MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
IEEE COMMUNICATIONS SOCIETY
http://www.comsoc.org/~mmc

E-LETTER

Vol. 5, No. 3, May 2010

IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from MMTC Chair

Dear MMTC members,

It is quite excited that we will have another MMC meeting coming soon in ICC 2010 at Cape Town, South Africa from May 23-28. I am looking forward to seeing all of you there for our MMC meeting, which has a draft agenda as follows.

0. Informal discussion and networking time
1. welcome new members /introduction
2. Last meeting minutes approval (Globecom 2009)
3. Report on Conferences activities CCNC 2010, ICC 2010, ICME 2010, Globecom 2010, ICC 2011, Globecom 2011
4. TAC Report
5. MMTC IGs Reports - all IG chairs
6. Sub-committees Report
7. Report for Newsletter activity
8. Report on recent status related to ICME conference and IEEE Transactions on Multimedia journal
9. Election of new officers of MMC
10. Suggestions & discussions – everyone
11. Adjourn

Please note that one of the very important items we will discuss in this MMC meeting is the election of new officers of MMC. A nomination committee that includes several past committee chairs has been formed. Nelson is serving as the nomination committee chair. Please actively join the election process.

As the chair of MMC during 2008-2010, I would like to take this opportunity on behalf of all the current officers to express our gratitude for all your strong supports and collaboration during the past 2 years.

Again, looking forward to seeing you in Cape Town, South Africa soon.

Cheers,
Qian Zhang
IEEE ComSoc MMTC Chair

Message from E-Letter Director

As reported, Ericsson's President and CEO Hans Vestberg, recently at the Annual General Meeting of Shareholders, envisioned 50 billion ($\approx 10*2^{32}$, decuple IPv4 addresses) Internet connected devices by 2020, the true Internet of Things (IoT); he also predicted growing opportunities in media industry. In fact, multimedia is becoming a more critical and attractive application in mobile and embedded devices.

This May issue of IEEE ComSoc MMTC bi-monthly E-Letter features a special topic and technology advances for embedded and mobile multimedia.

First, Dr. Mischa Dohler (CTTC, Spain) brings together a wonder special issue on multimedia over embedded systems, which contains six invited papers with diverse topics dealing with multimedia over wireless multimedia sensor networks (such as visual and video sensor networks) including: self-organized control in visual sensor networks of wireless cameras, cross-layer design and optimization in video sensor networks, leveraging distributed in-network processing to improve multimedia delivery in wireless sensor networks, secure multimedia in sensor networks, video transmission in 6LoWPAN/IEEE 802.15.4 multi-hop sensor networks, and research challenges and open issues for multimedia over embedded systems.

The technology advance column includes four invited papers and one published paper from editor's recommendation:

- "*Networking Coding – Enabling the Multimedia Wireless Internet*" from Marie-Jose Montprtit and Muriel Medard,
- "*Video Coding Solution for VANETs*" written by Mohammed Ghanbari, Martin Fleury and Nadia N. Qadri,
- "*Modeling and Resource Allocation for HD Videos over WiMAX Broadband Wireless Networks*" co-authored by Abdel Karim Al Tamimi, Chakchai So-In and Raj Jain
- "*Exploiting Channel Fading and SVC in Wireless Video Streaming* contributed by Honghai Zhang, Mohammad A. Khojastepour, Ravi Kokku, Rajesh Mahindra and Sampath Rangarajan.
- In addition, a recently published paper titled with "*Robust Video Transmission with Distributed Source Coded Auxiliary Channel*" by J. Wang, A. Majumdar, and K. Ramchandran is recommended by Dr. Guan-Ming Su.

I would like to thank all editors and authors for their great efforts and contributions to this issue. I hope you find some pieces interesting.

Chonggang Wang, Ph.D

Director of IEEE ComSoc MMTC E-Letter
InterDigital Communications, USA
cgwang@ieee.org

SPECIAL ISSUE ON EMBEDDED SYSTEMS

Multimedia Over Embedded Systems

Guest Editor: Mischa Dohler, CTTC, Spain
mischa.dohler@cttc.es

Wireless Sensor Networks (WSNs), primarily realized by means of embedded systems, have largely been designed and investigated in the past decade to carry small amounts of fairly delay-tolerant sensed data, albeit at the highest possible energy efficiency. This paradigm has driven various communities implicated in the design of said systems, including the signal processing, telecommunications and computer science communities. Algorithms and protocols were designed to facilitate the transmission of low-rate delay-tolerant data bundles at unprecedented energy efficiencies without (significantly) jeopardizing operational reliability.

This is until I.F. Akyldiz, *et al.*, came along and published a true visionary – and that time unthinkable – milestone paper (“Wireless Sensor Networks: A Survey,· in 2002) in which the use of embedded wireless sensor networks for multimedia applications has already been hinted at. Having guided our entire community, numerous contributions have emerged ever since. A quick search in IEEE Xplore reveals that almost 100 papers have been published in the last years on this subject.

A special issue dedicated this topic has hence been overdue. I have thus aimed at collecting some position and vision papers from leading scientists in the field, who stem from different communities and have some very differing backgrounds, and who deal or have been dealing with the topic at hand. This, so I hope, shall serve as a guiding hand for the multimedia and signal processing communities over the upcoming years. I also hope that this not only serves the academic community but also an industrial community so that embedded multimedia systems and applications are ubiquitous this time 5 years from now.

I have assembled these six papers in the following order. First, “Self-Organized Control for Visual Sensor Networks” by Naoki Wakamiya and Masayuki Murata advocates a self-organizing paradigm to improve multimedia transmission over embedded systems. Second, “Cross-layer Optimization in Video Sensor Networks” by Tommaso Melodia and Wendi Heinzelman outline some pending challenges in cross-layer designs. Third, “Harnessing Collective Power of Sensor Nodes as Distributed Embedded Systems” by Eylem Ekici shows that distributed processing is advantageous from a performance point of view. Fourth, “Open Issues in Secure Wireless Multimedia Sensor Networks” by Alfredo Grieco, Sabrina Sicari, and Gennaro Boggia focuses on security and trust issues which are important to numerous multimedia applications. Fifth, “Video Transmission Over A Standards-Based Wireless Multi-Hop Sensor Network” by Thomas Watteyne, Fabien Chraim, Nahir Sarmicanic, Chris Jian and Kristofer S. J. Pister from Berkeley shows that and how latest IEEE/IETF protocols facilitate transmission of multimedia information over embedded systems. Finally, “Current And Future Open Research Issues On Multimedia Over Embedded Systems” by Ilias Politis and Tasos Dagiuklas outlines some future research issues.

I hope you will enjoy reading these contributions of this special issue dedicated to multimedia over embedded systems.

Mischa Dohler
CTTC, Barcelona
May 2010

Mischa Dohler [www.cttc.es/home/mdohler] is now Senior Researcher with CTTC in Barcelona, working on wireless sensor, machine-to-machine, femto, cooperative, cognitive and docitive networks. He is also co-founder of Worldsensing [www.worldsensing.com] which provides real-time, mission-critical sensing and actuation applications; multimedia for security being one of such applications.

Prior to this, from June 2005 to February 2008, he has been Senior Research Expert in the R&D division of France Telecom. From September 2003 to June 2005, he has been lecturer at King's College London, Centre for Telecommunications Research. At that time, he has also been London Technology Network Business Fellow for King's College London, as well as Student Representative of the IEEE UKRI Section and member of the Student Activity Committee of IEEE Region 8 (Europe, Africa, Middle-East and Russia).

He obtained his PhD in Telecommunications from King's College London, UK, in 2003, his Diploma in Electrical Engineering from Dresden University of Technology, Germany, in 2000, and his MSc degree in Telecommunications from King's College London, UK, in 1999. Prior to Telecommunications, he studied Physics in Moscow. He has won various competitions in Mathematics and Physics, and participated in the 3rd round of the International Physics Olympics for Germany.

In the framework of the Mobile VCE, he has pioneered research on distributed cooperative space-time encoded communication systems, dating back to December 1999. He has published more than 120 technical journal and conference papers at a citation h-index of 21 and citation g-index of 43, holds several patents, co-edited and contributed to several books, has given numerous international short-courses, and participated in standardization activities. He has been TPC member and co-chair of various conferences, such as technical chair of IEEE PIMRC 2008 held in Cannes, France. He is and has been editor for numerous IEEE and non-IEEE journals and is Senior Member of the IEEE.

Self-Organized Control for Visual Sensor Networks

Naoki Wakamiya and Masayuki Murata, Osaka University, Japan
{wakamiya, murata}@ist.osaka-u.ac.jp

1. Introduction

A visual sensor network consisting of embedded visual sensor nodes, e.g. wireless cameras, is promising as a monitoring and surveillance system for the high visibility and the amount of information that video images provide. However, it suffers from limitations on the wireless communication capacity and the computational capacity [1]. For the limited bandwidth, an attempt to collect high-quality video data from all camera nodes to a monitoring center always fails and it results in corruption of perceived video images. Although mechanisms such as bandwidth allocation, retransmission control, and FEC (Forward Error Correction) have been proposed to maintain the quality of video images in a lightly loaded and moderately congested wireless network, they do not help much when the volume of video traffic is, for example, twice as much as the capacity of the network. Therefore, we need application-level control to regulate the amount of video traffic in addition to network-level congestion control. Furthermore, the mechanisms should be simple and light enough so that it can be easily implemented on powerless embedded nodes.

From a viewpoint of applications, not all video images are equally important. It implies that we can consider such a mechanism where the quality of video images is high at a camera node detecting a target, e.g. a suspicious person, while keeping the quality low at the other nodes to control the amount of video traffic not to exceed the network capacity [2,3]. At the same time, network-level congestion control is required to mitigate buffer overflow and packet loss caused by the heterogeneous and unbalanced video traffic. Such congestion control can be accomplished by allowing a node who has more packets in a buffer to have the higher transmission rate than other nodes [4]. Although it might be possible that a single central node determines and dictates the video quality and the transmission rate to all camera nodes, it involves the considerable control overhead in collecting the up-to-date information about the location of targets and the buffer occupancy and to disseminate the control message to all nodes. As such it works only in a small-scale network.

In this paper, we propose mechanisms of video quality control and congestion control which are scalable to the number of camera nodes and adaptive to the number, location, and velocity of targets. In our proposal, each node determines the video quality and the transmission rate based on local information that it obtains by exchanging messages with neighbors. As a consequence of mutual interaction among nodes, the globally organized application-level and network-level control emerges. That is, self-organization [5]. In a self-organizing system, the global pattern appears as a result of mutual interaction among simple agents behaving based on local information. We adopt a reaction-diffusion model [6] as the fundamental theory of self-organized control mechanisms on both of application and network levels in this paper. With our mechanisms, each node only need to evaluate the reaction-diffusion equations based on the information about itself and neighbors for video quality and congestion control.

2. Reaction-Diffusion Model

A reaction-diffusion model expresses chemical reactions of morphogens intra- and inter-cells. Alan Turing explained self-organization of periodic patterns on the surface of body of fishes and mammals by using the model [6]. A general form of a reaction-diffusion model of two virtual morphogens called activator and inhibitor is formulated by a pair of temporal differential equations as follows.

$$\begin{cases} \dfrac{\partial u}{\partial t} = F(u,v) + D_u \nabla^2 u \\ \dfrac{\partial v}{\partial t} = G(u,v) + D_v \nabla^2 v \end{cases}$$

where u and v are the concentrations of morphogens. The first term of the right-hand side is called a reaction term corresponding to chemical reactions within a cell formulated by functions F and G. The second term is called a diffusion term corresponding to chemical interactions between neighbor cells. D_u and D_v are the diffusion rates and ∇^2 is the Laplacian operator.

Depending on the form of reaction-diffusion equations and their parameters, a variety of patterns, such as stripes, maze, and spots, can be generated. In the reaction-diffusion model, the following two conditions must be satisfied to generate patterns. First, the activator activates itself and the inhibitor, whereas the inhibitor restrains itself and the activator. Second, the inhibitor diffuses faster than the activator ($D_v > D_u$). Now assume that the concentration of activator slightly increases in the field of homogeneous morphogen concentrations, the concentrations of activator and inhibitor are increased around the point by being activated by the activator. The generated inhibitor diffuses faster than the activator and restrains generation of activator at further areas. On the other hand, the activator stays at the point and the concentration of activator is kept higher than that of inhibitor. Consequently, the diversity in the concentration of activator emerges and a pattern of heterogeneous morphogen concentrations appears.

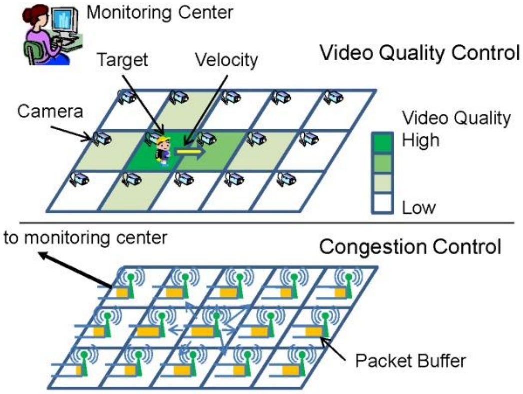


Fig. 1: Visual Sensor Network.

3. Reaction-Diffusion based Self-organized Control for Visual Sensor Network

Figure 1 illustrates our system model. Camera nodes with capability of wireless communication and motion detection are distributed in a monitored region. It is not necessary that nodes are arranged in a grid as far as a wireless network is connected where there is no isolated node, but we hereafter assume a grid layout for the ease of explanation. Each square corresponds to the area that a camera observes. A camera node can communicate with four neighbors in up, right, down, and left directions. It has a packet buffer for each of neighbors. Video images obtained by camera nodes are collected at a monitoring center, which is behind a wired connection, through multi-hop wireless communication among nodes.

Now, there is a target moving toward the right. The video quality at a camera node which has a target in its observation area must be as high as possible to have detailed video images of the target. A camera node of the neighbor area in the moving direction should provide high-quality images preparing for the future movement. Surrounding nodes also set the video quality at the moderate level, so that they can deal with the sudden and irregular movement of the target.

Consequently, we see a spatial pattern of heterogeneous distribution of video quality, i.e. a spot centered at the node detecting the target. Here, we can directly adopt the reaction-diffusion model in order to autonomously generate the spot pattern through local interaction among neighbors. Nodes maintain and calculate virtual morphogen concentrations and set the video quality in accordance with the concentrations. However we need to extend the model to have a spot spreading toward the moving direction of a target while keeping the total amount of video traffic within the wireless network capacity. In our mechanism, a node detecting a target in its observation area adds the small and constant amount of activator, called stimulus, in the reaction-diffusion equation to increase the activator concentration and generate a spot centered at the node. The stimulus propagates to neighbors and further nodes in the moving direction of a target while decreasing the amount.

As a result of the above-mentioned video quality control, there appears the concentration of video traffic on the path from the spot to the monitoring center. Then, packet losses would occur at a node on the path by exceeding the capacity of the packet buffer for the next-hop node to the monitoring center, while local buffers for the other neighbors and buffers of the other neighbors have room for more packets. Therefore, it is effective to distribute the load among local buffers and among neighbors to suppress packet loss. We can easily combine these two mechanisms by using the reaction-diffusion model. From a mechanistic viewpoint, the reaction term corresponds to local control within a node and the diffusion term realizes mutual interaction between neighbors.

We briefly explain the basic behavior of node adopting the reaction-diffusion model for video quality control and congestion control. A node monitors the buffer occupancy and the observation area by a camera. The state information is exchanged among neighbors by being embedded in a HELLO message at a regular HELLO interval. The information contains the concentrations of virtual morphogens, the amount of stimulus, and the total number of packets stored in buffers. A node evaluates two reaction-diffusion models, i.e. one for video quality control and the other for congestion control, based on the information of itself and neighbors. Equations are spatially and temporally discretized and transformed to integer arithmetic. For video quality control, the morphogen concentrations are updated and a node sets the video quality accordingly. If there is a target in the observation area, a node sets the stimulus. For congestion control, by the reaction term a node determines weights of the WRR (weighted round robin) scheduler to give more weight to a packet buffer with more packets than others and let packets leave the buffer more often. The diffusion term determines CW_{min} (the minimum of contention window) of IEEE 802.11 CSMA/CA in accordance with the relative buffer occupancy, where a node with more packets has a smaller CW_{min} and obtains more chances to find the available channel than neighbors.

4. Conclusions

In this paper, we briefly explain our idea to adopt the reaction-diffusion model to control both of video quality and congestion in a self-organized manner in visual sensor networks. Self-organization must lead to the scalability, adaptability, and robustness of the system. Although our preliminary results prove the performance of our idea, not shown for space limitation, we need to evaluate our proposal from the above aspects rather than the performance.

References

[1] Y. Charfi, N. Wakamiya, and M. Murata, “Challenging issues in visual sensor networks,” *IEEE Wireless Communications Magazine*, vol. 16, no. 2, pp. 44–49, April 2009.

[2] O. Ozturk, H. Tatsuya, Y. Toshihiko, and A. Kiyoharu, “Content-aware bit-rate and replenishment control for video transmission of wireless multi-camera surveillance systems,” in *Proceedings of ACM/IEEE ICDSC 2007*, September 2007, pp. 394–395.

[3] X. Zhu, E. Setton, and B. Girod, “Rate allocation for multi-camera surveillance over an ad hoc wireless network,” in *Proceedings of PCS-04*, December 2004.

[4] Y. Jin and G. Kesidis, “Distributed contention window control for selfish users in IEEE 802.11 wireless LANs,” *IEEE JSAC*, vol. 25, no. 6, pp. 1113–1123, August 2007.

[5] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999.

[6] A. Turing, “The chemical basis of morphogenesis,” *Philosophical Transactions of the Royal Society of London*, vol. B. 237, no. 641, pp. 37–72, August 1952.

Naoki Wakamiya received the M.E. and Ph.D. degrees from Osaka University in 1994 and 1996, respectively. He was a Research Associate of Graduate School of Engineering Science, Osaka University from April 1996 to March 1997, a Research Associate of Educational Center for Information Processing from April 1997 to March 1999, an Assistant Professor of Graduate School of Engineering Science from April 1999 to March 2002. Since April 2002, He is an Associate Professor of Graduate School of Information Science and Technology, Osaka University. His research interests include self-organized networking of overlay networks, sensor networks, and mobile ad-hoc networks. He received the 2nd IEEE ComSoc Asia-Pacific Young Researcher Award in 2005. He is a senior member of IEICE and a member of IPSJ, ACM, and IEEE.

Masayuki Murata received the M.E. and D.E. degrees in Information and Computer Science from Osaka University, Japan, in 1984 and 1988, respectively. In April 1984, he joined Tokyo Research Laboratory, IBM Japan, as a Researcher. From September 1987 to January 1989, he was an Assistant Professor with Computation Center, Osaka University. In February 1989, he moved to the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University. In April 1999, he became a Professor of Cybermedia Center, Osaka University, and is now with Graduate School of Information Science and Technology, Osaka University since April 2004. He has more than four hundred papers of international and domestic journals and conferences. His research interests include computer communication networks, performance modeling and evaluation. He is a member of IEEE, ACM and IEICE.

Cross-layer Optimization in Video Sensor Networks

Tommaso Melodia, SUNY Buffalo, USA
Wendi Heinzelman, University of Rochester, USA
tmelodia@eng.buffalo.edu, wheinzel@ece.rochester.edu

1. Introduction

Video Sensor Networks (VSN) (also referred to as *multimedia sensor networks* [1], or *visual sensor networks* [2]) are being made possible by the integration of low-power wireless networking technologies with inexpensive CMOS cameras and microphones. Video sensor networks are self-organizing, intelligent wireless systems of embedded and resource-constrained devices deployed to retrieve, distributively process, store, correlate, and fuse multimedia streams. VSNs can potentially constitute a viable alternative to ubiquitously deployed wired surveillance systems. Once properly regulated, the availability of affordable VSN systems will enhance the ability of private citizens and law enforcement officers to observe and monitor locations and events in an unprecedented way, enabling sophisticated real-time scene analysis. We envision that users will be able to gather information about the physical environment by issuing simple textual queries, accessing remote VSNs connected to the Internet through application-level gateways.

The characteristics of VSNs diverge considerably from wired network paradigms such as the Internet, and even from traditional sensor networks. VSN applications require the sensor network paradigm to be re-thought in view of the need for *adaptation* and *cross-layer optimization* to deliver video/multimedia content with predefined levels of quality of service (QoS). While minimizing the energy consumption has been the main objective in sensor network research, mechanisms to efficiently deliver application-level QoS (e.g., target video distortion), and to map these requirements to network-layer metrics have not been primary concerns.

In this position paper, we discuss some key research challenges in video sensor networks from a network design perspective. In particular, we discuss the need for adaptation and cross-layer optimization in protocol design, and the need for a tighter integration and co-design between networking functionalities and image/video sensing and processing.

2. Cross-Layer Networking

In multi-hop wireless networks the attainable capacity of each wireless link depends on the interference level at the receiver. This, in turn, depends on the interaction of functionalities that are distributively handled, such as power control, routing, and rate policies. Hence, capacity and delay attainable at each link are location dependent, vary continuously, and may be bursty in nature, thus making QoS provisioning a challenging task. Therefore, there is a strict interdependence among functions handled at all layers of the communication stack, which are inherently and strictly coupled due to the shared nature of the wireless communication channel. In addition to this, performance metrics in VSNs may be directly tied to the perceived video quality at the receiver rather than to traditional network metrics such as throughput. Furthermore, the required video quality may vary over time, and the network must adapt to meet these changes, providing the required video quality while minimizing resource utilization. Cross-layer protocol architectures, including information-sharing and layer fusion designs, enable such coupling of the protocols and allow adaptation for optimal resource utilization in VSNs.

Cross-layer Information-sharing. An information-sharing architecture that provides shared data repositories that all protocols can access is described in [3]. In this architecture, called X-Lisa, there is a common interface to the data repositories that allow protocols to update data and to read data when needed. This architecture includes middleware support such that application-level information can be shared with network protocols. This is vital for VSNs, where the QoS information (e.g., distortion) must be considered by the protocols for optimal resource utilization. Additionally, X-Lisa supports proactive event notification, such that protocols can subscribe to be notified of a change in any data stored in the repositories (for example, changes in link quality) or of any changes in application QoS. This enables the protocols to react immediately to important changes in the network or the application goals, ensuring efficient operation of the VSN while continuously meeting QoS goals.

Cross-layer architectures such as X-Lisa provide protocols with access to network-level and application-level information. One important challenge is how to use this information to adapt the protocols. For example, as link conditions or QoS requirements change, at the PHY layer, the transmit power or the packet length may be adjusted, while at the MAC layer, duty cycle or back-off windows may be adjusted. How best to adapt and coordinate the adaptation of the different layers and across different VSN nodes is an important area of research.

Video Quality-driven Cross-layer Optimization. An alternate approach that considers an integrated, cross-layer architecture for video streaming in VSNs is pursued in [4]. A multi-hop wireless network of video sensors deployed for surveillance applications is considered, and the focus is on reliable and real-time transport of video traffic. The objective is to design algorithms to *efficiently* and *fairly* share the common network infrastructure among the video streams generated by different video sensors, to deliver high-quality video on resource-constrained devices. To achieve this objective, the Distortion-Minimizing Rate Control (DMRC) algorithm is proposed, a decentralized cross-layer control algorithm that jointly regulates the end-to-end data rate, the video encoding rate, and the channel coding rate at the physical layer to minimize the distortion of the received video. The end-to-end data rate is chosen to avoid congestion while maintaining fairness in the domain of *video quality* (rather than data rate as in traditional rate control algorithms). Once the end-to-end data rate has been determined, the sender calculates the optimal proportion of video encoder rate and channel encoder rate based on the overall rate available and on the current quality of the wireless channel on the source-destination path, with the objective of minimizing the video distortion at the receiver. Cross-layer rate control algorithms designed to minimize the energy consumption while preserving video quality will be another important area of research.

3. Integration Of Video Sensing, Processing And Networking

Sensing and processing of multimedia content has mostly been approached as a problem isolated from the network design problem. Hence, research that addressed the content delivery aspects has typically not considered the characteristics of the source content. However, the sensing, processing and delivery of multimedia content are not independent, and their interaction has a major impact on performance. VSNs will support in-network processing algorithms operating on the sensed data. Hence, the QoS required at the application level will be delivered by means of a combination of cross-layer optimization and in-network processing of sensed data streams that describe the phenomenon of interest from multiple views, with different media, and using multiple resolutions. Thus, it is necessary to develop application-independent and self-organizing architectures to efficiently perform in-network processing of multimedia content. Examples of such interactions between sensing, processing and networking include compressed sensing paradigms as well as power-rate-distortion frameworks for resource allocation.

Video Encoding Based on Compressed Sensing. As an example of integration between sensing, processing and networking, video encoders based on the recently proposed compressive sensing (CS) paradigm [5] may offer a viable solution to the problems of encoder complexity and limited resiliency to channel errors that characterize predictive encoders. Compressed sensing is a new paradigm that allows the recovery of signals from far fewer measurements than methods based on Nyquist sampling. In particular, the main result of CS is that an N-dimensional signal can be reconstructed from M noise-like incoherent measurements as if one had observed the $M/log(N)$ most important coefficients in a suitable base [6]. Hence, CS can offer an alternative to traditional video encoders by enabling imaging systems that sense and compress data simultaneously *with low-complexity encoders*.

In [7], the performance and potential of CS-based video transmission in video sensor networks was evaluated. In CS, the transmitted samples constitute a random, incoherent combination of the original image pixels. This means that, unlike traditional wireless imaging systems, in CS no individual sample is more important for image reconstruction than any other sample. Instead, *the number of correctly received samples* is the only main factor in determining the quality of the received image. Hence, a peculiar characteristic of CS video is its *inherent and fine-grained spatial scalability*. The video quality can be regulated at a much finer granularity than traditional video encoders, by simply varying the number of samples per frame. Also, as shown in [7] a small amount of random channel errors does not affect the perceptual quality of the received image *at all*, since, for moderate BERs, the greater sparsity of the "correct" image will offset the error caused by the incorrect bit. CS image representation is completely *unstructured*: this fact *makes CS video* *more resilient than existing video coding schemes to random channel errors*. This simple fact has deep consequences on protocol design for end-to-end wireless transport of CS video (especially at the link and transport layers of the protocol stack), which are to be addressed in future research.

Power-Rate-Distortion Models. Another example of the need for tight integration of the networking with the sensing and processing is shown in the resource optimization frameworks that have recently been developed. Recognizing that VSNs are limited both in bandwidth and in energy, these frameworks extend the traditional rate-distortion (R-D) models for image processing to include power consumption. These power-rate-distortion (P-R-D) frameworks can be used to determine the optimal allocation of the constrained resources to the different components of a VSN node [8], [9]. Initial work on P-R-D frameworks for VSNs looked at the integrated optimization of power and rate for the compression and transmission modules [8]. More recent work includes the sensing in the optimization framework to ensure that all aspects of the VSN node are considered in resource allocation for a particular QoS (distortion) [9]. These works show the potential benefit of integrated optimization considering a simplified networking scenario. How to extend these frameworks considering multiple cameras' resources and medium access control issues is a challenging problem.

4. Conclusions

In conclusion, VSNs have unique challenges due to their high data rates and severe energy constraints coupled with the bandwidth-limited and time-varying nature of wireless networks. Cross-layer designs that integrate decisions not only of the network layers but also considering the sensing and processing are crucial to the efficient operation of future VSNs.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A Survey on Wireless Multimedia Sensor Networks," *Computer Networks (Elsevier)*, vol. 51, no. 4, pp. 921–960, Mar. 2007.

[2] S. Soro and W. Heinzelman, "A survey of visual sensor networks," *Advances in Multimedia*, 2009.

[3] C. Merlin, C.-H. Feng and W. Heinzelman, "Information-sharing Architectures for Sensor Networks: the State of the Art," in *ACM Mobile Computing and Communications Review (MC2R)*, Vol. 13, No. 4, Oct. 2009, pp. 26-38.

[4] S. Pudlewski and T. Melodia, "DMRC: Distortion-minimizing Rate Control for Wireless Multimedia Sensor Networks," in *Proc. of IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, Hong Kong S.A.R., P.R. China, October 2009.

[5] D. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[6] J. Romberg, "Imaging via Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14–20, 2008.

[7] S. Pudlewski and T. Melodia, "On the Performance of Compressive Video Streaming for Wireless Multimedia Sensor Networks," in *Proc. of IEEE Int Conf on Communications (ICC)*, Cape Town, South Africa, May 2010.

[8] Z. He and D. Wu, "Resource allocation and performance analysis of wireless video sensors," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, No. 5, pp. 590-599, May 2006.

[9] M. Marijan, W. Heinzelman, G. Sharma and Z. Ignjatovic, "Optimal Resource Allocation for Wireless Video Sensors with Power-Rate-Distortion Model of Imager," in *IEEE Midwest Symposium on Circuits and Systems (MWSCAS)*, 2009.

Tommaso Melodia [M2007] (tmelodia@buffalo.edu) is an Assistant Professor with the Department of Electrical Engineering at the University at Buffalo, The State University of New York (SUNY). He received his Ph.D. in Electrical and Computer Engineering from the Georgia Institute of Technology in 2007. He had previously received his Laurea (integrated B.S. and M.S.) and Doctorate degrees in Telecommunications Engineering from the University of Rome La Sapienza, Rome, Italy, in 2001 and 2005, respectively. He coauthored a paper that was recognized as the Fast Breaking Paper in the field of Computer Science for February 2009 by Thomson ISI Essential Science Indicators. He is an Associate Editor for the Computer Networks (Elsevier) Journal. He serves in the technical program committees of several leading conferences in wireless communications and networking, including IEEE Infocom, ACM Mobicom, and ACM Mobihoc. He was the technical co-chair of the Ad Hoc and Sensor Networks Symposium for IEEE ICC 2009. His current research interests are in modeling and optimization of multi-hop wireless networks, cross-layer design and optimization, cognitive radio networks, multimedia sensor networks, and underwater acoustic networks.

Wendi Heinzelman is an associate professor in the Department of Electrical and Computer Engineering at the University of Rochester. She hold a secondary appointment in the Computer Science Department at Rochester. Dr. Heinzelman also currently serves as Dean of Graduate Studies for Arts, Sciences and Engineering at the University of Rochester. Dr. Heinzelman received a B.S. degree in Electrical Engineering from Cornell University in 1995 and M.S. and Ph.D. degrees in Electrical Engineering and Computer Science from MIT in 1997 and 2000, respectively. Her current research interests lie in the areas of wireless communications and networking, mobile computing, and multimedia communication. Dr. Heinzelman received the NSF CAREER award in 2005 for her research on cross-layer architectures for wireless sensor networks, and she received the ONR Young Investigator Award in 2005 for her work on balancing resource utilization in wireless sensor networks. She is an Associate Editor for the IEEE Transactions on Mobile Computing, an Associate Editor for the ACM Transactions on Sensor Networks, and an Associate Editor for Elsevier Ad Hoc Networks Journal. Dr. Heinzelman is a senior member of the IEEE and the ACM, and she is co-founder of the N^2 Women (Networking Networking Women) group.

Harnessing Collective Power of Sensor Nodes as Distributed Embedded Systems

Eylem Ekici, The Ohio State University, USA
ekici@ece.osu.edu

1. Introduction

With the evolution of the MEMS technology and the availability of low cost communication and computation hardware, Wireless Sensor Networks (WSNs) have been transformed from conceptual paradigms to reality [1] and are being used for environmental monitoring, urban safety, traffic monitoring, and battlefield surveillance. A big step forward in the realm of WSNs is the move from simple sensor hardware to more complex hardware to capture multimedia content. The so-called Multimedia Wireless Sensor Networks (MMWSNs) capture still images, video, and audio content from the observed area [2, 3, 4]. The multimedia content gathered by MMWSNs does not only increase the precision of applications including target tracking, environmental monitoring, and vehicle traffic monitoring and control, but also enable new applications such as accurate in-home patient monitoring, vision-based safety systems, and automated assistance for elderly and disabled in public areas.

Delivery of multimedia content has been investigated for wired networks extensively over the past two decades. As in wired networks, high bandwidth demand and low and consistent delay requirements of multimedia streams are also primary concerns for MMWSNs. Additionally, low processing power of sensor nodes, energy constraints, and limited bandwidth availability must be considered while sustaining multimedia traffic in MMWSNs. Therefore, communicating multimedia content in sensor networks is a challenging task that requires further research. In this article, we highlight the interdependence of in-network processing and the real-time communication requirements inherent to MMWSNs. Based on these observations, we advocate the use of distributed in-network processing to reduce the communicated information and to meet application-imposed latency constraints.

2. Multimedia Delivery over MMWSNs

MMWSNs are special types of WSNs that consist of small wireless nodes equipped with multimedia sensors such as microphones, still image, video, and infrared cameras. Surveillance is the primary application in a MMWSN. An example surveillance scenario with multimedia sensor is depicted in Figure 1. In this example, let the area marked with the ellipse be observed with multimedia sensors to detect and classify intruders. The information is obtained in streaming mode by multiple sensors. Different streams can optionally be processed in the network for aggregation to reduce the communicated data volume exploiting correlations in multimedia streams. Alternatively, multimedia data can also be processed according to application requirements. The resulting data stream is encoded using a Multiple Description Code [5] and delivered to the sink over multiple hops subject to delay and reliability requirements.

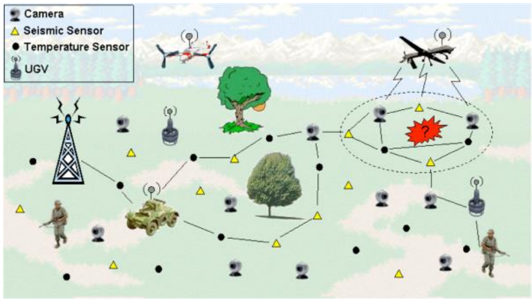
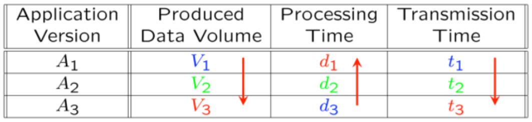


Fig. 1: Example Surveillance Scenario with MMWSN.

The requirements of the information delivered to the sink are directly determined by the application. Consider an example application where targets need to be classified. On the one extreme, information can be minimally processed and delivered to the sink for actual processing. Here, the raw data volume is high, but the processing required of sensors is minimal.

On the other extreme, the entire classification can be done in the network, and only the outcome can be delivered to the sink. This results in significant processing in the network, yet very small communication overhead. The different levels of processing in the network resulting in different data volumes are referred to as versions of the same application. The effect of in-network processing and how communication and computation latencies can be accounted for is exemplified in Figure 2, where red arrows indicate an increase in the magnitude.



Application Version	Produced Data Volume	Processing Time	Transmission Time
A_1	V_1	d_1	t_1
A_2	V_2	d_2	t_2
A_3	V_3	d_3	t_3

$$d_i + t_i \leq D \; \forall i \in \{1, 2, 3\}$$

Fig. 2: Effect of In-Network Processing on Performance

The crucial question that remains is how to implement in-network processing in a resource limited environment. Since most multimedia processing algorithms require significant computational and storage resources, this may seem to be an impossible task. However, it is possible to distribute the processing of information to various nodes in the network so as to minimize the execution time and reduce the processing as well as storage requirements for individual sensor platforms. Deferring the tradeoff between processing and communication to future investigations, we will focus on how a version of a complex algorithm can be implemented in a distributed manner in MMWSNs.

3. Distributed Processing in MMWSNs

Information aggregation in WSNs has been recognized as a powerful method to reduce the load on the network and prolong its lifetime. Reduction of communicated multimedia volume is an important step to reduce energy consumption in MMWSNs, as well. The processing power required to process multimedia content can be obtained through coordinated use of multiple sensor nodes. The processing can be done to aggregate correlated multimedia streams, to compress data, or extract application-specific information. As an example, two still images obtained from overlapping areas can be collated to generate a smaller volume image using image registration algorithms. Similarly, images obtained from multiple cameras can be processed in the network to deliver the location of an object rather than raw image data. Here, we propose to distribute processing to multiple sensor nodes. The applications are divided into tasks and assigned to sensor nodes. Our proposed solutions schedule execution of tasks and communication events to exchange intermediate results between sensors. The communication and computation scheduling is done such that the execution time and energy consumption is minimized.

Let an application be represented as a Directed Acyclic Graph (DAG) $T = (V,E)$, where the vertex set $V = \{v_i\}_{i=1}^{N}$ denotes the tasks to be executed and the edge set $E = \{e_{i,j}\}_{i,j \in 1,\ldots,N}$ denotes the communication events from v_i to v_j, where N is the number of tasks of the application. The weight of vertexes represents computation cost in number of clock cycles. The weight of edges corresponds to the data volume in bits that must be transmitted between sensors if two dependent tasks are executed on different sensors. An example DAG is shown in Figure 3(a). DAGs can represent different multimedia processing applications in a unified format and allow the use of the same scheduling algorithms.

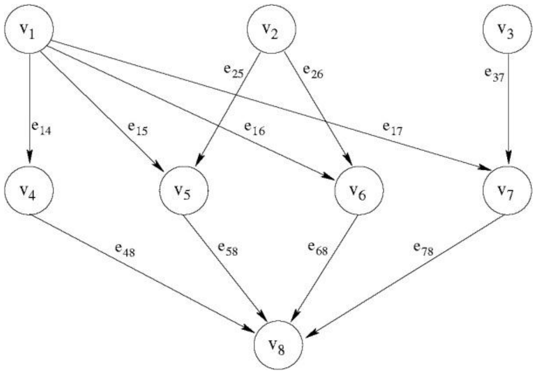
Given an application represented by a DAG, task matching (determining the location of task execution) and scheduling (determining the execution sequence) should be performed such that an objective function is minimized while satisfying a set of QoS constraints. Although this problem resembles traditional task matching and scheduling in high performance computing systems, WMSNs carry two important differences: (i) Sensors do not have dedicated point-to-point links and communicate over a shared wireless medium. (ii) Sensors have limited power supplies and communication events consume significant amounts of energy.

Since all communication must go over the same wireless medium, it is only logical to represent the wireless channel as a special kind of processor, denoted by C, that can only process communication events. However, unlike sensors that can execute one task at a time, the exclusive access to the wireless channel is subject to spatial and temporal conditions. For instance, communications of two source-destination pairs ($s_1 \rightarrow d_1$) and ($s_2 \rightarrow d_2$) can only be scheduled simultaneously if the distances $| s_1, d_2 |$ and $| s_2, d_1 |$ are both large enough to avoid harmful interference. To accommodate different channel models and to model the effect of interference on data fidelity, we introduce a penalty function $pen(s; t_1; t_2)$ that returns the penalty of the sensor s transmitting data over the wireless channel in the time period $[t_1, t_2]$. The penalty function is used in the task matching and scheduling algorithms to regulate the channel access. Under the simple unit disc graph model, $pen(s, t_1, t_2)$ reduces to $pen(s,t_1,t_2) =$

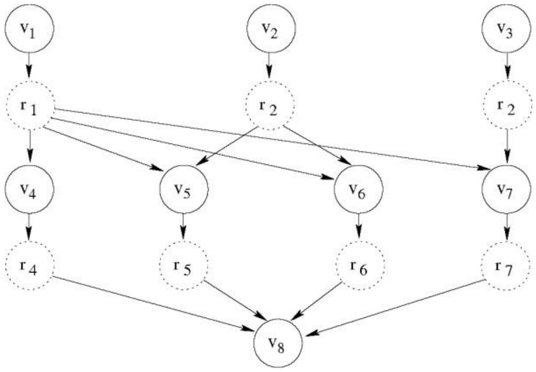
$$\begin{cases} \infty & \text{if another } s', |s, s'| < \text{R, receives data in}[\text{t1, t2}] \\ 0 & \text{otherwise} \end{cases}$$

where R is the communication range and $|s,s'|$ is the distance between sensors s and s'.

While communication over the wireless medium presents challenges, the broadcast nature of the wireless communication can also be leveraged to deliver information to multiple destinations in a single transmission. If the successors of a task are executed in sensors in a single hop neighborhood, then a single transmission suffices to communicate intermediate results. This method reduces the energy expenditure in the network. Note that communication between nodes beyond a single hop neighborhood requires multiple transmissions.



(a) Directed Acyclic Graph (DAG)



(b) Corresponding Hyper-DAG

Fig. 3: Example Graph Representations of an Application

At this point, we would like to analyze a special case where all sensors are located in a single hop cluster. In this case, the penalty function $pen(s,t_1,t_2)$ reduces to a form where the channel processor can only be accessed exclusively. Furthermore, to represent the broadcast delivery of results in a single hop neighborhood, we replace the vertexes between tasks with *nets* [1] that connect the predecessor and successor tasks over directed edges. With this representation, the task graph is converted to a *hypergraph* with directed edges, which we refer to as *Hyper-DAG*. In Figure 3(b), the Hyper-DAG that corresponds to the DAG of Figure 3(a) is shown. In the Hyper-DAG representation, the channel node C can only execute communication tasks represented as nets (dotted circles), whereas tasks represented as vertexes are executed in sensor nodes. Since the communication cost is shifted to the nets (denoted by r_i), the links are not associated with weights. Considering nets as special vertexes which can only be executed on *C*, Hyper-DAGs are used to schedule communication and computation tasks *jointly* with the same algorithm. Based on this model, various scheduling algorithms can be executed to determine the schedule of communication and computation tasks and their distributions to different nodes.

This approach has been tested in single [6] and multihop [7] WSN clusters with image-based localization algorithms, where four images are captured in a setting shown in Figure 4. Our implementations show significant gains in terms of energy consumption and schedule length over central processing of the information inside the network. Moreover, it has also been shown that reduction in data volume from hundreds of kB to tens of bytes significantly saves energy when communicating the outcome to the sink.

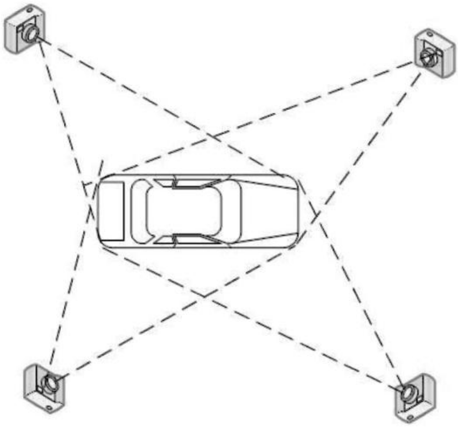


Fig. 4: Example Application Scenario.

4. Conclusions

In this article, a new view on the multimedia delivery and processing for WSNs has been proposed. The main idea is to process the raw data in the network according to application requirements to significantly reduce the communicated data volume. This also allows for a very efficient framework to trade latency for energy consumption. A proof-of-concept implementation is one of our most immediate next steps.

---

[1] Structures that connect multiple vertexes in a hypergraph.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaninam, and E. Cayirci, “Wireless Sensor Networks: A Survey,” Computer Networks Journal (Elsevier), vol. 38, no. 4, pp. 393–422, March 2002.

[2] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “A Survey on Wireless Multimedia Sensor Networks,” Computer Networks (Elsevier), vol. 51, no. 4, pp. 921–960, Mar. 2007.

[3] J. Campbell, P. Gibbons, S. Nath, S. Seshan, and R. Sukthankar, “IrisNet: An Internet-scale Architecture for Multimedia Sensors,” Proceedins of ACM Multimedia Conference, 2005.

[4] W. Feng, B. Code, E. Kaiser, W. Feng, and M. L. Baillif, “Panoptes: Scalable Low-Power Video Sensor Networking Technologies,” ACM Transactions on Multimedia Computing, Communications, and Applications, Jan. 2005.

[5] Y. Wang, A.R. Reibman, and S. Lin, “Multiple Description Coding for Video Delivery,” Proceedings of the IEEE, vol. 93, no. 1, pp. 57–70, Jan. 2005.

[6] Y. Gu, Y. Tian, and E .Ekici, “Real-Time Multimedia Processing in Video Sensor Networks,” Signal Processing: Image Communication Journal (Elsevier), vol. 22, no. 3, pp. 237-251, March 2007.

[7] Y. Tian and E .Ekici, “Cross-Layer Collaborative In-Network Processing in Multi-Hop Wireless Sensor Networks,” IEEE Transactions on Mobile Computing, vol. 6, no. 3, pp. 297-310, March 2007.

Eylem Ekici has received his BS and MS degrees in Computer Engineering from Boğaziçi University, Istanbul, Turkey, in 1997 and 1998, respectively. He received his Ph.D. degree in Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta, GA, in 2002. Currently, he is an associate professor in the Department of Electrical and Computer Engineering of The Ohio State University, Columbus, OH. He is an associate editor of IEEE/ACM Transactions on Networking, Computer Networks Journal (Elsevier), and ACM Mobile Computing and Communications Review. Prof. Ekici is the recipient of 2008 Lumley Research Award of the College of Engineering at OSU. His current research interests include cognitive radio networks, wireless sensor networks, vehicular communication systems, and nano-communications systems, with a focus on routing and medium access control protocols, resource management, and analysis of network architectures and protocols. He is a member of IEEE and ACM.

Open Issues in Secure Wireless Multimedia Sensor Networks

L. A. Grieco and G. Boggia, Politecnico di Bari, Italy
S. Sicari, Università degli Studi dell'Insubria, Italy
a.grieco@poliba.it, sabrina.sicari@uninsubria.it, g.boggia@poliba.it

1. Introduction

Nowadays, technology is mature enough to allow the production of multimedia wireless sensors, i.e., wireless nodes able to acquire and process audio/video signals. A Wireless Multimedia Sensor Network (WMSN) relies on a set of collaborating multimedia wireless sensors to provide distributed multimedia sensing services (see Fig. 1). Potential WMSN application domains range over indoor/outdoor surveillance systems, traffic monitoring and control systems for urban and sub-urban areas, systems supporting telemedicine, attendance to disable and elderly people, environment monitoring, localization and recognition of services and users, monitoring and control of manufacturing processes in industry [1]. The richness of information retrieved and delivered by multimedia monitoring applications poses new interesting problems to afford in order to transform WMSN potentials in real revenues. In fact, beside the need of providing the desired Quality of Experience (QoE) using power and computational constrained nodes [2], there are also important issues to face regarding security and privacy [3,4,13] (for instance, consider the high critical, sensitive and reserved information managed by networks that support telemedicine or monitoring services in countries threatened by terrorism).

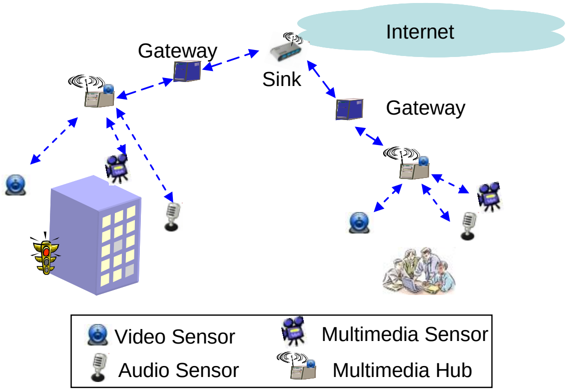


Fig. 1: Wireless Multimedia Sensor Network.

The transition from WMSNs to Secure WMSNs (SWMSNs) requires that security and privacy policies be combined with complex algorithms for compression and distributed multimedia processing. For that purpose, new tools and methodologies have to be conceived, but radically different from those used for classic wireless sensor networks. The innovative features of this research field are also underlined by the promising technological innovations in the area of cognitive communication architectures and nano-technologies, which can significantly increase the transmission and computational capabilities of sensors [5,6,7].

It is worth to note that SWMSN architecture perfectly fits the Internet of Things (IoT) framework [8] extending the set of supporting technologies (e.g., RFIDs, sensor/actuators) necessary to realize IoT vision and then improving the intelligence level of things. In fact, the ability to interface with the physical realm is achieved through the presence of devices able to sense physical phenomena as SWMSN (thereby providing information on the current context and/or environment), as well as through the presence of devices able to trigger actions having an impact on the physical realm (through suitable actuators).

This letter focuses on SWMSNs, highlighting the most important issues to afford in order to boost their development in real scenarios. In particular, five important areas are described, in which novel approaches are required to solve open problems of SWMSNs.

2. Open Issues in SWMSNs

The path to SWMSNs is well summarized in Fig. 2: in our vision simple WMSNs have to be enriched with a set of new security-aware techniques able to improve the QoE provided to final users.

Secure management of QoS and QoE

At present, Quality of Service (QoS) and QoE are being addressed by means of cross layer approaches and ad-hoc scheduling algorithms. The standards IEEE 802.11e and 802.15.3 [9] represent the reference solutions, since they are equipped at MAC level with those mechanisms required to manage the data flow under different QoS requirements. *How secure are these mechanisms? Can they be exploited in subtle attacks?* These questions represent the first obstacle to afford in order to enable secure management of QoS and QoE. Furthermore, the use of new communication technologies based on cognitive radio networks

[6,7] should be also investigated to boost network performance.

Secure Aggregation of Multimedia Contents

Compression techniques and aggregation algorithms for multimedia contents target strong reductions of transmitted/processed data (and spent energy) in WMSNs. The problem of aggregating multiple compressed frames coming from different multimedia sensors while guaranteeing the expected security and quality levels is still an open research area. The literature provides a great variety of end-to-end and hop-by-hop secure aggregation protocols [10], but they can be hardly applied to multimedia data. For instance, the encryption of images is a highly power consuming task, hence the most innovative solutions try to adopt selective encryption schemes for the multimedia contents [11]. In such contexts, it is possible to describe an image as composed of different and mutual integrable qualitative levels, and to encrypt only the data of the basic level, by making useless any attack oriented to the theft of no encrypted transmitted data. This approach allows innovative strategies be conceived to discard less important frames when the network become congested. Another possible research direction is based on exploiting compression algorithms that do not strictly require entropy coding, such as Set-Partitioning In Hierarchical Trees image. In fact, since entropy coding requires a high computational effort, the energy efficiency of the WMSN would be improved [12].

Privacy & trust management

In many WMSN scenarios, such as telemedicine or of military surveillance, data are sensitive and are required to be adequately protected. The privacy solutions available in literature focus on the specific aspects of data cloaking, secure communication channel, definition of privacy policies [3]. Each solution satisfies only specific requirements for ad-hoc problems, in other words, no single proposal is able to provide a complete privacy solution for WMSNs. The study of integrated theoretical solutions, as well as the development of HW/SW platforms supporting them, represents a great challenge for the scientific community. A possible approach towards the achievement of this goal passes through the definition of a privacy model for WMSNs. In fact, such a model should support the definition of privacy enabling mechanisms that overcome the limits of WMSN, and the definition of enforcement schemes that guarantee the correct and automatic application of the privacy policies defined for WMSNs. Such enforcement mechanisms aim at verifying the compliance of the processing activities with the privacy policy and indicate the actions that are required to be executed in case behavioral anomalies are identified. It is really important notice that the model should be used in combination with both data cloaking mechanisms and some other privacy policy based approaches.

Moreover, in a distributed and collaborative environment like a WMSN, trust management becomes a real challenging aspect. All the data exchanged among nodes have to be trusted, above all information related to authentication and localization data. For this reason, research should focus on the development of a flexible framework usable in several application scenarios. However, the definition of an effective model of trust becomes a complex task in a highly distributed environment characterized by strict performance requirements. Each node should be equipped with an autonomous evaluation and analysis capabilities that aim at measuring the trust relationships with the other members of the network; notice that such relationships depend on the communication and cooperation needs of the nodes. In other words, it is required to move from the classic centralized and static approach proposed for the most widely used trust management solutions, to adopt a fully distributed and dynamic approach that assumes that no trust relationship is defined a priori among the nodes of the network. At present only few solutions are available [14], but they cannot be applied to WMSNs due to the relevant computational effort required by the multimedia traffic and to the real-time constraints that are not suited to the limited power resources of current sensor nodes.

Development of nano-technology

Nano-technology [6,7] should allow to overcome the constrains imposed by the currently available technologies to supply distributed and secure monitoring services based on WMSNs. Upcoming solutions coming from this promising field should be encouraged and timely exploited in order to design next generation monitoring applications.

Computer-aided Tools for SWMSNs Design and Performance Evaluation

Technological and architectural innovations described above, if conceived, have to be properly integrated and evaluated to provide the expected performance gains. To this end, computer-aided design methodologies play a key role. SWMSN simulators should be developed in order to encompass all facets of these sensing platforms, from PHY details to secure distributed multimedia processing aspects. The efforts of the community are currently very spread on many simulation platforms, each one focusing a tightly bounded view of the entire problem. With this letter we also encourage choral efforts targeting reference solutions for evaluating SWMSN performance.

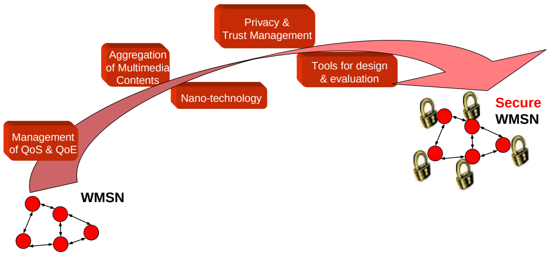


Fig. 2: Path from WMSNs to SWMSNs.

References

[1] I. F. Akyildiz, T. Melodia, K. R. Chowdhury "A survey on wireless multimedia sensor networks," Computer Networks 51 (2007) 921–960.

[2] I. F. Akyildiz, T. Melodia, K. R. Chowdhury "Wireless Multimedia Sensor Networks: Applications and Testbeds ," Proc. of the IEEE, vol. 96 (10), Oct. 2008.

[3] M. Guerrero-Zapata, R. Zilan, J. M. Barcelo-Ordinas, K. Bicakci, and B. Tavli, The Future of Security in Wireless Multimedia Sensor Networks: a position paper, accepted by the Special Issue "Secure Multimedia Services" in Telecommunications System Journal (Springer), 2009

[4] J. P. Walters, Z. Liang, W. Shi, and V. Chaudhary, "Wireless Sensor Network Security: A Survey", Security in Distributed, Grid, Mobile, and Pervasive Computing - chap. 7, Auerbach Publications, CRC Press, 2006.

[5] I. Downes, L. B. Rad, H. Aghajan, "Development of a mote for wireless image sensor networks," in Proc. of COGnitive systems with Interactive Sensors, COGIS, Paris, France, Mar. 2006.

[6] J. P. M. She and J. T. W. Yeow, "Nanotechnology-Enabled Wireless Sensor Networks: From a Device Perspective," IEEE Sensors Journal, Vol. 6 (5) Oct. 2006.

[7] I.F. Akyildiz, F.Brunetti, and C.Blazquez,, "NanoNetworking: A New Communication Paradigm", Computer Networks Journal, (Elsevier), June 2008.

[8] "*Internet of Things in 2020 - A Roadmap for the Future*", Paper by INFSO D.4 NETWORKED ENTERPRISE & RFID INFSO G.2 MICRO &NANOSYSTEMS in co-operation with the RFID WORKING GROUP OF THE EUROPEAN TECHNOLOGY PLATFORM ON SMART SYSTEMS INTEGRATION (EPOSS), 05 September, 2008.

[9] B. H. Walke, S. Mangold, and L. Berlemann, IEEE 802 Wireless Systems. NJ, USA: John Wiley & Sons, 2006.

[10] S. Ozdemir and Y. Xiao. Secure data aggregation in wireless sensor networks: a comprehensive overview. Computer Networks, 53, 2009.

[11] S. Misra, M. Reisslein, and X.Guoliang , "A survey of multimedia streaming in wireless sensor networks," IEEE Communications Surveys & Tutorials, Vol. 10(4), Fourth Quarter 2008.

[12] L.W. Chew, L. M. Ang, K. P. Seng, "Survey of Image Compression Algorithms in Wireless Sensor Networks", in Proc. of Int. Symp. on Information Technology (ITSim 2008), Aug. 2008.

[13] Y. Wang, G. Attebury, B. Ramamurthy, "A survey of security issues in wireless sensor networks," IEEE Commun. Surveys & Tutorials, vol.8, no.2, pp.2-23, Second Quarter 2006.

[14] L. A. Grieco, G. Boggia, S. Sicari, and P. Colombo, " Secure Wireless Multimedia Sensor Networks: a Survey", *Proc. of The Third International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM*, Sliema, Malta, Oct., 2009.

Luigi Alfredo Grieco received, with honors, the Dr. Eng. Degree in Electronic Engineering in October 1999 from 'Politecnico di Bari', Italy. He received the Ph.D. in Information Engineering on December 2003 from 'Università di Lecce'. Since January 2005 he holds an assistant professor position in Telecommunications at 'Politecnico di Bari', 'Dip. Elettrotecnica ed Elettronica (DEE)'. From March to June 2009, he has been visiting researcher at the INRIA (Planete Project, Sophia Antipolis, France), working on the topics "Internet Measurements" and "Scheduling in WiMax Networks". His main research interests are: congestion control in packet switching networks, QoS and Service Discovery in wireless networks, video delivery in the Internet, Internet measurements, and real-time video processing using Cellular Nonlinear Networks. He has authored more than 80 scientific papers about these topics, published on international journals and conference proceedings.

Sabrina Sicari was born on September 18, 1977 in Catania, Sicily, Italy. She received her laurea degree in Electronical Engineering,110/110 cum laude, from University of Catania, Catania, Italy, in 2002. In March 2006 she got her Ph.D. in Computer and Telecommunications Engineering at the same university, under the guidance of Prof. Aurelio La Corte. From September 2004 to March 2006 she has been a

Visiting Scholar at Dipartimento di Elettronica e Informatica, Politecnico di Milano, Italy, where her research concerned risk assessment methodology and web service security (European project Secse), under the guidance of Prof. Carlo Ghezzi. Since May 2006 she works at Dipartimento di Informatica e Comunicazione, Università degli Studi dell'Insubria, in software engineering group. She is reviewer of Pervasive and Mobile Computing (Elsevier), Computer networks (Elsevier) IEEE Transactions on Vehicular Technology, International Journal of Computer Applications in Technology (IJCAT), ACM-Monet, ICC'09, ICC'10 (IEEE International Conference on Communications), IEEE ISIE'10, S-Cube'09 and WiOpt'09 (Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks)and TPC member of Q2SWinet 2009. Dr. Sicari is an Editor for Computer Networks (Elsevier) journal since 2008. She is the general co-chair of S-Cube 2009 and guest editor for the ACM Monet Special Issue, named " Sensor, system and Software". She is a Steering Committee member of S-Cube 2010. She is TPC member of the international conference IEEE Globecom'10, Q2SWinet'09, the international workshop SESENA 2010 (co-located with ICSE'10), Q2SWinet'10 and ICC'11. Her research interests are on: definition of risk assessment methodologies, privacy policy modelling, integration of heterogeneous wireless network, quality of service for wireless and VoIP (Voice over IP) systems, IoT (Internet of Things), wireless sensor network (WSN, WMSN), in particular secure localization algorithms by means of a cross layer approach, low power consumption localization algorithm for underwater sensor network, secure data aggregation algorithm for wireless sensor networks (WSNs) and wireless multimedia sensor networks (WMSNs).

Gennaro Boggia received, with honors, the Dr. Eng. Degree in Electronics Engineering in July 1997 and the Ph.D. in Electronics Engineering in March 2001, both from the ‘Politecnico di Bari’, Italy. Since September 2002, he has been with the department of ‘Elettrotecnica ed Elettronica’ at the ‘Politecnico di Bari’, Italy, where he is currently Assistant Professor. From May 1999 to December 1999, he was visiting researcher at the ‘TILab’, TelecomItalia Lab (formerly CSELT, Centro Studi e Laboratori Telecomunicazioni), Italy, where he was involved in the study of the Core Network for the next releases of universal mobile telecommunications system (UMTS). In 2007, he was visiting researcher at FTW (Vienna), where he was involved in activities on passive and active traffic monitoring in 3G networks. He has authored or coauthored more than 70 papers in international journals or conference proceedings. His research interests span the fields of Wireless Networking, Multimedia Systems, Cellular Communication, Queueing Networking, and Network Performance Evaluation.

Video Transmission Over A Standards-Based Wireless Multi-Hop Sensor Network

Thomas Watteyne, Fabien Chraim, Nahir Sarmicanic, Chris Jian, Kristofer S. J. Pister,
University of California, Berkeley, USA
{watteyne,chraim,cajian,pister}@eecs.berkeley.edu, nahira@berkeley.edu

Abstract

Video transmission combines large quantities of data with real-time requirements, two constraints which are hard to meet in low-power wireless multi-hop networks. This position paper presents experimental results of multi-hop video transmission in an IEEE802.15.4-based wireless network, using a protocol stack based solely on standards which are being finalized. This practical look allows us to quantify the performance one can expect from such a system, and to underline the areas where further investigation is needed.

1. Opportunities and Challenges

In most applications, Wireless Sensor Networks (WSNs) carry small amounts of sensor data to a sink node, with the duration between two sensor reports which varies from minutes to days. Video transmission sits at the opposite end of the spectrum, and hence puts new challenges on the protocol stack, especially on the Medium Access Control (MAC) layer. This letter shows how Time-Synchronized Channel Hopping (TSCH) – a MAC technology being standardized by the IEEE802.15.4e working group – meets those requirements and can be used for video transmission.

Using a wireless multi-hop network of small low-power embedded devices for transmitting video opens up a new range of possibilities. Following an earthquake, micro autonomous robots could enter a collapsed building and drop off video-enabled sensors to help rescue teams map the rubble and assess the presence of people. Other application areas include surveillance, traffic monitoring and advanced health care [1].

The main challenges are low data rate and multi-hop operation:

- IEEE802.15.4 radios (the *de-facto* standard for such networks) communicate at 250kbps. A 128-byte-long packet (the largest size handled by those radios) hence takes just over 4ms to be sent. Taking into account processing, radio turnaround time and link layer acknowledgments, in practice, a packet is sent every *10ms* or so, causing the useful data rate to drop to *100kbps*.
- Let's assume a multi-hop path $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$, with source node A streaming video data to destination node E. One expects every link to be active continuously, i.e. while A sends a packet to B, B is relaying the previous packet to C. Yet, because radios are half-duplex, when A sends to B, B can not send to C, causing the effective data rate to be further reduced to *50kbps*[2].

At such low effective data rates, it is important to trade off image size (i.e. compression quality and pixel size) with the frame rate. Fig. 1 illustrates this by taking the canonical case of the network transmitting a succession of JPEG images. It shows how the quality of the images impacts their size, which in turn impacts the maximum frame rate – expressed in frame per second, *fps*. These images were collected using with the Python-based software used in the experiments described in Section 4.

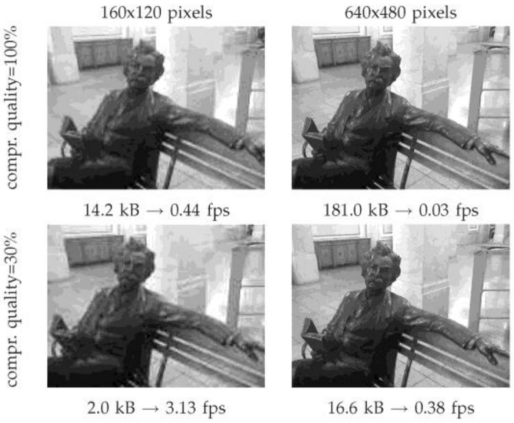


Fig. 1: The quality of the image impacts the frame rate. JPEG compression obtained using the Python Imaging Library (PIL), and an off-the-shelf webcam.

The remainder of this letter is organized as follows. Section 2 provides an overview of the related work, illustrating how multimedia transmission over

---

[2] It is sometimes assumed that when A sends to B, C can not send to D or else B will be exposed, causing the data rate to fall to *33kbps*. We show in Section 4 how frequency agile protocols can alleviate that problem.

WSNs can be tackled at all layers of a communication stack. Section 3 presents a protocol stack composed solely of to-be-finalized standards. This stack couples a Time Synchronized Channel Hopping MAC protocol (to enable video transmission) with 6LoWPAN/IPv6/UDP (to enable seamless integration within the Internet). Section 4 details the experimental setup and discusses the results. The areas that we believe require further investigation are outlined in Section 5, which also concludes this letter.

2. Related Work

Enabling video transmission over WSNs impacts the design of all layers in the protocol stack. In this section, we describe related work from the physical layer up to the application layer.

Several projects have looked at designing video-enabled daughter cards which plug into existing wireless motes. One example is the Cyclops project [2], which proposes a camera daughter card for the mica2 mote capable of performing simple inter-frame compression using an 8-bit Atmel micro-controller. Another, more recent, is the CITRIC project which proposes a camera daughter card for the TelosB mote which uses a 32-bit Intel microprocessor to locally process captured images before sending them through the network [3].

MAC layer design traditionally advantages low-power operation over efficient use of the available bandwidth. As a result, contention-based approaches such as preamble sampling or MAC protocols with common active periods suffer from network collapse at data rates exceeding a few *kbps* [4]. While suitable for very low throughput applications, this makes them ill-suited for video transmission.

Most experimental studies have hence opted for single-hop communication [2], [3], or used higher throughput radio technologies such as IEEE802.11 [5]. To the best of our knowledge, this is the first work to demonstrate multi-hop video communication on a low-power IEEE802.15.4-based network.

The MAC protocol approach used is Time Division Multiple Access (TDMA). Coupled with channel hopping, this technique – called Time Synchronized Channel Hopping (TSCH) – combats external interference and multi-path fading [6]. It has been used in proprietary solutions for industrial WSNs, and is being standardized by the IEEE through its IEEE802.15.4e working group. This position paper shows how it can be efficiently used for video transmission.

The way multi-hop routes are established influences the transmission of multimedia streams. Chen *et al.* [7] establish that single path routing is not suitable because of the unreliable nature of the wireless links and the fact that they are bandwidth limited. [7] therefore proposes a geographic routing scheme which allows data to flow over separate multi-hop routes. The performance of the network is evaluated by simulation.

Another important aspect in video over WSNs is compression, which directly impacts the amount of data that needs to be carried. He *et al.* [8] study the impact of video compression on power consumption and the quality of transmitted data. The authors extend the notion of Rate Distortion (R-D) to include Power (P), therefore developing an analytic P-D-R model for data compression over WSNs. Their analysis shows that efficient video compression is crucial in conserving network bandwidth and power consumption.

In traditional MPEG-x or H.26*x* video encoding schemes, encoding is more computationally intensive than decoding. In a WSN, the encoding source node is usually a mote while the decoding destination is a more powerful computer. The PRISM architecture [9] therefore proposes a compression scheme with can balance the computational load between source and destination. It is evaluated by simulation.

The SensEye project aims at designing a complete camera sensing network [5]. It uses a multi-tier topology (three in this case), where the lower tiers have more motes in the network. The low tier performs object detection and localization, and in turn uses that information to wake up the appropriate (closest) mid-tier motes for higher resolution object recognition, and if necessary wakes up the top tier camera attached to a computer. [5] shows how the hierarchical approach of SensEye consumes less energy than a single-tiered approach by a factor of 33, with only a 6% decrease in sensing reliability. It is to be noted however, that images are not transmitted at the lowest tier, but rather at the upper, IEEE802.11 enabled, tier.

[10] proposes Rate Controlled Variable Bit Rate (RCVBR), which uses the packet queue size to vary the transmission rate: when the queue gets full, the video quality is reduced . It is coupled with

Region Of Interest (ROI) encoding to reduce the amount of data transmitted. Simulation results show an average decrease of 40% in dropped frames, along with a 2.5dB increase in Peak Signal to Noise Ratio (PSNR). At rates of 10kbps, no loss is observed along good video quality. However, the authors observe a sharp decrease in overall network bandwidth, essentially caused by the concepts of hidden and exposed terminals.

The interested reader is referred to [1] which provides an in-depth discussions on important aspects such as collaborative in-network processing, multimedia sensor hardware, and cross-layer design.

3. A Standards-Based Protocol Stack

Major standardization bodies such as the IEEE and the IETF are finalizing standards for Wireless Sensor Networks. Fig. 2 depicts the protocol stack we believe will equip the WSNs of tomorrow. It is based on the IEEE802.15.4 PHY layer, and is composed of IEEE802.15.4e Time Synchronized Channel Hopping at the MAC layer, and IETF "Internet" standards at upper layers. IETF 6LoWPAN is the adaptation layer used to compact long IPv6 headers into short IEEE802.15.4 frames.

4	transport	IETF	UDP
3	routing		RPL-*like*
	adaptation		6LoWPAN
2	medium access	IEEE	802.15.4e
1	PHY		802.15.4-2006

Fig. 2: The OpenWSN standards-based protocol stack.

Note that all of the standards in Fig. 2 – with the exception of UDP – are in the process of being finalized. We have implemented this stack with TinyOS on the TelosB platform as part of Berkeley's OpenWSN project[3].

In IEEE802.15.4e, nodes are synchronized on a common time slotted structure. Slots are grouped into a superframe of length L slots, each slot having a duration d; the slotframe constantly repeats over time. A slot is long enough for a node to transmit a packet to the next hop, and for the next hop node to acknowledge correct reception; a retransmission policy is invoked when no acknowledgment is received. Each transmission can happen on any of the 16 available frequencies on the 2.4GHz band. A scheduling algorithm is used to assign each of the 16 × L cells to pairs of neighbor nodes.

We tune the IEEE802.15.4e scheduling algorithm to enable video transmission. Fig. 3 depicts the resulting IEEE802.15.4e schedule, which consists of a superframe of length L = 5. The ADV slot is required by IEEE802.15.4e to exchange advertisement packets for neighbor discovery and to keep the network synchronized when no data is exchanged. The SERIAL slots are used for the source nodes to send/receive image data and status information over the serial port (for debugging purposes). TXRX slots are used to exchange the actual video data. When a node receives a packet in slot 1 (resp. 2), it retransmits it in slot 2 (resp. 1).

ADV	TXRX	TXRX	SERIAL	SERIAL
0	1	2	3	4

Fig. 3: The IEEE802.15.4e superframe organization used. L = 5 slots and d = 30ms; the superframe continuously repeats over time.

A hash function is used to translate a node's MAC address into one of the 16 available IEEE802.15.4 channels. A node listens to its own channel and transmits on the channel of the next hop's node. When it has nothing to send in a TXRX slot, a node listens.

Fig. 4 depicts a topology similar to the one used experimentally, and which we use here to illustrate the schedule. Node *B* is the destination; a node's channel is the rank of its identifier in the alphabet. The RPL routing protocol identifies the most efficient route from *C* to *B* to pass through *I* and *A* (details about RPL can be found in [11]). The schedule executes as follows. *C* chooses to send a packet to *I* on slot 2. Node *I* then retransmits that data to *A* on slot 1 using channel 1. As a result, *A* can send to *B* while *C* send to *I* (*I* is not exposed as transmission happen on a different channel), and *B* receives a packet every slotframe. Note that, as illustrated in Fig. 1, a single image consists of multiple packets.

---

[3] The open-source code and detailed documentation on the different standards is available at http://openwsn.berkeley.edu/.

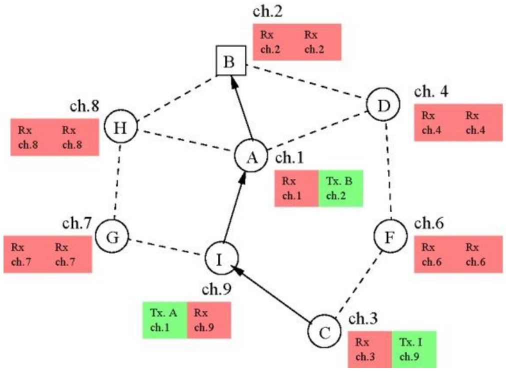


Fig. 4: An example topology, arrows indicate the multihop path from *C* to. Every node is depicted with the activity it has in its TXRX slots.

4. Experimental Setup

We show experimentally how a standards-based protocol stack can be tuned to efficiently transmit video, without requiring a paradigm shift. We strongly believe that the cornerstone to tackling this challenge is the Medium Access Control (MAC) protocol.

Similar to Fig. 4, we deploy a network of 8 TelosB motes running the OpenWSN stack. The nodes form a multi-hop network of diameter 3 hops. We attach a computer to the sink node (node *B* in Fig. 4) to display the received images. Another laptop equipped with a webcam is used to generate images at any of the other nodes in the network. Video is successfully transmitted from any node in the network to the sink node.

By default, we transmit 160x120 pixel gray-scale JPEG images compressed at a 30% compression quality. This yields an average image size of 2.1kB, the equivalent of around 20 packets. The sink receives one packet every superframe, which translates in an image every 2-3 seconds (.33-.5fps).

The energy consumption of a communicating node can be well represented by the ratio of time the radio is on; we call this the network's duty cycle. Table 1 shows the radio on-time for every type of slot. When a node sits idle (it is neither generating nor relaying video), the average duty cycle is 5.8%. On a TelosB mote (which is powered by a pair of 2400mAh AA batteries and which consumes 81mW when the radio is on), this translates in an average node lifetime of 64 days. A node which relays video (it receives in one of its TXRX slots and transmits in the other), the average duty cycle is 12.9%, which translates into an average lifetime of 29 days.

Table 1: Radio On-Time as a Function Slot Type.

Type of Slot	Transmitter	Receiver
ADV	4.76ms on average	
TXRX (w. communication)	6.86ms	7.66ms
TXRX (w.o. communication)	0.00ms	2.00ms
SERIAL	0.00ms	0.00ms

Latency depends on the number of hops. In one superframe, a mote receives a packet and transmits one. This means that, on average, it takes a packet half the duration of the superframe to travel one hop. Assuming a 2kB image, it is composed of 20 packets. The latency between the moment an image is taken to the moment it is displayed at the receiver is the time to transmit 20 frames ($20 \cdot L \cdot d$), but the time it takes the last packet to travel over 3 hops ($3/2 \cdot L \cdot d$), or 3.2s.

Table 2: Number of Dropped and Corrupted Frames as a Function of Image Resolution.

Number of Hops	Image Resolution	Frames Transmitted	Frames Dropped	Frames Corrupted
1	160 × 120	100	1	7
1	192 × 144	100	1	3
1	224 × 168	100	4	10

Finally, Table 2 shows the percentage of dropped and corrupted frames as a function of the image resolution.

5. Open Challenges

In this letter we have addressed the topic of Video over Wireless Sensor Networks from a practical perspective. With a fully standard-based network stack, rates up to .5fps were observed along with a success rate above 90%, in a multi-hop environment. It was mainly the TSCH characteristic of the network that allowed us to address issues such as resource constraints and limited channel capacity. Using multiple channels increased the overall bandwidth while making the links more robust.

Many features of the network stack can be improved and are deemed as open challenges. Starting with the physical layer, recent IEEE802.15.4 radios offer a higher 2Mbps data rate. This improvement would clearly allow lower latency and higher image resolutions. Unfortunately, higher data rates are not part of the IEEE802.15.4 standard.

Moving up the stack, it should be noted that we have used a standard routing protocol. It could be imagined that, in a dense network, using two (or more) disjoint paths could double the bandwidth of the network (the destination node could receive a packet at every TXRX slot).

Disjoint paths call for a transport protocol capable of handling out-of-order delivery of packets, while end-to-end reliability would result in no frame loss at all. There is a clear need for a transport protocol for WSNs, which could also ensure that MAC resource allocation is performed according to application and transport layer requirements.

Finally, and looking at the application itself, it is noticed that inter-frame compression would make more sense than intra-frame compression since the video sequences in question are mostly static with bursts in changes. It would therefore be preferable to try to identify and compress what changes and later transmit it in order to conserve bandwidth and reduce delay.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “A Survey on Wireless Multimedia Sensor Networks,” *ACM International Journal of Computer and Telecommunications Networking*, vol. 51, no. 4, pp. 921–960, 2007.

[2] M. Rahimi, R. Baer, O. i. Iroezi, J. C. Garcia, J. Warrior, D. Estrin, and M. Srivastava, “Cyclops: In Situ Image Sensing and Interpretation in Wireless Sensor Networks,” in *Conference on Embedded Networked Sensor Systems (SenSys)*, San Diego, CA, November 2005.

[3] P. W.-C. Chen, P. Ahammad, C. Boyer, S.-I. Huang, L. Lin, E. J. Lobaton, M. L. Meingast, S. Oh, S. Wang, P. Yan, A. Yang, C. Yeo, L.-C. Chang, D. Tygar, and S. S. Sastry, “CITRIC: A Lowbandwidth Wireless Camera Network Platform,” in *ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*, Stanford University, California, USA, 7-11 September 2008.

[4] A. Bachir, M. Dohler, T. Watteyne, and K. K. Leung, “MAC Essentials for Wireless Sensor Networks,” *IEEE Communications Surveys and Tutorials*, vol. to appear., p. to appear., August 2009.

[5] P. Kulkarni, G. Deepak, P. Shenoy, and Q. Lu, “SensEye: A Multitier Camera Sensor Network,” in *ACM International Conference on Multimedia*, Singapore, 6-11 November 2005, pp. 229–238.

[6] T. Watteyne, A. Mehta, and K. Pister, “Reliability Through Frequency Diversity: Why Channel Hopping Makes Sense,” in *6th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN)*, Tenerife, Canary Islands, Spain, 26-30 October 2009.

[7] M. Chen, V. C. M. Leung, S. Mao, and Y. Yuan, “Directional Geographical Routing for Real-Time Video Communications in Wireless Sensor Networks,” *ACM Computer Communications*, vol. 30,

no. 17, pp. 3368–3383, November 2007.

[8] Z. He and D.Wu, “Resource Allocation and Performance Analysis of Wireless Video Sensors,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 5, pp. 590–599, May 2006.

[9] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, “Distributed Video Coding in Wireless Sensor Networks,” *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 94–106, July 2006.

[10] A. Zainaldin, I. Lambadaris, and B. Nandy, “Adaptive Rate Control Low bit-rate Video Transmission overWireless Zigbee Networks,” in *Internationl Conference on Communications (ICC)*. Beijing, China: IEEE, 19-23 May 2008.

[11] *RPL: IPv6 Routing Protocol for Low power and Lossy Networks*, ROLL IETF Internet-Draft, 8 March 2010, draft-ietf-roll-rpl-07 [work in progress].

Thomas Watteyne is a postdoctoral researcher at the Berkeley Sensor and Actuator Center, University of California in Berkeley, working with Prof. Kristofer S.J. Pister on reliable low-power communication for Wireless Sensor Networks. From October 2005 to September 2008, he was a research engineer at France Telecom R&D/Orange Labs working on Energy Efficient Self-Organizing Wireless Sensor Networks, together with the CITI Laboratory, Lyon, France. At that time, he was also member of the Student Activity Committee and Electronic Communications Coordinator of IEEE Region 8 (Europe, Africa, Middle-East and Russia). He obtained his PhD in Computer Science (2008) and MSc in Telecommunications (2005) from INSA Lyon, France. He has published several journal and conference papers, holds two patents, has contributed to two books and participated in standardization activities. He has been TPC member and member of the organizing committee of various conferences.

Fabien Chraim earned a Bachelor’s degree with honors in Electrical and Computer Engineering from the American University of Beirut in 2009. He is expected to graduate with an M.S. degree in Civil Systems Engineering from the University of California at Berkeley in 2010 after which he will join the PhD program in Electrical Engineering and Computer Sciences. He is a student researcher working with Prof. Kristofer S.J. Pister at the Berkeley Sensor and Actuator Center since January 2010.

Nahir Sarmicanic is an undergraduate in Electrical Engineering and Computer Sciences at the University of California, Berkeley, working with Prof. Kristofer S.J. Pister at the Berkeley Sensor and Actuator Center. His work there involves integrating wireless motes with cell phones. His research interests range from embedded systems to integrated circuits. He is expected to graduate in May 2010.

Chris Jian is an undergraduate in Electrical Engineering and Computer Sciences at the University of California, Berkeley, working with Prof. Kristofer S.J. Pister at the Berkeley Sensor and Actuator Center. His work there involves writing firmware to interface with 802.15.4 radios as well as wireless mote drivers. His research interests include micro-autonomous robots and MEMS devices. He is expected to graduate in May 2010.

Kristofer S.J. Pister received the B.A. degree in applied physics from the University of California, San Diego, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1989 and 1992. From 1992 to 1997, he was an Assistant Professor of electrical engineering with the University of California, Los Angeles. In 1997, he joined the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, where he is currently a Professor and a Co-Director of the Berkeley Sensor and Actuator Center. He coined the term Smart Dust and pioneered the development of ubiquitous networks of communicating sensors, a concept that has since become a virtual sector of technology R&D. During 2003 and 2004, he was on industrial leave as CEO and then CTO of Dust Networks, a company that he co-founded to commercialize low power wireless mesh networking for sensors. In addition to wireless sensor networking, his research interests include MEMS-based micro-robotics and low power circuit design.

Current And Future Open Research Issues On Multimedia Over Embedded Systems

Ilias Politis, University of Piraeus, Greece
Tasos Dagiuklas, Technological Educational Institute (TEI) of Mesolonghi, Greece
ipolitis@ece.upatras.gr

The dominant popularity of multimedia applications in wireless and wired communications, has led to a rapid transition of the multimedia service delivery from desktop PCs to mobile platforms such as PDAs, smartphones and multimedia sensors. As a result, multimedia devices have become one of the most growing areas in the embedded market [1]. Hence, it is hard to imagine the development of consumer electronics without various multimedia capabilities.

Nevertheless, multimedia applications incorporate a number of different algorithms and mechanisms that require high computational complexity with heavy memory access [2]. Such algorithms include optimized encoding schemes that achieve high compression rates with minimum distortion, intelligent packetization and transmission techniques that ensure QoS levels and service continuity over error prone communication channels, and transmission error detection and correction mechanisms for increasing the perceived QoS Therefore, the need for high performance embedded processors is increasing rapidly, simultaneously with widely spreading mobile devices. However, aiming at simply increasing the performance can be trivial when dealing with small wireless multimedia devices, due to important limitations in computational power, memory are, battery life, miniaturization, real-time processing and so on.

The research community will need to focus on addressing new challenges in order to open the way for enhancing our communication experiences, boosting individual and social creativity and productivity in a Future Media Internet environment and ultimately changing the way we live [3]. The issues that need to be addressed include among others: Scalable multimedia compression, transmission, concealment, Network coding and streaming, Content & context fusion for improved multimedia access (seamless and secure), Immersive multimedia experiences, Multimedia, multimodal & deformable objects search, Content with memory and finally, behavior and Power-awareness [4].

In the converged networking era, there is heterogeneity in terms of wireless network capabilities (i.e. network bandwidth, protocol support, capabilities-seamless mobility, security), diversity in end-user devices (i.e. PDAs, cell phones, digital cameras) and diversity of content media formats (i.e. MPEG-1, MPEG-2, H.264 AVC/SVC). In this environment, future research challenges should be addressed in the following directions:

1. New methods of design in order to achieve the Optimisation, Parallelism, Complexity/Design Productivity and Reconfigurability. There are drawbacks at Reference Software specifications in current Video Coding source projects (i.e. MPEG-4, H.264/AVC, etc). Such software tools comprise large non-optimized Software packages that have developed by many contributors. The vast majority of specifications followed the monolithic approach so that there is a insufficiency in terms of using these tools in different application environments. Optimization can be accomplished by using H/W S/W Co-Design methods where complex functional blocks of the encoder (i.e. motion estimation, Inverse Transform, de-block filtering) can be implemented in hardware and the rest functional blocks can be implemented in software [4]. Parallelism can be applied bys using Multiprocessor System-on-Chip as core process unit to carry out video-encoder tasks. Finally, reconfigurability aims to provide ”higher levels” description formalism to speed-up verification and development of new standards and increase the portability of implementations based on common blocks and interfaces.

2. Power-aware systems: In the converged networking era, there is heterogeneity in terms of wireless network capabilities (i.e. network bandwidth, protocol support), diversity in end-user devices (i.e. PDAs, cell phones, digital cameras) and diversity of content media formats (i.e. MPEG-1, MPEG-2, H.264 AVC/SVC etc). In this heterogeneous environment, power-awareness is a challenging task [5].

- Affected by the following factors: non-linear behaviour of battery discharge, strict QoS requirements in terms of delay, jitter and packet loss, the dynamic nature of wireless network conditions and the mobile activity of the end-user.

- Hardware solutions can achieve additional power savings by employing clock gating, a power-saving technique used in system-on-chip designs. Clocks can be deactivated for functions when they are not required, thereby reducing their power consumption. Implemented during synthesis, automatic clock gating can achieve up to a 70% reduction in the overall power consumption. Although extremely effective, it cannot reduce the dynamic power consumption of unused blocks to absolute zero because it is unable to de-activate the clock for all of the functions within the block.

- Regulate battery according to the encoder-decoder functionality complexity (e.g. DCT, entropy coding, motion estimation), adapt on the fly encoding parameters according to network conditions, terminal capabilities (i.e. screen size, media content format support).

In conclusion, this paper outlines some of the key issues regarding the optimization of embedded systems in order to support efficiently real time multimedia applications. As new multimedia services become available to consumers and new technologies constantly emerging, this research will continue to vibrant and active.

References

[1] Ian F. Akyildiz, Tommaso Melodia and Kaushik R. Chowdhury, "A survey on multimedia sensor networks,"Elsevier Computer Networks, vol. 51, no. 4, 2007.
[2] Chang W. Chen and Zhihai He, "Signal processing challenges in next generation multimedia communications," China Communications, vol. 4 no. 5, Oct. 2006.
[3] Future Media Internet-Task Force, "Future media internet research challenges and the road ahead," April 2010.
[4] Liu Yan, Li Renfa, Xu Cheng and Yu Fei, "HW-SW framework for multimedia applications on MPSoC:practice and experience," Journal of Computers, vol. 4, no. 3, March 2009.
[5] R. Vidhyapriya and P.Vanathi, "Energy efficient data compression in wireless sensor networks," International Arab Journal of Information Technology, vol. 6, no. 3, July 2009.

Ilias Politis received his BSc in Electronic Engineering from the Queen Marry College London in 2000, his MSc in Mobile and Personal Communications from King's College London in 2001 and his PhD in Multimedia Communications from University of Patras Greece in 2009. His is currently an Adjunct Lecturer at the Dept. of Computer Science of University of Piraeus, Greece, Scientific Associate with the Dept. of Telecommunications Systems and Networks, at Techological and Educational Institute of Messolonghi, Greece and a post-doc research associate at University of Patras, Greece. Dr. Politis research interests include multimedia communications with emphasis on video transmission optimisation, video coding and heterogeneous networking. His is a member of the IEEE, FITCE and the Technical Chamber of Greece.

Tasos Dagiuklas (www.tesyd.teimes.gr/cones) received the Engineering Degree from the University of Patras-Greece in 1989, the M.Sc. from the University of Manchester-UK in 1991 and the Ph.D. from the University of Essex-UK in 1995, all in Electrical Engineering. Currently, he is employed as Assistant Professor at the Department of Telecommunications Systems and Net- works, Technological Educational Institute (TEI) of Mesolonghi, Greece. He is the Leader of the Converged Networks and Services Research Group. He is also Senior Research Associate within the Wireless Telecommunications Laboratory of the Electrical and Computer Engineering Department at the University of Patras, Greece. Past Positions include teaching Staff at the University of Aegean, Department of Information and Communications Systems Engineering, Greece, senior posts at INTRACOM and OTE, Greece. He has been involved in several EC R&D Research Projects under FP5, FP6 and FP7 research frameworks, in the fields of All-IP network and next generation services. Currently, he is the Technical Manager of the FP7-ICT-PEACE project. He was the Conference General Chair of the international conference, Mobile Multimedia 2007 (ACM Mobimedia 2007), Technical Co- Chair of MMNS Conference of MANWEEK 2008, IMS Workshop Chair as part of ACM Mobimedia 2008 and Workshop Chair for ACM Mobimedia 2009. He has served as TPC member to more than 30 international conferences. His research interests include Future Internet architectures and converged multimedia services over fixed-mobile networks. Dr Dagiuklas has published more than 80 papers at international journals, conferences and standardisation for a in the above fields. He is a member of IEEE and Technical Chamber of Greece.

TECHNOLOGY ADVANCES

Network Coding: Enabling the Multimedia Wireless Internet

Marie-José Montpetit and Muriel Médard (IEEE Fellow), MIT Research Lab for Electronics, USA
{mariejo,medard}@mit.edu

1. Introduction

Network coding (NC) is rapidly migrating from concepts and simulations into software and embedded implementations. This evolution is driven by the realization that network coding offers a solution to improve throughput on congested networks. There is ample published work on NC that has shown NC achieves maximum throughput (min-cut) on a network path by using re-encoding at the intermediate nodes. The phenomenal growth of the mobile Internet and its increasing use for multimedia traffic specifically video traffic is another driver for the move of NC to technology. NC enables more efficient distribution on topologies with peer-to-peer (P2P) with user experience as will be seen in the rest of this paper. While some of the aspects of NC were addressed in the March 2010 Multimedia Communications Technical Committee e-newsletter [1] In this short paper we intend to show that Network Coding will be at the center of the mobile content revolution and propose a few examples of potential implementations.

2. Network Coding Basics

NC is a suite of techniques that allows different strategies across different network topologies, applications and multiple physical media. There are implementations of NC from the physical layer [2] to the IP layer [3]. Network coded packet flows use bottleneck resources for data downloads and dissemination and enable hop-by-hop stateless control of these flows in core and edge networks. With the rise of the Multimedia-rich Wireless Internet, NC is increasingly seen as providing solution to better serve better the new and future network traffic requirements this with manageable added complexity, given the performance of current CPU.

While a number of papers in [1] defined what network coding is, it worth restating what are its salient features. Figure 1 shows a common but very simplified wireless network with sources S in the base stations (BTS) and the destinations D in a smartphone and a netbook. The figure also shows a relay R that could either be another BTS or, in a P2P topology, any end user device. What Figure 1 illustrates is the traditional problem of scheduling and routing in relay and multi-hop networks: any intermediate node becomes a bottleneck. NC provides a mechanism by which the data is combined at the relay: thus no intermediate data queuing or keeping of states is necessary in order to reach all destinations. This reduces delay and operational complexity as well as augmenting throughput in streaming applications. NC achieves this by considering data as algebraic information that can be combined and multiplied using Galois Field arithmetic. By using random linear coefficients [4], essentially a random code, this provides a simple encoding mechanism for data dissemination at the expense of some decoding complexity at the destination. And while Figure 1 shows a traditional *butterfly* network it is easy to generalize it to any multi-hop network, unicast or multicast.

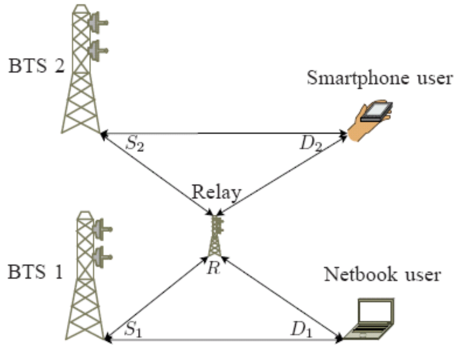


Fig. 1: Simplified Topology for Wireless Network Coding

Our group has contributed to a considerable amount of published work showing how NC can provide invaluable features in wireless networking [4-9]. Important results include error and erasure resilience, reduced dissemination time and the property to achieve min-cut (essentially the highest achievable rate) on these networks. Some of these results are detailed in the next section.

3. Why Network Coding?

large efforts has been dedicated to architecting the next generation wireless networks and defining its services, until recently little had been done to address the performance aspects of pushing huge amounts of data on bandwidth limited and noisy networks. This is changing partly because of the CISCO mobile video predictions [10] of 66 times more video traffic on wireless networks in the next few years. Increasingly, the users are mobile or nomadic and they watch multimedia content on an ever-growing ecosystem of computers and portables, mobile and fixed. The new *social networking* requirements linking people and devices necessitate the new networking paradigms [11-12] that are provided by NC.

Recent technological advances in embedded networked devices help push the NC from the core to the edge of the network and into the end devices, even when accounting for the NC's encoding and decoding complexities. Ultimately this approach could be implemented in current networks nodes as well as in user devices from routers to gateways to PC and set-top boxes. NC strategies are not monolithic and can adapt to the specifics of the network topology and the device ecosystem as will be seen in the next sections.

Downloads

Web surfing, Video on Demand as well as video streaming on the Internet rely on file transfer. Mechanisms like progressive downloads have been designed to compensate for routing delays that can impair the user experience. Yet, little has been done to compensate for erasures of packets due to for example, wireless network limitations like dead spots. Traditionally, erasure channel codes such as Reed-Solomon are used to recover from erasures. The decoding delays associated with these codes can however be very large, as whole blocks of information need to be received before decoding. In addition, source-based solutions burden a network along a route with transmitting redundant information when only the edges requires added reliability; this can contribute to more congestion that will further impact user experience.

NC can recover packets efficiently since the *lost* packet is part of a linear combination of transmitted packets. As can be seen in Figure 2 [7], when comparing current WIFI transmission to network coded ones, the time for completion of a file transfer when the packets are coded is significantly reduced especially at the lower rates that are representative of many wireless networks.

Overcoming IP Protocol Limitations

TCP, the underlying protocol below HTTP and many streaming protocols, uses feedback to provide rate and congestion control by acknowledging (ACK) received packets. The ACKs are used to control the transmission window size and when they are lost the protocol assumes it is because of congestion. The automatic response is to reduce the window size and use a *slow start* mechanism to reduce the number of sent packets. While this is appropriate for low loss networks it does not reflect the characteristics of wireless networks where losses are more likely due to noise than to congestion.

But it would appear that NC solutions are inappropriate for TCP-based protocols since decoding will delay the TCP acknowledgements and mislead the source into congestion avoidance. The work presented in [3] has shown that by encoding and decoding groups of packets in a progressive manner and introducing a concept of *seen* packet for fast acknowledgement, TCP throughput can be improved considerably in a noisy multi-hop network. The *TCP/NC* inserts a layer of network coding between TCP and IP. In this scheme, the random linear NC masks link losses from TCP. Lost packets can be recuperated from the coded information. A packet can be ACK'ed even if partially *revealed* (decoded) hence keeping a steady flow of ACKs and operating TCP under optimal conditions. Moreover, the TCP/NC does not need to be implemented end to end but only at certain nodes to profit from the throughput gains.

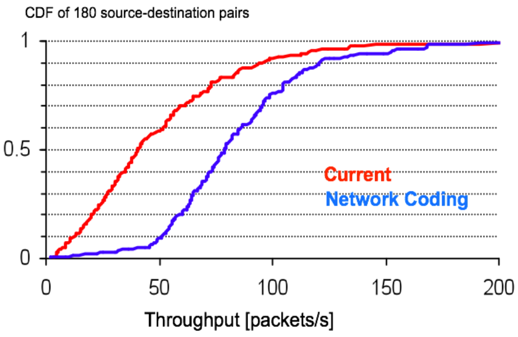


Fig. 2: Without NC at 50 packets/s, 50% of transfers take more than 5 minutes [7].

As shown in Figure 3, this approach results in higher goodput (the measure of delivered packets to the application) when the loss rate is high hence this mechanism offers much promise for the wireless and the peer-to-peer Internet.

FTP simulations over a very lossy medium (Figure 4) show that TCP NC (orange and green lines) outperforms TCP (pink line) and that the in-network intermediary re-encoding (orange line) outperforms the end-to-end operations in terms of overall throughput. Figure 4 also highlights that with TCP/NC there are fewer interruptions which leads to a better user experience in multimedia applications.

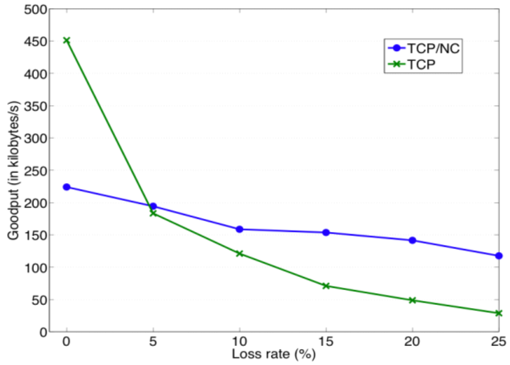


Fig. 3: Goodput vs. Loss rate - TCP and TCP/NC [3]

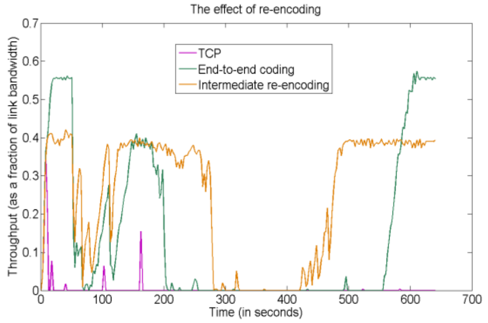


Fig. 4: TCP/NC Simulation over a lossy medium [6]

Layered Approaches

The solutions presented above rely on a node to mix packets from a single stream. Recent work has investigated coding across multi-resolution streams [13] typical of some recent codecs like Scalable Video Coding (SVC) or of multimedia traffic. The goal is to serve a heterogeneous receiver ecosystem by using *pushback* mechanisms about the number of layers a receiver can decode. The pushback information carries the maximum number of layers, the receiver's min-cut, based on a receiver instantaneous condition. While the results presented in [13] are for a multicast network, the work also applies to the unicast case. In a real system the pushback information could be carried over device discovery mechanisms or MAC layer acknowledgements.

Figure 5 shows some of the results presented in [13] with a network of 25 nodes and 12 receivers. *Happy Nodes* are those who achieve min-cut hence received their requested layers from the source. The blue curve represents the non-coded end-to-end solution that considers each multicast paths as unicast, the brown curve is from a Steiner tree (optimal routing) and the black curve is the per-layer random network coding. The red curve implements one version of the inter-layer coding where the information pushed back to the a node is the minimum number of requested layers over its multicast tree, driven by the worst node. Finally the pink curve represents the inter-layer encoding where the encoding is done over a number of layers determined by sub-trees.

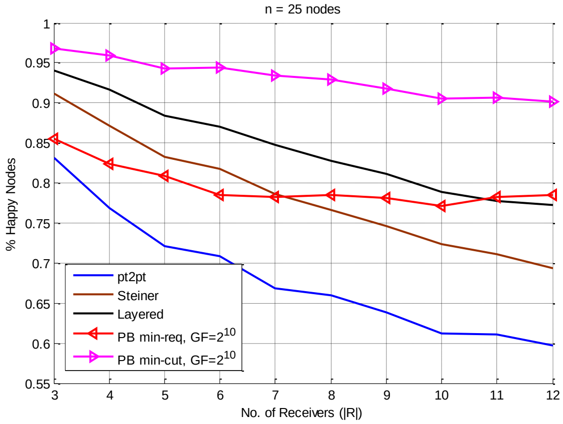


Fig. 5: Cross-layer NC for multi-resolution multicast. [6]

Figure 5 shows that Network Coding (black and pink) out performs the non-coded versions all the time but that the use of an adequate cross-layering (pink) algorithm increases the performance further. This result is encouraging as more and more wireless networks are used to transmit video conferencing or other converged applications that can be considered layered (video with user feedback or personalized messaging for example).

Conclusion

This paper introduced some of the NC solutions that we are proposing to improve wireless networks performance and enable next generation multimedia services. By better using bandwidth resources and allocate these resources only where they are needed, the challenges of high rate and video-rich applications can be met. Our future developments will address *distributed virtualized* *nodes* combining peer to peer transmission and local storage as well as *network composition* using NC to combine heterogeneous network features for higher performance and lower cost.

The growing wireless and mobile ecosystem challenges current networks and needs new paradigms for the provision of network services. We believe that Network Coding will provide the network support the new *Wireless Multimedia Internet* requires.

References

[1] Comsoc Multimedia Communications Technical Committee e-Letter, Special Issue on Network Coding, March 2010. Available from http://www.comsoc.org/~mmc

[2] Katabi, D., Rahul, H., Jakubczak, S. ”Softcast: one video to serve all video receivers”, MIT Technical Report available from http://hdl.handle.net/1721.1/44585.

[3] J.K. Sundararajan, S. Devavrat, M. Médard, M. Mitzenmacher and J. Barros, "Network coding meets TCP", Proceedings of IEEE INFOCOM 2009, Rio de Janeiro, Brazil, April 2009, pp. 280-288.

[4] S. Deb, M. Médard, C. Choute, “On Random Network Coding Based Information Dissemination”, International Symposium on Information Theory, September 2005.

[5] D.S. Lun, T. Ho, N, Ratnakar, R. Koetter, R., and M. Médard, “Network Coding in Wireless Networks -A survey of techniques for efficient operation of coded wireless packet networks,” Cooperation in Wireless Communications: Principles and Applications, Springer, Editors: F. Fitzek and M. Katz, 2007.

[6] A. Eryilmaz, A. Ozdaglar and M. Médard, M., “On the Delay and Throughput Gains of Coding in Unreliable Networks”, *IEEE Transactions on Information Theory*, Volume 54, Issue 12, December 2008, pp:5511 - 5524.

[7] D. Katabi, C. Fragouli, A. Markopoulou, H. Rahul and M. Médard, “Wireless Network Coding: Opportunities and Challenges”, MILCOM, October 2007

[8] D.S. Lun, M. Médard, R. Koetter, “Network Coding for Efficient Wireless Unicast,” invited paper, IEEE International Zurich Seminar on Communications, pp. 74-77, February 2006.

[9] M.J. Montpetit and Muriel Médard, "Video-centric Network Coding Strategies for 4G Wireless Networks: An Overview", Proceedings of CCNC'10, Las Vegas January 2010.

[10] [CISCO 2009] CISCO. Cisco visual networking index: Forecast and methodology, 20082013. White paper. CISCO. June 2009. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf

[11] MIT, “The Power of Random”, http://web.mit.edu/newsoffice/2010/network-coding-part1.html

[12] MIT. “Rethinking Networking”, http://web.mit.edu/newsoffice/2010/network-coding-part2.html

[13] MinJi Kim, Daniel Lucani, Xiaomeng Shi, Fang Zhao and Muriel Médard “Network Coding for Multiresolution Multicast”, Infocom 2010, March 2010.

Muriel Médard is a Professor in the Electrical Engineering and Computer Science at MIT. She was previously an Assistant Professor in the Electrical and Computer Engineering Department and a member of the Coordinated Science Laboratory at the University of Illinois Urbana-Champaign. From 1995 to 1998, she was a Staff Member at MIT Lincoln Laboratory in the Optical Communications and the Advanced Networking Groups. Professor Médard received B.S. degrees in EECS and in Mathematics in 1989, a B.S. degree in Humanities in 1990, a M.S. degree in EE 1991, and a Sc D. degree in EE in 1995, all from the Massachusetts Institute of Technology (MIT), Cambridge. She has served as an Associate Editor for the Optical Communications and Networking Series of the IEEE Journal on Selected Areas in Communications, as an Associate Editor in Communications for the IEEE Transactions on Information Theory and as an Associate Editor for the OSA Journal of Optical Networking. She has served as a Guest Editor for the IEEE Journal of Lightwave Technology, the Joint special issue of the IEEE Transactions on Information Theory and the IEEE/ACM Transactions on Networking on Networking and Information Theory and the IEEE Transactions on Information Forensic and Security: Special Issue on Statistical Methods for Network Security and Forensics. She serves as an associate editor for the IEEE/OSA Journal of Lightwave Technology. She is a member of the Board of Governors of the IEEE Information Theory Society and currently Second Vice-President. Professor Médard's research interests are in the areas of network coding and reliable communications, particularly for optical and wireless networks. She was awarded the 2009 Communication Society and Information Theory Society Joint Paper Award, the 2009 William R. Bennett Prize in the Field of Communications, the IEEE Leon K. Kirchmayer Prize Paper Award 2002 and the Best Paper Award for G. Weichenberg, V. Chan, M. Médard. She received a NSF Career Award in 2001 and was co-winner 2004 Harold E. Edgerton Faculty Achievement Award, established in 1982 to honor junior faculty members "for distinction in research, teaching and service to the MIT community." In 2007 she was named a Gilbreth Lecturer by the National Academy of Engineering.

Marie-José Montpetit is an invited scientist in the Electrical Engineering and Computer Science at MIT focusing on network coding for video transmission. She was previously an invited scientist at the MIT Media Laboratory where she is still involved in a class on converged video applications. Her work on converged and social TV was highlighted in the May-June 2010 MIT Technology Review as one of 10 technologies that will change the world. Dr. Montpetit received a Ph.D. in EECS from the Ecole Polytechnique in Montreal, Canada. She is a member of the IEEE Standing Committee on DSP and a collaborator to the ETSI BSM working group on aspects of convergence. She was the recipient of the Motorola Innovation Prize in 2007 for the development of a multi-screen and multi-network video mobility system. Her work on converged video applications and multi-screen IPTV has gotten her many invited papers and keynote presentations. She is a reviewer for the European Union for proposal and projects in the wireless networking and future Internet fields as well as the editor of many journals and publications and has served on numerous conference program committees. Dr. Montpetit is a Senior Member of the IEEE.

Video Coding Solutions for VANETs

Mohammed Ghanbari (IEEE Fellow), Martin Fleury and Nadia N. Qadri, University of Essex, Colchester, UK
{ghan,fleum,nnawaz}@essex.ac.uk

1. Introduction

The impending widespread deployment of Vehicular Ad Hoc Networks (VANETs) has created an opportunity for multimedia communication not only as part of vehicle safety provision, traffic management, and emergency response but also in the value-added 'infotainment' domain. The key feature of video distribution to the passengers of vehicles or crew members of emergency vehicles in the VANET is robustness and reliability, as the environment is highly error prone. As detailed below, complex movement patterns of vehicles and non-line-of-sight wireless propagation add to the challenge of multi-hop routing. One response to this challenge is through the error-resilience features of the H.264/AVC (Advanced Video Coding) codec, which together with multipath streaming, including peer-to-peer streaming [1], provide source coding solutions. Application-layer forward error correction (AP-FEC) as a general solution may replicate physical layer channel coding (unless the AP-FEC acts as a concatenated code), while error control through ARQ, apart from the additional latency introduced over multi-hop network paths, is unreliable when network links are constantly being broken.

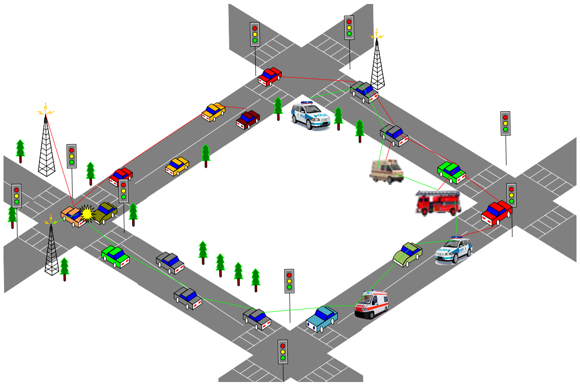
In retrospect, three events may be singled out in the rise of VANET multimedia communication. Firstly in 1999, the US FCC allocated 75 MHz bandwidth of the 5.9 GHz spectrum to Dedicated Short Range Communication (DSRC), essentially for wireless communication between vehicles and from a vehicle to the network infrastructure, normally via roadside units. Secondly, the term Vehicular Ad Hoc Network (VANET) was first applied in the 2004 ACM international workshop of that name and since then academic activity has burgeoned. And thirdly, in 2005 video streaming over VANETs was suggested as a way of reporting traffic congestion and accidents [2], as captured by roadside cameras. In 2005 also, an early feasibility study [3] in Japan was testing streaming video between two vehicles. Imaginative ways of responding to urban emergencies [4] by streaming video to responding vehicles are one of a number of VANET initiatives by Mario Gerla's research group at the University of California, Los Angeles.

Multi-wireless-interface vehicles placed within cellular networks are already proposed [5] as a way to relieve congested cells. As new 'push' multimedia services are introduced into 3G cellular networks, the same services may be extended into VANETs. VANETs may also support the exchange or sharing of personal video clips (as occurs in social networks). Roadside sources of multimedia content [6], possibly linked in a backbone network, can disseminate pre-encoded video or serve to notify the passengers of a passing vehicle of available video sequences in circulation within the VANET.

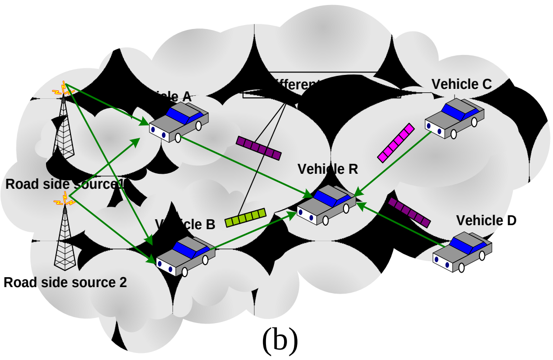
2. VANET Streaming Characteristics

Though video streaming for Mobile Ad Hoc Networks (MANET) has been long investigated, for example [7], there are some important differences [8] between MANETs and VANETs. The mobility model often used in MANETs is random waypoint which is unconstrained either by the presence of buildings that occur in an urban road topology or the linear nature of a highway VANET. High speeds on highways may cause network fragmentation. VanetMobiSim [9] is openly available and includes modeling of driver behavior and deceleration/acceleration of vehicles when overtaking or changing lanes, especially within a city or the suburbs. Though car manufacturers are leading the way with microscopic level simulation modeling [10], there are generic features of vehicle mobility such as the nature of road obstacles including lane closures, uphill gradients, and potholes, which produce the same reactions the world over. Another difference is that in MANET research line-of-sight signal propagation models such as the two-ray ground propagation model are common, whereas it is becoming increasingly obvious that an urban environment will introduce reflections, diffraction and scattering due in part to the presence of urban 'canyons'. IEEE 802.11p based VANETs (there are some Code Division Multiple Access networks proposed) will be able to take advantage of the vehicle as an energy source. This does not mean that the control overhead of streaming across the network can be neglected, as this overhead still acts as a form of congestion, but it does remove a computational and storage restriction that affects MANET devices.

In Fig. 1a, a crash has occurred which can be captured by roadside cameras (on masts in the Figure) or by patrol cars, which commonly are equipped with cameras. Emergency vehicles such as the fire engine and ambulance are shown making their way to the accident.



(a)



(b)

Fig. 1: a) Simplified scenario of VANET multi-path streaming from a crash scene b) in P2P streaming of video clips a receiver conceptually receives four versions of the stream, two from vehicles that have already acquired a version of the video clip and two from vehicles capturing or relaying from roadside units.

An early view of the scene is a way that the responder vehicles can prepare while on route. As roadside, possibly high-rise buildings reflect the signal, propagation is likely to be along the roads and likewise diffraction of the signal will occur round corners. Therefore, it is difficult for emergency vehicles to communicate directly with each other even if they are within range. Consequently, we have proposed that to relay the video stream, these vehicles form a multicast group embedded within a VANET formed by vehicles in the vicinity. Network coding has been introduced [11] as a way of protecting confidentiality in this type of situation but selective encryption of (say) motion vectors is another possibility which does not require action by non-emergency vehicles. The Figure shows two streams following multi-hop paths from the accident. In networking terms, path diversity helps to balance the load but in coding terms there is a natural mapping to Multiple Description Coding (MDC).

In Fig. 1b, P2P streaming is proposed as a solution to another application, when video distribution is less time-critical, as it indeed might be in 'infotainment' applications. Because passing vehicles may not linger sufficiently for a full video sequence to be transferred from a roadside unit, partial storage in any one vehicle may occur. Vehicles with partial video sequences may also later park or leave the vicinity.

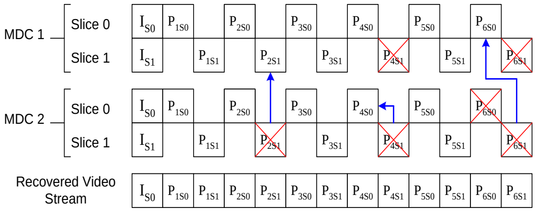


Fig. 2: An example of the proposed slice compensation scheme with MDC and FMO, with arrows indicating the relationship "can be reconstructed from"

Fig. 2 is an example of our P2P slice compensation scheme for MDC. The *same* video stream transported in MDC form is available from two sets of peers (MDC 1 and 2). That is the MDC 1 and 2 streams are duplicates of each other that are transported using MDC and are NOT two descriptions of the same video. Each frame within a video stream (MDC 1 or 2) is further split into two slices (slices 0 and 1) to form two descriptions.

3. Source Coding Options

We have experimented with temporal MDC [12] using a type of Video Redundancy Coding (VRC) scheme [13] without back channel, which avoids the need for the specialist codecs associated with some other forms of MDC. VRC maintains synchronization, which was a problem with an earlier scheme. An issue with VRC is the extraction of independent odd and even frame streams, which reduces the coding efficiency derived from motion estimation. In contrast, H264/AVC's spatial Flexible Macroblock Ordering (FMO) [14] with the dispersed or checkerboard selection macroblock (MB) pattern as a form of MDC is attractive. If one of the descriptions for a frame is lost then the other can be used, as missing MBs can be lost concealed using motion copy from adjacent MBs. Moreover, there is no duplication of data in the two streams, and mapping overhead is only apparent when at very low error rates. The ‘bursty’ output rate associated with distributing large I-frames is a problem for VANETs, because of the sudden influx of packets into the ad hoc network. H.264/AVC provides intra-MB refresh [15], which can be combined with FMO to reduce the latter problem. When intra-MB refresh is operated in row cyclic order, then synchronization occurs after every cycle. Another potential solution to achieving MDC is to employ redundant slices or pictures [16]. In this scheme redundant slices are encoded at a higher compression ratio than the slices they accompany in the stream. This can take place with two ‘redundantly-sliced’ streams of even and odd frames, using MB-refresh to avoid error propagation across the IPPPP…….. Group of Picture (GoP) structure. Compared to other forms of MDC, we found that redundant slices suffered heavier losses under burst error conditions. However, intuitively either the packet losses occur for redundant-slice bearing packets or the reduction in video quality from replacement by redundant slices is not as damaging as the complete loss of some frames.

An insight from the experience with redundant slice MDC is that in a VANET in particular the extra transmission energy consumed is not necessarily a handicap. For instance, with scalable video coding, we have proposed [17] that a redundant or duplicate base layer may be a way to effectively provide MDC. In this scheme, in a counter-intuitive way, the enhancement layer(s) *and* the original base layer are transported in one stream and the redundant base layer in another stream. Each stream follows a different route. In fact, on-going work for the H.264/SVC (Scalable Video Coding) extension will further extract just the key pictures as another stream. (In an 8-picture SVC GoP, the key picture content can be as low as 5% of the total data depending on spatio-temporal configuration, bearing in mind that the base layer only contains these key pictures and is more coarsely encoded than the enhancement layers.) If packets from key picture frames are lost then other packets bearing data from the SVC predictive structure have to be discarded, which without this robustness makes SVC problematic in the highly error-prone VANET environment.

4. Conclusion

However, before any of these robust schemes can be properly validated it is important to utilize mobility models that incorporate all the important factors and to develop propagation models that show how on average multimedia streaming will respond in a VANET environment. To that end we have elaborated an existing ray-tracing model for simulation purposes. This enhanced model can include the distance over which a signal is reflected and add the effect of roadside scattering such as from signs, street ‘furniture’ and foliage. This does mean that expertise in wireless *and* coding is required for VANET multimedia solutions, but in return for this versatility there is a potential rich range of coding possibilities that hopefully this Letter has outlined.

References

[1]. N. Qadri, M. Fleury, M. Altaf, and M. Ghanbari, “Resilient P2P multimedia exchange in a VANET,” in *IEEE/IFIP Wireless Days Conf.*, Dec. 2009.

[2]. M. Guo, M. H. Ammar, and E. W. Zegura, “V3: A vehicle-to-vehicle live video streaming architecture,” in *3rd IEEE Int’l Conf. on Pervasive Computing and Communs.*, Mar. 2005, pp. 171-180.

[3]. P. Bucciol, E. Masala, N. Kawaguchi, K. Takeda, and J. C. de Martin, “Performance evaluation of H.264 video streaming over inter-vehicular 802.11 ad hoc networks,” in *IEEE 16th Int. Symp. on Personal, Indoor and Mobile Radio Comms.*, Sept. 2005, pp. 1936-1940.

[4]. M. Roccetti, M. Gerla, C. E. Palazzi, S. Ferretti and G. Pau, “First responders’ crystal ball: How to scry the emergency from a remote vehicle,” in *IEEE 26th Int’l Conf. on Performance of Computing and Communs.*, April 2007, pp. 556-56.

[5]. H. Wu, C. Qiao, S. De, and O. Tonguz, “Integrated cellular and ad hoc relaying service: iCAR,” *IEEE J. on Sel. Areas in Communs.*, pp. 2105-2113, vol. 19, no. 10, 2001

[6]. C. Lochert, B. Scheuerrmann, C. Wewetzer, A. Luebke, and M. Mauve, “Data aggregation and roadside unit placement for a traffic information system,” *5th ACM Workshop on VehiculAr Inter-NETworking*, Sept. 2008 , pp. 58-65.

[7]. W. Wei and A. Zakhor, “Multipath unicast and multicast video communication over wireless ad hoc networks”, in *Int’l Conf. on Broadband Networks*, Oct. 2004, pp. 496-505.

[8]. N. Qadri, M. Fleury, M. Altaf, and M. Ghanbari, “Multi-source video streaming in a wireless vehicular ad hoc network,” accepted for IET Communications, Special Issue on Video Communications over Wireless Networks, July, 2010.

[9]. M. Fiore, J. Härri, F. Filali, and C. Bonnet, “Vehicular mobility simulation for VANETs,” in *40th Ann. Simulation Symp.*, Mar. 2007, pp. 301-307.

[10]. B. Liu, B. Khorashadi, H. Du, D. Ghosal, C-H. Chuah, and M. Zhang, “VGSim: An integrated networking and microscopic vehicular mobility simulation platform,” *IEEE Communs. Mag.*, vol. 47, no. 5, pp. 134-141, May 2009.

[11]. J.-S. Park, U. Lee, S. Y. Oh, M. Gerla, and D. Lun, “Emergency related video streaming in VANET using network coding,” in *3rd Int’l Workshop on VANETs*, 2006, pp. 102-103.

[12]. N. Qadri, M. Altaf, M. Fleury, and M. Ghanbari, “Robust video communication over an urban VANET,” *Mobile Information Systems*, vol. 6, no. 3, 2010.

[13]. S. Wenger, “Video Redundancy Coding in H.263+”, Packet Video Workshop, 1997.

[14]. P. Lambert, W. de Neve, Y. Dhondt, and R. Van de Walle, “Flexible macroblock ordering in H.264/AVC,” *J. of Visual Commun. and Representation*, vol. 17 , no. 2, pp. 358-378, Apr. 2004

[15]. R.M. Schreier, and A. Rothermel, “Motion adaptive intra refresh for the H.264 video coding standard,” *IEEE Trans. Consumer Electronics*, vol. 52, no. 1, pp. 249 – 253, 2006.

[16]. I. Radulovic, Y-K. Wang, S. Wenger, A. Hallapuro, M. H. Hannuksela, and P. Frossard, “Multiple description H.264 video coding with redundant pictures*”,*in *Int’l Workshop on Mobile Video*, Sept. 2007, pp. 37-42.

[17]. N. Qadri, M. Altaf, M. Fleury, M. Ghanbari, "P2P layered video streaming over wireless ad hoc networks,” *MobiMedia*, article no. 23, 2009.

Mohammed Ghanbari (M’78, SM’96, F’01) is best known for his pioneering work on two-layer video coding for ATM networks (which earned him an IEEE Fellowship in 2001), now known as SNR scalability in the standard video codecs. He has served as an Associate Editor to IEEE Trans. on Multimedia (IEEE-T-MM from 1998-2004) He has registered for eleven international patents on various aspects of video networking and was the co-recipient of A.H. Reeves prize for the best paper published in the 1995 Proc. of IEE on the theme of digital coding. He is the co-author of “Principles of Performance Engineering”, a book published by IET press in 1997, the author of “Video Coding: An Introduction to Standard Codecs”, a book also published by IET press in 1999, which received the year 2000 best book award by the IEE, and the author of “Standard Codecs: Image Compression to Advanced Video Coding” also published by the IET press in 2003. Prof. Ghanbari has authored or co-authored about 500 journal and conference papers, many of which have had a fundamental influence in this field.

Martin Fleury holds a degree in Modern History (Oxford University, UK) and a Maths/Physics based degree from the Open University, Milton Keynes, UK. He obtained an MSc in Astrophysics from QMW College, University of London, UK in 1990 and an MSc from the University of South-West England, Bristol in Parallel Computing Systems in 1991. He holds a PhD in Parallel Image-Processing Systems from the University of Essex, Colchester, UK. He is currently employed as a Senior Lecturer at the University of Essex. Martin has authored or co-authored over one hundred and fifty articles on topics such as document and image compression algorithms, performance prediction of parallel systems, software engineering, reconfigurable hardware and vision systems. His current research interests are video communication over MANS, WLANs, PANs, BANs, MANETs, and VANETs. He has published a book on high-performance computing for low-level image- and signal-processing algorithms and a number of book chapters.

Nadia N. Qadri (Student Member 07) is currently a PhD student at School of Computer Science and Electronics Engineering, University of Essex, UK. She received her Masters of Engineering (Communication Systems and Networks) and Bachelors of Engineering (Computer Systems), from Mehran University of Engineering and Technology, Jamshoro, Pakistan in 2004 and 2002 respectively. She has more than four years of teaching and research experience at renowned universities of Pakistan viz. Mehran University of Engineering & Technology, Fatima Jinnah Womens University and COMSATS Institute of Information Technology. Her research interests include video streaming for mobile ad hoc networks and vehicular ad hoc networks, along with P2P streaming.

Modeling and Resource Allocation for HD Videos over WiMAX Broadband Wireless Networks

Abdel Karim Al Tamimi, Chakchai So-In, and Raj Jain (IEEE Fellow), Washington University in St. Louis, St. Louis, MO
{aa7,cs5,jain}@cse.wustl.edu

1. Introduction

Mobile video is considered a major upcoming application and revenue generator for broadband wireless networks like WiMAX and LTE. Therefore, it is important to design a proper resource allocation scheme for mobile video, since video traffic is both throughput consuming and delay sensitive.

In order to compare resource allocation schemes for mobile video, it is necessary to have an accurate model of the video traffic that represents real mobile videos. For limited-resource networks like WiMAX, it is essential to maximize the resources utilization. An accurate video model can provide the basis for a reliable traffic predictor that is the core component of any dynamic resource allocation scheme.

MPEG video frames are known to have seasonal characteristics. As shown in the Figure 1, MPEG video frames are divided into three types: I, P, and B-frames. These frame types differ in their function and size. For example, I-frames are the largest in size, and B-frames are the smallest in size.

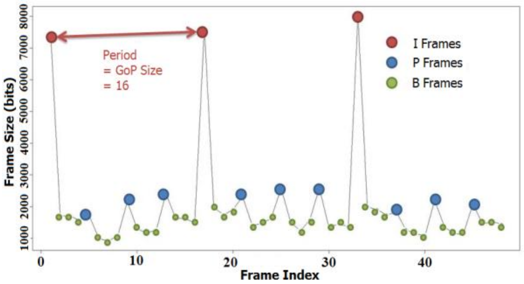


Fig.1: Seasonal Characteristics of MPEG Video

Typically, the pattern of video frames is repeated every "*s*" frames, where *s* is the Group of Picture (GoP) size. This observation justifies our approach to model MPEG videos as a time series.

2. The SAM Model

Our mathematical model is based on the seasonal autoregressive integrated moving average (SARIMA) models [1,2]. SARIMA models aim to achieve better modeling by identifying both non-seasonal and seasonal parts of the data series.

The process of modeling a data series includes two steps: identifying the model order, and then estimating the model coefficients. The first step requires human intervention due to the significant statistical requirements in determining the best model. The second step uses algorithms like maximum likelihood (ML) to estimate the model's coefficients. SARIMA can be described as follows:

$$SARIMA = (p, d, q) \times (P, D, Q)^{s} \qquad (1)$$

where *p* is the order of the autoregressive (AR) part; q is the order of the moving average (MA) part; d is the order of the differencing part. The parameters *P*, *Q*, and *D* are the corresponding seasonal order, respectively. The parameter *s* denotes the seasonality of the series. Our proposed model: Simplified Seasonal ARIMA model (SAM), as a SARIMA model, can be written as follows:

$$SAM = (1,0,1) \times (1,1,1)^{z} \qquad (2)$$

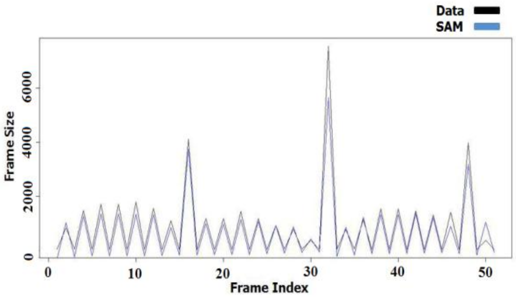
where *z* is the video seasonality, which in most cases is equal to the used GoP size. This simplification means that SAM does not require any human intervention, and needs only 4 coefficients to be estimated. These coefficients are: AR coefficient (φ), MA coefficient (θ), seasonal AR or SAR coefficient (Φs), and seasonal MA or SMA coefficient (Θs). SAM can be written in its difference form as :

$$\begin{aligned} X_t &= X_{t-1} + \varphi X_{t-1} - \varphi X_{t-2} + \Phi_s X_{t-s} - \varphi \Phi_s X_{t-s-1} \\ &- \Phi_s X_{t-s-1} + \varphi \Phi_s X_{t-s-2} - \theta \varepsilon_{t-1} - \Theta_s \varepsilon_{t-s} \\ &+ \theta \Theta_s \varepsilon_{t-s-1} + \varepsilon_t \end{aligned} \qquad (3)$$

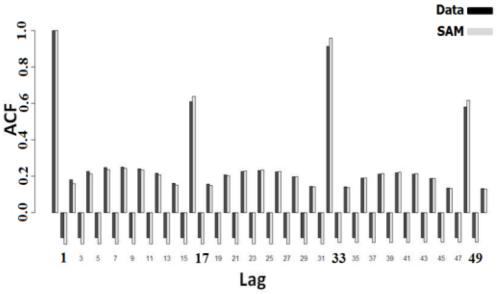
SAM provides a unified approach to model video traces encoded with different video codec standards, using different encoding settings [1-3]. In recent research results [1, 2, 7], it was shown that SAM is capable of capturing the statistical characteristics of video traces within 5% of the optimal model for these video traces.

Figure 2 shows the modeling results of Star Wars IV with an AVC-encoded video trace. The model has been tested against video traces encoded using three different encoding settings and standards: MPEG-Part2, MPEG4-Part10/AVC (H.264), and AVC's scalable extension to support temporal scalability (SVC-TS).

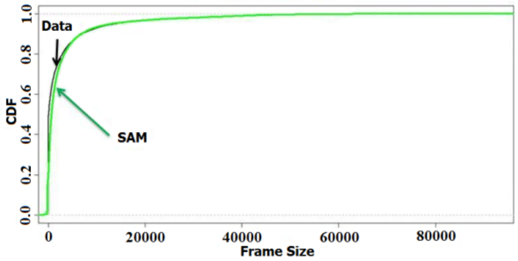
In addition, in [3], we demonstrated that SAM has a clear edge in modeling high definition (HD) video traces over both AR modeling, and the automatic SARIMA estimation algorithm proposed in [4], that implements a unified approach to specify the model's order using a step-wise procedure.



(a) Close-up Comparison



(b) ACF Comparison



(c) CDF Comparison

Fig. 2: SAM Modeling Results for Star Wars IV Movie Trace

Based on our SAM model, we developed a trace generator that can be used to evaluate the performance of video based simulations. The trace generator provides the researchers with the ability to produce user-defined length traces that resemble the desired statistical attributes [2,5].

In [1], we presented the results of our analysis of various scheduling methods for Mobile WiMAX networks using the SAM traffic generator. These are Earliest Deadline First (EDF), Deficit Round Robin (DRR) and Easiest Deadline First with Deficit Round Robin (EDF-DRR). The simulation results for EDF, DRR, and EDF-DRR show that EDF is the most unfair. While EDF-DRR is an improvement, DRR is the most fair and provides the best performance for real-time mobile video traffic.

3. Video Traffic Prediction

SAM, as an accurate source model, can be used to predict future traffic based on the available short-term history of the incoming video frames. Using SAM difference's equation (3), future incoming video frames can be easily forecasted. We demonstrated in [6] that SAM provides 55% improvements on average, in terms of prediction accuracy compared to AR, and 53% improvements over the automatic SARIMA model estimation algorithm [4].

3. HD Video Traces Collection

We collected more than 50 HD video traces from the HD section of YouTube website that represents a wide variety of statistical characteristics. We encoded these traces using AVC codec with the most common settings, confirming experts' recommendations[6]. These traces provide the research community with the means to test and research new methods to optimize network resources. All the video traces are available to the research community through our website [5].

Our modeling and prediction comparisons are based on our HD video traces collection. The comparison results are available to the research community along with our developed tools [5].

In addition, we performed a full statistical analysis on our video traces collection. Our analysis included factor analysis using principle component analysis (PCA), and cluster analysis using *k-means* clustering [3].

4. Conclusions

The SAM model provides a convenient and accurate approach to model, generate and predict video traffic. It may be considered for practical solutions to solve the dynamic resource allocation challenge for live video streams, due to its ability to provide accurate results for the most common video codecs. This is especially important for networks with limited resources like WiMAX and LTE.

References

[1] Al Tamimi, A., So-In, C., and Jain, R. 2010 Modeling and Resource Allocation for Mobile Video over WiMAX Broadband Wireless Networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, Special Issue on Wireless Video Transmission, vol. 28, no. 3, pp. 354-365, April 2010.

[2] Al Tamimi, A., Jain, R., and So-In, C. 2010 Modeling and generation of AVC and SVC-TS mobile video traces for broadband access networks. *In Proceedings of the First Annual ACM SIGMM Conference on Multimedia Systems* (Phoenix, Arizona, USA, February 22-23, 2010). MMSys '10. ACM, New York, NY, 89-98.
DOI= http://doi.acm.org/10.1145/1730836.1730848.

[3] Al Tamimi, A., Jain, R., and So-In, C. 2010 Statistical Analysis and Modeling of High Definition Video Trace. *In Proceedings of the 2010 IEEE international Conference on Multimedia and Expo* (Singapore July 19-23, 2010), in press.

[4] Hyndman, R. and Khandakar, Y. 2008 Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, vol. 27, no. 3, July 2008.

[5] SAM model Traces website.
URL=[http://www.cse.wustl.edu/~jain/sam/index.htm]

[6] Al Tamimi, A., Jain, R., and So-In, C. 2010 Modeling and Prediction of High Definition Video Traffic: A Real-World Case Study. *In Proceedings of the Second International Conferences on Advances in Multimedia*, MMEDIA 2010, in press.

[7] Al Tamimi, A., Jain, R., and So-In, C. 2008 SAM: A Simplified Seasonal ARIMA Model for Mobile Video over Wireless Broadband Networks. *IEEE International Symposium on Multimedia 2008*, pp 178-183.

Abdel Karim Al Tamimi received his Bachelor's degree in Computer Engineering from Yarmouk University, Jordan in 2004. He received his Master's degree from Washington University in St. Louis in 2007. He is a final year PhD candidate in Computer Engineering at Washington Univerisity at St. Louis. His research interests include network systems, multimedia modeling, dynamic resource allocation, traffic engineering, and wireless networks.

Chakchai So-In received the B.Eng. and M.Eng. degrees in computer engineering from Kasetsart University, Bangkok, Thailand in 1999 and 2001 respectively. He is currently working toward a Ph.D. degree at the Department of Computer Science and Engineering, Washington University in St. Louis, MO, USA. His research interests include architectures for future wireless networks/next generation wireless networks; congestion control in high speed networks; protocols to support network and transport mobility, multihoming, and privacy; and quality of service in broadband wireless access networks.

Raj Jain is a Fellow of IEEE, a Fellow of ACM, a winner of ACM SIGCOMM Test of Time award, CDAC-ACCS Foundation Award 2009, and ranks among the top 50 in Citeseer's list of Most Cited Authors in Computer Science. Dr. Jain is currently a Professor of Computer Science and Engineering at Washington University in St. Louis. He is the author of "Art of Computer Systems Performance Analysis," which won the 1991 "Best-Advanced How-to Book, Systems" award from Computer Press Association. His fourth book entitled " High-Performance TCP/IP: Concepts, Issues, and Solutions," was published by Prentice Hall in November 2003. He has co-edited "Quality of Service Architectures for Wireless Networks: Performance Metrics and Management," published in March 2010.

Exploiting Channel Fading and SVC in Wireless Video Streaming

Honghai Zhang, Mohammad A. Khojastepour, Ravi Kokku, Rajesh Mahindra, and Sampath Rangarajan,
NEC Laboratories America, USA
{honghai,amir,ravik,rajesh,sampath}@nec-labs.com

1. Introduction

Driven by the increasing wireless transmission rates and the advances in video compression technologies, streaming video over wireless networks has received increasing popularity in recent years. A key challenge in this type of applications is how to efficiently and fairly allocate radio resources among multiple users in a wireless network sharing the same spectrum frequency.

Several researchers [1,2] have studied the radio resource allocation problem for video streaming, although they did not consider the fading characteristics of wireless channels. Some researchers [3,4] developed algorithms to optimize realtime radio resource allocation. However, they only warranted asymptotic convergence, and did not consider the hard deadline constraint and bursty rate requirement of video applications.

In this work, we consider streaming video using SVC-encoded video sequences in wireless networks. An SVC (Scalable Video Coding) video stream contains a base layer and multiple enhancement layers. As long as the base layer is received, the receiver can decode the video stream. As more enhancement layers are received, the received video quality improves. We exploit both the *scalable* feature of SVC-encoded videos and the *fading* characteristics of wireless channels to improve the received video quality of video streaming.

2. Video Quality Model

As in [5] and [6], we use PSNR (Peak Signal to Noise Ratio) as a measure of video quality, and develop an empirical model to relate the rates and the PSNR values of video sequences truncated from an SVC-encoded video stream. The PSNR S of a video stream can be described as a piece-wise linear function of the rate r:

$$S_i(r) = \begin{cases} S_i^0 + L_i(r - r_i^0), & \text{if } r < r_i^0 \\ S_i^0 + K_i(r - r_i^0), & \text{if } r \le r_i^0 < r_i^{\max} \\ S_i^{\max}, & \text{if } r \ge r_i^{\max} \end{cases}$$

where r_i^0, S_i^0 are the rate and the PSNR of base layer and $r_i^{\max}$, $S_i^{\max}$ are the rate and PSNR of all layers. In Figure 1, we plot both the sample points of rates and PSNRs and the model we obtained for eight video sequences: News, Hall, Silent, City, Foreman, Crew, Harbour, and Mobile, all downloaded from [7]. It is easy to see that $L_i > K_i > 0$ for all video sequences, and therefore, PSNR is a concave, non-decreasing function of rate.

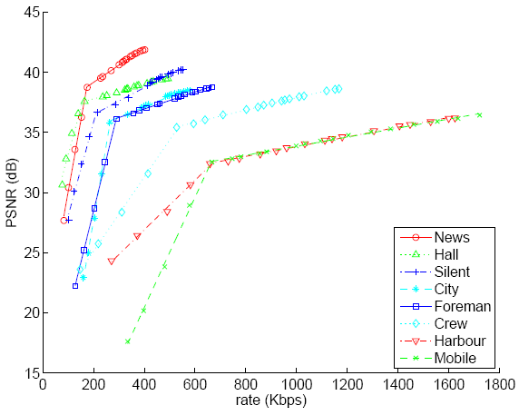


Fig. 2: Sample points and piece-wise linear regression model of the rate and PSNR.

3. A Static Scheduling Scheme

We first consider a static scheduling scheme that is solely based on long-term channel conditions and average video rate in making a decision of video rate control. In a non-fading wireless network, the problem can be solved via Time Division Multiple Access (TDMA), e.g. [2, 8]. In a fading channel, wireless networks often employ some channel-dependent scheduling algorithms to improve the average throughput and fairness (e.g. [9]). In this case, we denote the feasible region as **R** and attempt to solve the following optimization problem:

$$\text{Maximize} \sum_{i=1}^{n} w_i S_i(r_i) \qquad (1)$$

$$\text{s.t.} \qquad (r_1, r_2, \ldots, r_n) \in \mathrm{R}$$

It is shown in [3, 10] that the feasible region of the problem is convex, bounded and closed, and that the optimal solution has the following **maximal scheduling rule**:

At each time slot, the user with the largest u_iC_i is chosen for scheduling, (2)

where C_i is the channel capacity for user i and u_i is a constant for user i.

Given a vector $\vec{u} = (u_1, u_2, ..., u_n)$, the achieved user rate for user *i* is

$$r_i(\vec{u}) = E[C_i I(u_i C_i > u_j C_j, \text{ for all } j \neq i)] \quad (3)$$

We now formulate the optimization problem as

$$\text{Maximize } Y = \sum_{i=1}^{n} w_i S_i(r_i(\vec{u})) \qquad (4)$$

To solve this problem, we develop a gradient-based algorithm and prove that it converges to the optimal solution in [6] even though the problem (4) is not convex in the variable space $\vec{u}$. After obtaining the optimal solution, we compute the rate r_i according to Eq. (3) and truncate the video from top layers until the remaining video rate is below or equal to r_i. This solution is termed as the static scheduling scheme.

4. A Dynamic Scheduling Scheme

To address the burstiness and delay requirement of video traffic, we develop a dynamic scheduling solution as follows. First, we convert the burstiness and delay requirement into rate requirements. Assuming that for user *i*, the total size of packets with deadline T_j is L_j, we define the target rate $\overline{r_i}$ for user *i* as

$$\overline{r_i} = \max_j \frac{\sum_{k \le j} L_k}{T_j - t},$$

where *t* is the current time. We next apply the same maximal scheduling rule as in Eq. (2), and consider the following max-min problem:

$$\text{Maximize } \min \frac{r_i(\vec{u})}{\overline{r_i}} \qquad (5)$$

Essentially, if the optimal solution to (5) $\eta \geq 1$, it indicates that all users can achieve their required rate using the maximal scheduling rule with the optimal solution $\vec{u}$ in (5). Otherwise, the required rate for all users cannot be supported, and some packets may need to be dropped. In [6], we also develop an algorithm to solve problem (5) optimally.

Let the optimal solution to (5) be η. We design the following packet dropping rule. If $\eta < \underline{\eta}$, the scheduler marks some packets to be dropped in the order of priority (which is based on their layer numbers). These marked packets are dropped unless they are un-marked before their deadline is passed. If $\eta > \overline{\eta}$, the scheduler un-marks some packets if there exist marked packets whose deadline has not passed. Typically, $\underline{\eta} < 1 < \overline{\eta}$.

5. Performance Evaluation

We perform extensive simulations to evaluate the performance of the proposed algorithms. We consider three reference schemes. The first scheme is *Maximum capacity scheduling w/ FD* (FD refers to frame dropping), where the user selection is based on maximum channel capacity and packets are dropped based on their priority at the time of buffer overflow in [11]. The buffer limit for each link is 110KBytes as used in [11]. Packets with the lowest priority are dropped first at the time of buffer overflow. The other two reference schemes are variations of the *Maximum capacity w/ FD* where different user scheduling algorithms are employed. The second scheme uses proportional fairness scheduling [9], and the third uses Modified Largest Weighted Delay First (M-LWDF) scheduling [3].

Figure 2 shows the average video quality with 8 video sequences (as in Figure 1), with average wireless channel SINR values randomly generated with uniform distribution from 5dB to 20dB for each video client. Figure 3 shows the average video quality with the same 8 video sequences but their average channel SINRs are all equal and take values from 4dB to 20dB. Simulation results suggest that both the static and dynamic scheduling schemes significantly improve video quality over existing wireless streaming solutions.

6. Implementation Alternatives

Implementing the solutions proposed in this paper in a real system, however, requires considering two design alternatives and addressing several interesting challenges. The solutions may either be implemented within the base station or on a gateway (e.g. the ASN Gateway in a WiMAX network). For implementation within the base station, one has to consider the overhead of deep packet inspection to identify video packets and their payload type or priority. Typically, there are much more base stations than gateways, and hence the cost of additional hardware per base station adds significant overhead for a network operator.

For implementation on the gateways, there are two design considerations: (1) appropriate channel information has to be fed back to the gateways to assist in choosing the best user at each time. (2) the channel information has to be fed back frequently for leveraging channel fading characteristics effectively. While the former may require agreement through standardization between the base station and gateway manufacturers, the latter requires the tradeoff study of the frequency of information, the amount of information, the resulting efficacy of the scheduling algorithms with coarser feedback, and the scalability of gateways to a large number of basestations. Our current work is focusing on implementation of the algorithms on the gateways in a WiMAX testbed, and we are studying above issues.

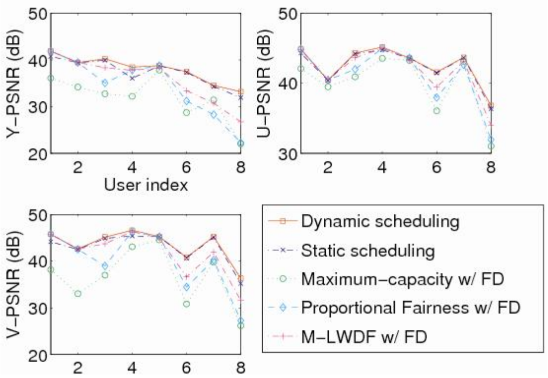


Fig. 3: Average PSNR of different schemes for each user.

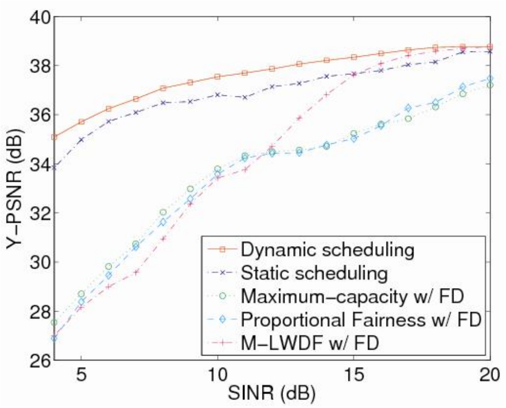


Fig. 4: Average PSNR of all users for each scheme

7. Conclusion

This paper proposes cross-layer solutions for streaming scalable video in wireless networks.

We exploit both the scalable video characteristics of video sequences and the fading property of wireless channels to optimize the video transmissions. A static scheduling scheme and a dynamic one are developed to optimize the received video quality. Simulation results show that the proposed algorithms significantly improve the video quality.

References

[1] W. Tu, J. Chakareski, and E. Steinbach. Rate-distortion optimized frame dropping for multiuser streaming and conversational videos. Advances in Multimedia, 8(2), Jan. 2008.

[2] M. Kalman and B. Girod. Optimized transcoding rate selection and packet scheduling for transmitting multiple video streams over a shared channel. Proc. IEEE International Conference on Image Processing, ICIP-2005, September 2005.

[3] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting. Providing quality of service over a shared wireless link. IEEE Communication Magazine, Feb. 2001.

[4] S. Shakkottai and R. Srikant. Scheduling real-time traffic with deadlines over a wireless channel. In WoWMoM, 1999.

[5] H. Zhang, Y. Zheng, M. A. Khojastepour, and S. Rangarajan. Scalable video streaming over fading wireless channels. In IEEE Wireless Communications & Networking Conferences, April 2009.

[6] H. Zhang, Y. Zheng, M. A. Khojastepour, and S. Rangarajan. Cross-layer optimization for streaming scalable video over fading wireless networks. IEEE J-SAC, April 2010.

[7] Video test sequences. http://trace.eas.asu.edu/yuv/.

[8] W. Kuo and W. Liao. Utility-based resource allocation in wireless networks. IEEE Transactions on Wireless Communcations, 6(10), Oct. 2007.

[9] D. Tse. Multiuser diversity in wireless networks. http://www.eecs.berkeley.edu/~dtse/stanford416.ps, 2001.

[10] A. Stolyar. On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation. Operations Research, 53(1), 2005.

[11] Gunther Liebl, Thomas Schierl, Thomas Wiegand, and Thomas Stockhammer. Advanced wireless multiuser video streaming using the scalable video coding extensions of H.264/MPEG4-AVC. In IEEE ICME 2006.

Honghai Zhang received the B.S. degree from the University of Science and Technology of China and the Ph.D. degree from the Department of Computer Science at University of Illinois at Urbana-Champaign. He was a recipient of Vodafone Fellowship during his Ph.D. study. He is currently a Member of Technical Staff in NEC

Laboratories America. Prior to the current position, he was in the Wireless Advanced Technology Laboratory of Alcatel-Lucent. His research interests include scheduling, interference mitigation, and video streaming in mobile wireless networks, and cognitive radio networks.

Mohammad Ali (Amir) Khojastepour (S'02-M'05) received the B.Sc. and M.Sc. degrees in electrical and computer engineering from Shiraz University, Shiraz, Iran, in 1993 and 1996, respectively, and the Ph.D. degree from Rice University, Houston, TX, in 2004. Since 2004, he has been a Member of Technical Staff with the Mobile and Signal Processing Department, NEC Research Laboratories America, Princeton, NJ. Prior to that, he was a Researcher with different research institutes, including MATN Research Center, Tehran, Iran, Bell Laboratories/Lucent Technologies, Holmdel, NJ, and Mitsubishi Electric Research Laboratories, Cambridge, MA. His research interests are in the areas of information theory and coding, communication theory and signal processing with emphasis on multiuser communications, and wireless communication networks.

Ravi Kokku is a research staff member at NEC Laboratories America, Princeton, USA. He received his M.S. and Ph.D. in Computer Sciences from The University of Texas at Austin (2000-05), and B.Tech in Computer Sciences from Indian Institute of Technology, Kharagpur (1993-97). From 1997 to 2000, he worked at Hughes Software Systems (Gurgaon, India) working on the design and development of satellite communication systems. His current research interests include designing and prototyping innovative systems in WiMAX networks. He is a recipient of several awards including the IBM doctoral fellowship (2001-2004) and Best paper award at BROADNETS (2008).

Rajesh Mahindra received his B. Tech. degree from the Department of Electronics and Communications Engineering, M.S.Ramaiah Institute of Technology, Bangalore, India in 2005 and his M.S. degree from the School of Electrical and Computer Engineering, Rutgers University, NJ. He is currently working as a assistant research associate at NEC Labs America. His research interests include video streaming over wireless, wireless network virtualization and wireless resource management.

Sampath Rangarajan received the MS degree in electrical and computer engineering in 1987 and the PhD degree in computer science in 1990, both from the University of Texas at Austin. He heads the Mobile Communications and Networking Research Department at NEC Laboratories America in Princeton, New Jersey. Previously, he was with the Networking Research Center at Bell Laboratories in Holmdel, New Jersey. Prior to that, he was a cofounder and vice president of technology at Ranch Networks, a venture funded startup in the IP networking space. Earlier, he was a researcher in the Systems and Software Research Center at Bell Laboratories in Murray Hill, New Jersey. Before joining Bell Laboratories, he was an assistant professor in the Electrical and Computer Engineering Department at Northeastern University in Boston, Massachusetts. His research interests span the areas of mobile communications, mobile networks and distributed systems. He is a senior member of the IEEE. He has been on the editorial boards of IEEE Transactions on Computers and ACM Mobile Computing and Communications Review and currently is a member of the editorial board of the IEEE Trans. on Parallel and Distributed Systems.

Editor's Selected Paper Recommendation

Editor: Guan-Ming Su, Marvell Semiconductors, USA

J. Wang, A. Majumdar, and K. Ramchandran, "Robust Video Transmission With Distributed Source Coded Auxiliary Channel," IEEE Trans. on Image Processing, vol. 18, no.2, p.2695-2705, Dec. 2009.

Video transmission over wireless networks has been rapidly increasing in the past few years, spurred on by the widespread availability of 3G networks and the increasing market penetration of smartphones. Video delivery over wireless networks, however, poses some unique challenges primarily related to the underlying unreliability of wireless channels. Specifically, packet losses are far more frequent and bursty than in wire-line networks. Since modern video codecs are prediction-based, packet drops during transmission lead to predictor mismatch or ``drift" between encoder and decoder, which causing unpleasant visual artifacts.

Traditionally, to provide a measure of error resilience for video, Automatic Repeat Request (ARQ) (i.e. re-transmission of packets) or Forward Error Correction codes (FEC) have been used. However, both ARQ and FEC schemes are not suitable for certain applications with very tight latency requirements. In particular, ARQ based schemes require packet retransmission, which incurs an additional delay of at least one round-trip time. ARQ schemes also require a feedback channel, and are not well-suited to multicast/broadcast scenarios. In the case of FEC, because long block lengths are required for strong codes, the usefulness of FEC-based schemes is limited by delay constraints. Further, with finite block length, FEC-based schemes can only mitigate the probability of error. When errors occur, they propagate until receiving next intra-coded data. Therefore, even when FEC codes are used, there still is a need to mitigate the effects of possible drift.

In this paper, the authors send additional information alongside a regular predicatively coded video bitstream for the purpose of correcting drift. The auxiliary encoder and decoder are designed based on the principles of distributed source coding [1]. The key point to note here is that while the decoded frame may be in error due to drift, it is still highly correlated to the correct reconstruction. Thus, the decoded frame can be used as side information at the auxiliary decoder for the purpose of decoding the auxiliary bitstream. Due to the high degree of correlation between the decoded frame and the correctly reconstructed frame, the auxiliary encoder will typically need to send very few bits to correct the errors.

In the theoretical distributed source coding setup, the correlation between the source and the side information is assumed to be known. In practice, the correlation between the original source frame and the erroneously decoded frame needs to be estimated since real-world video sources are highly non-stationary. The major contributions of this paper are firstly to design an analytically tractable correlation estimation algorithm that dynamically incorporates the effect of channel loss and secondly to devise a rate-distortion optimization scheme to allocate the bit rate used by the auxiliary encoder among the encoding blocks within a video frame.

The authors compare the proposed system with the traditional approach of using FEC to protect the video bitstream and show that as the average burst error length increases the proposed system outperforms the traditional approach in terms of PSNR as well as visual quality.

Future research on generalizing DSC over auxiliary channel(s) in multiuser/ cooperative/ multihop/ cognitive wireless environment where multiple channels exist can be exploited. The multicast scenario can be also further studied [2].

[1] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Trans. Information Theory, vol. 22, pp. 1–10, Jan. 1976.
[2] J. Wang and K. Ramchandran, "Receiver-driven multicast over wireless with distributed source coding and FEC", Proc. wireless communications and mobile computing, 2007.

E-Letter Editorial Board

DIRECTORS

EDITOR

MMTC Officers

CHAIR

VICE CHAIRS

SECRETARY