trans. Net.*, vol. 12, no. 3, June 2004, pp. 493–506.

[8] S. K. Jayaweera and M. L. Chebolu," Virtual MIMO and Distributed Signal Processing for Sensor Networks — An Integrated Approach," *Proc. IEEE ICC*, Seoul, Korea, May 2005.

[9] M. C. Vuran and I. F. Akyildiz, "Cross-Layer Analysis of Error Control in Wireless Sensor Networks," *Proc. IEEE Intl. Conf. on Sensor and Ad Hoc Commun. and Networks*, Reston, VA, Sept. 2006.

[10] S. Vural, Y. Tian, and E. Ekici," QoS-Based Communication Protocols in Sensor Networks," A. Boukerche, ed., *Algorithms and Protocols for Wireless Ad Hoc and Sensor Networks*, Wiley, 2006.

[11] E. Felemban, C.-G. Lee, and E. Ekici," MMSPEED: Multipath Multi-SPEED Protocol for QoS Guarantee of Reliability and Timeliness in Wireless Sensor Networks," *IEEE Trans. Mobile Comp.*, vol. 5, no. 6, June 2006, pp. 738–54.

[12] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed Source Coding for Sensor Networks," *IEEE Sig. Processing*, vol. 21, no.5, Sept. 2004, pp. 80–94.

[13] S. Misra, M. Reisslein, and G. Xue," A Survey of Multimedia Streaming in Wireless Sensor Networks," 2006.

Biographies

IAN F. AKYILDIZ (ian@ece.gatech.edu) received his B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nürnberg, Germany, in 1978, 1981, and 1984, respectively. Currently, he is the Ken Byers Distinguished Chair Professor with the School of Electrical and Computer Engineering (ECE), Georgia Institute of Technology, Atlanta, and director of the Broadband Wireless Networking Laboratory. He is an Editor-in-Chief of *Computer Networks Journal* (Elsevier), as well as the founding Editor-in-Chief of *Ad Hoc Network Journal* (Elsevier). His current research interests are in next-generation wireless networks, sensor networks, and wireless mesh networks. He received the Don Federico Santa Maria Medal for his services to the Universidad of Federico Santa Maria in 1986. From 1989 to 1998 he served as a national lecturer for the Association for Computing Machinery (ACM) and received the ACM Outstanding Distinguished Lecturer Award in 1994. He received the 1997 IEEE Leonard G. Abraham Prize Award (IEEE Communications Society) for his paper entitled "Multimedia Group Synchronization Protocols for Integrated Services Architectures" published in *IEEE Journal of Selected Areas in Communications* in January 1996. He received the 2002 IEEE Harry M. Goode Memorial Award (IEEE Computer Society) with the citation "for significant and pioneering contributions to advanced architectures and protocols for wireless and satellite networking." He received the 2003 IEEE Best Tutorial Award (IEEE Communication Society) for his paper entitled "A Survey on Sensor Networks," published in *IEEE Communications Magazine* in August 2002. He also received the 2003 ACM Sigmobile Outstanding Contribution Award with the citation "for pioneering contributions in the area of mobility and resource management for wireless communication networks." He received the 2004 Georgia Tech Faculty Research Author Award for his "outstanding record of publications of papers between 1999–2003." He also received the 2005 Distinguished Faculty Achievement Award from the School of ECE, Georgia Tech. He has been a Fellow of the ACM since 1996.

TOMMASO MELODIA (tmelodia@eng.buffalo.edu) received his Laurea in telecommunications engineering and his doctorate in information and communication engineering from the University of Rome La Sapienza, Italy, in 2001 and 2005, respectively. He also received his Ph.D. in electrical and computer engineering after working as a research assistant at the Broadband Wireless Networking Laboratory (BWN-Lab), Georgia Institute of Technology in 2007. He joined the Department of Electrical Engineering at the State University of New York at Buffalo in 2007. His main research interests are in wireless sensor networks, underwater acoustic sensor networks, and cognitive mesh networks. He is the recipient of the BWN-Lab Researcher of the Year Award for 2004, and is the author of about 40 publications in leading conferences and journals on wireless networking.

KAUSHIK R. CHOWDHURY [S] (kaushikc@ece.gatech.edu) received his B.E. degree in electronics engineering with distinction from VJTI, Mumbai University, India, in 2003. He received his M.S. degree in computer science from the University of Cincinnati in 2006, graduating with the best thesis award. He is currently a research assistant in the Broadband Wireless Networking Laboratory and pursuing his Ph.D. degree at the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta. His current research interests include multichannel medium access protocols, dynamic spectrum management, and resource allocation in wireless multimedia sensor networks.

Algorithms for both inter-media and intra-media data aggregation and fusion must be developed, because simple distributed processing schemes developed for existing scalar sensors are not designed for multimedia contents.