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

E-LETTER

Vol. 7, No. 8, November 2012
IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from MMTC Chair

Dear Fellow MMTC Members,

During the past few months, we have been working hard to invite new MMTC leader team members on board. I am very happy to introduce the new team to you.

The Board Leaders are:

E-letter Board
Director: Shiwen Mao, Aburn University, USA
Co-Director: Guosen Yue, NEC Labs, USA
Co-Director: Periklis Chatzimisios, Alexander Technological Educational Institute of Thessaloniki, Greece

R-letter Board
Director: Irene Cheng, University of Alberta, Canada
Co-Director: Xianbin Wang, University of Western Ontario, Canada

Award Board
Director: Mung Chiang, Princeton University, USA
Co-Director: Vincent Wong, University of British Columbia, Canada

Services & Publicity Board
Director: Joel Rodrigues, University of Beira Interior, Portugal
Co-Director: Honggang Wang, University of Massachusetts Dartmouth, USA

Membership Board
Director: Ci Song, University of Nebraska-Lincoln, USA
Co-Director: Ping Wang, Nanyang Technological University, Singapore

The IG (Interest Groups) leaders are:

(QoEIG) QoE for Multimedia Communications
Chair: Weisi Lin, Nanyang Technological University, Singapore
Vice-chairs: Tasos Dagiuklas, Technological Institute of Messolonghi
Adlen Ksentini, IRISA - University of Rennes
Advisor: Chang Wen Chen, University of Buffalo, USA

(ASPIG) Acoustic and Speech Processing for Communications
Chair: Zhiqiang Wu, Wright State University, USA
Vice-chairs: Xihong Wu, Peking University, China
Pei-Jung Chung, University of Edinburgh, UK
Advisor: Jerry Gibson, UCSB, USA

(VAIG) Visual Analysis, Content Management and Interaction for Communications
Chair: Benoit Huet, EURECOM, France
Vice-chairs: Xu Dong, NTU, Singapore
Gene Cheung, NII, Japan
Advisor: Xian-Sheng Hua, Microsoft Research Asia, China

(IMVIMEC) Image, Video & Mesh coding
Chair: Petros Daras, CERTH/Information Technologies Institute, Greece
Vice-chairs: Aljoscha Smolic, Disney Research Zurich, Switzerland

Paulo Nunes, University Institute of Lisbon/Instituto de Telecomunicações, Portugal
Ebroul Izquierdo, Queen Mary University of London, UK
Advisor: Aggelos Katsaggelos, Northwestern University, USA

(MSIG) Media Streaming

Chair: Nabil J. Sarhan, Wayne State University, USA
Vice-chairs: Di Wu, Sun Yat-Sen University, China
Sanjeev Mehrotra, Microsoft Research, USA
Advisor: John Apostolopoulos, HP Labs, USA

(SecIG) Media Processing and Communications Security

Chair: Jiangtao (Gene) Wen, Tsinghua University, China
Vice-chairs: Haiyong Xie, Huawei Research Labs, USA
Advisor: Wenjun Zeng, University of Missouri, USA

(3DRPCIG) 3D Rendering, Processing and Communications

Chair: Maria Martini, Kingston University-London, UK
Vice-chairs: Cha Zhang, Microsoft Research, USA
Jianfei Cai, Nanyang Technological University, Singapore
Web Chair: Chaminda Hewage, Kingston University-London, UK
Advisor: Chang Wen Chen, University of Buffalo, USA

(WTIG) Wireless Technology for Multimedia Communications

Chair: Fen Hou, Macao Polytechnic Institute
Vice-chairs: Ali Tajer, Wayne State University, USA
Ping Wang, Nanyang Technological University, Singapore
Lin Cai, HUAWEI, USA
Advisor: Sherman Shen, University of Waterloo, Canada

(GMCIG) Green Multimedia Communications

Chair: Kyle Guan, Bell Labs, Alcatel-Lucent, USA
Vice-chairs: Jaime Llorca, Bell Labs, Alcatel-Lucent, USA
Wee Peng Tay, Nanyang Technology University, Singapore

(DSNIG) Distributed and Sensor Networks for Mobile Media Computing and Applications

Chair: Min Chen, Huazhong University of Science and Technology, China
Vice-chairs: Jun Wu, Tongji University, China
Yevgeni Koucheryavy, Tampere University of Technology, Finland
Advisor: Marie-Jose Montpetit, MJMontpetit.com, USA

(MENIG) Multimedia Services and Applications over Emerging Networks

Chair: Chuan Wu, The University of Hong Kong, Hong Kong
Vice-chairs: Zhou Su, Waseda University, Japan
Chin-Feng Lai, National Ilan University, Taiwan, ROC
Advisor: Dah-Ming Chiu , Chinese University of Hong Kong, Hong Kong

(CLDIG) Cross-layer Design for Multimedia Communications

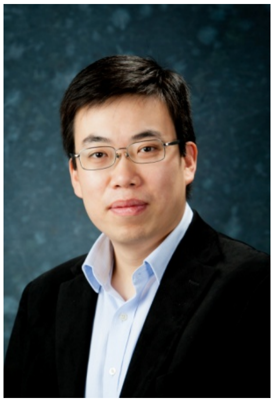
Chair: Dusit (Tao) Niyato, Nanyang Technological University, Singapore
Vice-chairs: Xiaoqing Zhu, Cisco, USA
Advisor: Ekram Hossain, University of Manitoba, Canada

(MCDIG) Multimedia Content Distribution: Infrastructure and Algorithms

Chair: Zongpeng Li, University of Calgary, Canada
Vice-chairs: Zuqing Zhu, University of Science and Technology of China, China
Ying Li, Samsung, USA
Advisor: Keith W. Ross, Polytechnic Institute of NYU, USA

I am very excited to have such an excellent team in place, and I look forward to working closely with the team as well as every MMTC member to make the TC an even stronger one.

Regards,



Jianwei Huang

Chair, IEEE ComSoc Multimedia Communication Technical Committee
http://jianwei.ie.cuhk.edu.hk/

EMERGING TOPICS: SPECIAL ISSUE ON “Wireless Multimedia Communications”

Wireless Multimedia Communications

Guest Editors: Periklis Chatzimisios[1] *and Albert Banchs*[2 3]

[1]*Alexander TEI of Thessaloniki, Greece, pchatzimisios@ieee.org*
[2] *Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganes, Spain,*
[3] *Institute IMDEA Networks, Avda. del Mar Mediterraneo, 28912 Leganes, Spain*
banchs@it.uc3m.es

This Special Issue gathers a collection of five selected articles that present a variety of concerns and latest advances relating to Wireless Multimedia Communications.

The first article “Energy-Efficient Delay-Critical Communication in Unknown Wireless Environments” presents a summary of novel reinforcement learning approaches to achieve energy-efficient point-to-point transmission of delay-critical multimedia data in unknown wireless environments. In particular, the proposed solutions can be applied to significantly more complex transmission strategies and network scenarios for more complex video traffic.

The second article “An Enriched Multimedia Experience for Wireless Networks in Horizon 2020 and Beyond” the paper focuses on future mobile terminals and present their view on new multimedia technologies, as the crucial ingredients for the immersive, collaborative, emotional and interactive user experience. The authors also introduce a new notion of collaborative multimedia, 3D audio & video formats and evolution of the corresponding codecs. Furthermore, this article provides some insights into augmented reality for several applications (e.g. gaming, advertising, localization, and orientation).

The third article “Towards QoE-Driven Resource Allocation for Multimedia Services in LTE” introduces a solution for a domain wide QoE-driven resource allocation of multimedia services based on the negotiation and calculation of a per-session Media Degradation Path (MDP). The authors present the simulation tool that was developed for evaluating the proposed model, considering the notion of 3GPP-based Evolved Packet System (EPS) over Long Term Evolution (LTE) wireless networks. Finally, the article includes the validation of the proposed approach by using their simulation results as input for the publicly available LTE-Sim simulator.

The fourth article “Radio Resource Management for Video Traffic in IEEE 802.16 Networks” presents two proposed solutions for intra-cell radio resource management optimization for video traffic. The authors specifically focus on IEEE 802.16 and they propose an enhanced multi-model CAC function based on measurements as well as a prediction based “long-term” bandwidth reservation scheme. Subsequently, they provide a performance evaluation that demonstrates the positive impact of the two solutions on both Grade of Service (GoS) and Quality of Service (QoS).

The last paper “Watermarking Technologies in Wireless Multimedia Sensor Networks” studies the existing methods of using watermarking technologies in order to provide security protection for multimedia data streaming over Wireless Multimedia Sensor Networks (WMSNs). In particular, the authors classify the watermarking based schemes in two categories (quality-aware schemes and resource-aware schemes) and then discuss new research issues that can further enhance the related applications.

We would like to thank all the authors for their contribution and hope these articles will stimulate further research on wireless multimedia communications and help by providing an up-to-date sketch of currently hot topics on this research area.

Finally, we want to express our gratitude to the Vice Chair of Letter & Member Communications, Dr. Kai Yang, as well as to Shiwen Mao and Guosen Yue (Director and co-Director of the IEEE MMTC E-letter, respectively) for their invaluable support in coordinating this Special Issue.

Enjoy your reading!

Periklis Chatzimisios is an Assistant Professor with the Department of Informatics at the Alexander TEI of Thessaloniki. He holds a Ph.D. degree in Computer Engineering from Bournemouth University. He currently serves as a Member for the IEEE Communication Society (ComSoc) Standards Board. He also serves as Organizing/TPC member and co-Chair for several conferences and he holds editorial board positions for many IEEE and non-IEEE journals. He has been an Organizer/Invited Speaker for many dissemination events, talks and seminars.

Dr. Chatzimisios is the author/editor of 8 books and more than 70 peer-reviewed papers and book chapters. His published research work has received more than 600 citations by other researchers. His current research activities are mainly focused on wireless communications, multimedia communications (mainly Quality of Service and Quality of Experience) and security.

Albert Banchs received his degree in telecommunications engineering from the Polytechnic University of Catalonia in 1997, and his Ph.D. degree from the same university in 2002. He received a national award for the best Ph.D. thesis on broadband networks. He was a visiting researcher at ICSI, Berkeley, in 1997, worked for Telefonica I+D, Spain, in 1998, and for NEC Europe Ltd., Germany, from 1998 to 2003. He has been with the University Carlos III of Madrid since 2003. Since 2009, he also has a double affiliation as Deputy Director of the institute IMDEA Networks.

Albert Banchs has authored over 80 publications in peer-reviewed journals and conferences and holds six patents. He is area editor for Computer Communications and has been senior and associate editor for IEEE Communications Letters and guest editor for IEEE Wireless Communications, Computer Networks and Computer Communications. He has served on the TPC of a number of conferences and workshops including IEEE INFOCOM, IEEE ICC and IEEE GLOBECOM, and has been TPC chair for European Wireless 2010, IEEE HotMESH 2010 and IEEE WoWMoM 2012. He is senior member of IEEE.

Energy-Efficient Delay-Critical Communication in Unknown Wireless Environments

Nicholas Mastronarde and Mihaela van der Schaar***
**State University of New York at Buffalo, USA, **University of California Los Angeles, USA*
nmastron@buffalo.edu, mihaela@ee.ucla.edu

1. Introduction

Delay-critical multimedia applications such as videoconferencing, surveillance, medical health monitoring, etc. often operate in dynamic wireless environments where they experience time-varying and a priori unknown channel conditions and traffic loads. In [1][2][3], we propose to learn the impact of these dynamics on the user's utility using a novel class of online reinforcement learning techniques that do not require a priori specified models of these environments.

Many existing solutions for transmitting delay-critical packets under energy constraints react to the experienced dynamics in a "myopic" way, by optimizing the transmission strategies based only on information about the current traffic and channel dynamics [9][10]. However, our work in [1][2][3] shows that, to achieve the optimal power-performance tradeoff, users should adopt foresighted strategies that account for the fact that current decisions impact their immediate and future performance. Meanwhile, many existing foresighted solutions assume that the underlying dynamics are known [11][12]; however, when they are unknown, only heuristic solutions are provided, which cannot guarantee optimal performance because they ignore side information about the traffic and channel dynamics. Alternative methods for optimizing transmission in unknown environments rely on reinforcement learning [13][14], where the goal is to learn the state-value function, which provides a measure of the expected long-term performance (utility) when the transmitter is acting optimally in a dynamic environment. However, these methods have to learn the state-value function in every possible state; as a result, they incur large memory overheads for storing the state-value function and they are slow to adapt to new or dynamically changing environments, especially when the state space is large. This limitation makes them unsuitable for wireless multimedia communication, where *low-delay and energy-efficiency* are paramount. Our methods and algorithms proposed in [1][2][3] are able to fulfill these two requirements by exploiting the specific structures and unique properties of the considered transmission policies, as well as their utilities and experienced dynamics.

The remainder of this letter is organized as follows. In Section 2, we formulate the transmission problem. In Section 3, we describe the learning algorithms proposed in [1][2][3]. In Section 4, we present our numerical results and conclude in Section 5.

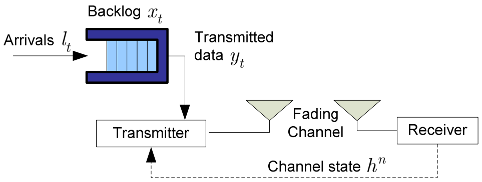


Figure 1. Point-to-point transmission model.

2. Illustrative point-to-point transmission model

We consider a simple model for the energy-efficient transmission of delay-critical data over a fading channel to illustrate our learning solutions. We assume that time is slotted into discrete-time intervals of length Δt. We let $h_t \in \mathcal{H}$ denote the fading coefficient over the point-to-point link in slot t as illustrated in Figure 1. We assume that the set of channel states $\mathcal{H}$ is discrete and finite and that the sequence of channel states can be modeled as a Markov chain with transition probabilities $p_h(h_{t+1} \mid h_t)$ [1][2][13][14]. Importantly, $p_h(\cdot \mid \cdot)$ is unknown a priori.

The source injects $l_t \in \{0,1,\ldots\}$ packets into the FIFO transmission buffer in each slot. The arrivals are assumed to be i.i.d. with distribution $p_l(l_t)$, which is also unknown a priori. The buffer state x_t evolves recursively: i.e., $x_{t+1} = Purge(x_t - y_t + l_t, D)$, where x_0 is the initial buffer state, $0 \leq y_t \leq x_t$ is the amount of data transmitted in slot t, and packets whose delay deadlines (denoted by D) have passed are purged from the buffer. Based on this buffer recursion and the arrival distribution $p_l(\cdot)$, the sequence of buffer states $\{x_t : t = 0,1,\ldots\}$ can be modeled as a controlled Markov chain with transition probabilities $p(x_{t+1} \mid x_t, y_t)$ [1][2][13]. When y_t packets are transmitted in slot t, the immediate utility received by the transmitter is denoted by $u(x_t, y_t) \geq 0$ and the incurred transmission cost is denoted by $c(h_t, y_t) \geq 0$. The utility can be defined in terms of the delay and the costs can be the exerted transmission energy.

Constrained Markov decision process formulation

We define the state of the system at time t as $s_t = (x_t, h_t)$ and the action at time t as y_t. Hence, the sequence of states $\{s_t : t = 0,1,\ldots\}$ can be modeled as a controlled Markov chain with transition probability function $p(s_{t+1} \mid s_t, y_t)$. The objective for the user is to determine the transmission action in each slot to maximize its long-term utility subject to a constraint on its long-term cost: i.e.,

$$\begin{aligned} \max_{0 \le y_t \le x_t, \forall t} & E\left[\sum_{t=0}^{\infty} \alpha^t u(x_t, y_t)\right] \\ \text{s.t. } & E\left[\sum_{t=0}^{\infty} \alpha^t c(h_t, y_t)\right] \le \overline{c} \end{aligned} \quad (1)$$

where $\alpha \in [0,1)$ is the discount factor and $\overline{c}$ is the average cost constraint [1][2]. We can reformulate the constrained optimization in (1) as an unconstrained optimization by introducing a Lagrange multiplier $\lambda \ge 0$ associated with the cost constraint. For a fixed λ, the optimal solution to the unconstrained optimization satisfies the following Bellman equation:

$$J^{*,\lambda}(s) = \max_{0 \le y \le x} \begin{bmatrix} u(x,y) - \lambda c(h,y) + \\ \alpha \sum_{s' \in S} p(s' \mid s, y) J^{*,\lambda}(s') \end{bmatrix}, \forall s. \quad (2)$$

We refer to $J^{*,\lambda}$ as the optimal *state-value function* (SVF). The optimal policy $\pi^{*,\lambda}(s)$, which gives the optimal action to take in each state, can be determined by taking the action that maximize the right-hand side of (2) (assuming the dynamics are known).

Post-decision state based dynamic programming

A key element of our approach is the use of an intermediate state called the post-decision state $\tilde{s} = (\tilde{x}, \tilde{h})$ (hereafter PDS) and of a value function over the PDSs that we refer to as the post-decision state-value function V^{λ} (hereafter PDSVF). The relationship between the PDS $\tilde{s}_t = (\tilde{x}_t, \tilde{h}_t)$ and the conventional state $s_t = (x_t, h_t)$ is as follows:

$$\tilde{x}_t = x_t - y_t, \tilde{h}_t = h_t, x_{t+1} = Purge\left(\tilde{x}_t + l_t, D\right). \quad (3)$$

The buffer's PDS $\tilde{x}_t = x_t - y_t$ characterizes the buffer state after the data is transmitted, but before new data arrives; and the channel's PDS is the same as the channel state at time t. In other words, the PDS incorporates all of the known information about the transition from state s_t to state s_{t+1} after taking action y_t. Meanwhile, the next state incorporates all of the unknown dynamics that were not included in the PDS (i.e., the packet arrivals and next channel state).

The conventional SVF J^{λ} and PDSVF V^{λ} are related through the following Bellman equations:

$$J^{*,\lambda}(x,h) = \max_{0 \le y \le x} \begin{bmatrix} u(x,y) - \lambda c(h,y) \\ +\alpha V^{*,\lambda}(x - y, h) \end{bmatrix}, \quad (4)$$

$$V^{*,\lambda}(x,h) = \mathbf{E}_{l_t, h' \mid h}\left[J^{*,\lambda}\left(Purge(\tilde{x}_t + l_t, D), h'\right)\right], \quad (5)$$

Equation (5) shows that the PDSVF $V^{*,\lambda}$ is obtained from the conventional SVF $J^{*,\lambda}$ by taking the expectation over the possible traffic arrivals and channel transitions. Meanwhile, (4) shows that the conventional SVF can be obtained from the PDSVF $V^{*,\lambda}$ by maximizing over the scheduling actions.

Studying the Bellman equations for the conventional SVF (i.e., (2)) and those for the PDSVF (i.e., (4) and (5)), we notice: First, in the latter the expectation over the channel states and packet arrivals is separated from the maximization. This allows us to separate learning and decision making (see PDS learning in Section 3). Second, because the channel and traffic dynamics are independent of the queue backlog, we can develop a batch online update of the PDSVF to enable fast learning (see virtual experience learning in Section 3).

3. Learning the optimal policy

Post-decision state learning

In slot t, given the state $s_t = (x_t, h_t)$, the scheduling action y_t should be selected to maximize the right-hand side of (4). However, we cannot compute the PDSVF $V^{*,\lambda}$ using (5) because we do not know $p_l(\cdot)$ and $p_h(\cdot|\cdot)$; instead, we must update it online using reinforcement learning [8]. PDS learning uses two simple update steps: the first updates the conventional SVF and the second updates the PDSVF:

$$J_t^{\lambda}(x_t, h_t) \leftarrow \max_{0 \le y \le x_t} \begin{bmatrix} u_t(x_t, y_t) - \lambda c_t(h_t, y_t) \\ +\alpha V_t^{\lambda}(x_t - y, h_t) \end{bmatrix}, \quad (6)$$

$$\begin{aligned} V_{t+1}^{\lambda}(\tilde{x}_t, h_t) \leftarrow & (1 - \beta_t) V_t^{\lambda}(x_t, h_t) + \\ & \beta_t J_t^{\lambda}\left(Purge(\tilde{x}_t + l_t, D), h_{t+1}\right) \end{aligned}, \quad (7)$$

where β_t is a learning rate factor. It can be shown that the above updates converge to the optimal PDSVF $V^{*,\lambda}(x,h)$ [1][2][13]. Equation (7) computes the PDSVF by replacing the expectation in (5) with a sample average obtained from interactions with the environment. Note that λ^* can be found using stochastic subgradient methods [1][2][13].

Virtual experience learning

PDS learning only updates one PDS in each time slot [13]. In most wireless communication settings, however, the data arrivals and channel state transition are independent of the queue backlog x_t (e.g., [1][2][11][12][13][14]) such that the traffic arrival l_t and next channel state h_{t+1} would have been realized at any possible backlog x with the same probability. Hence, instead of updating the PDSVF only at the PDS $(\tilde{x}_t, h_t)$ as in [13], we update the PDSVF at all backlog states with the same channel state h_t (i.e.,

$\{(\tilde{x},h_t),\forall\tilde{x}\}$). We refer to this as virtual experience (or batch) learning because we did not actually visit all of the buffer states that we update [1][2].

In Figure 2, we compare the Q-learning [8], PDS learning [13], and virtual experience learning algorithms [1] with the optimal policy. The x-axis is in log-scale to highlight how the cumulative average power and delay change over time. The virtual experience learning algorithm achieves optimal delay performance within 300 slots (corresponding to 3 seconds for 10 ms slots) and optimal power performance within 1000 slots; the PDS learning algorithm achieves optimal delay and power performance within 10,000 and 50,000 slots, respectively; and within 50,000 slots Q-learning neither converges to the optimal power-performance tradeoff nor meets the 30 mW power constraint.

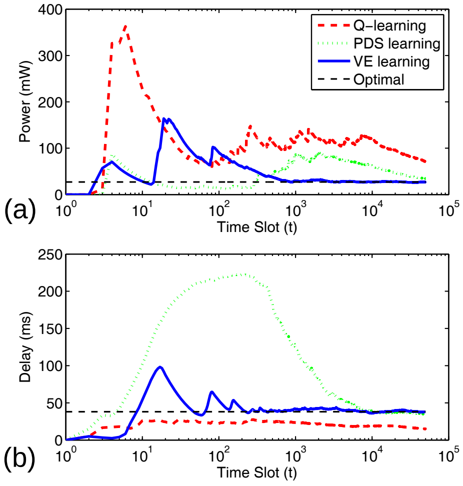


Figure 2. Learning algorithm performance comparison ($\alpha = 0.98$). (a) Average power. (b) Average delay.

Approximation-based learning

The complexity of virtual experience learning increases linearly in the number of buffer states and therefore may be computationally prohibitive if the capacity of the queue is large. To address this, we propose an approximation-based learning algorithm that allows us to systematically tradeoff the complexity of the learning algorithm and its achievable performance, without sacrificing the convergence rate.

The approximation-based learning algorithm is made possible by the fact that the typical utility/cost functions in the considered multimedia transmission problems satisfy certain properties of discrete functions, namely supermodularity [15] and multimodularity [16], which impose specific structures on the optimal policy and SVF: in particular, (i) the monotonic structure of the optimal scheduling policy (i.e., the optimal scheduling action $\pi^{*,\lambda}(x)$ is non-decreasing in the buffer state) and (ii) the concavity of the extensions of the SVF $J^{*,\lambda}(x,h)$ and PDSVF $V^{*,\lambda}(x,h)$, where the extension is defined as the linear interpolation of the functions over $x \in \mathbb{R}_+$ [16].

Because of the concavity of the PDSVF's extension, we are able to compactly represent it using a piece-wise linear function approximation, thereby decoupling the learning update complexity from the (potentially large) number of buffer states. Our preliminary results show that the energy increase due to approximation is *less than 5%* for a fixed delay, while the number of states updated in each slot *decreases by two orders of magnitude* [2]. Hence, the proposed approximate approach can be implemented on a user employing limited processing and memory capacity, with minimal performance penalties in terms of delay and energy.

Instead of updating all the states for the PDSVF, we can exploit the approximation of the concave PDSVF $V_t^{\lambda}(\tilde{x},h)$ and only update a limited number of states in each slot, which can be determined by an adaptive method. Specifically, we can obtain an approximately optimal scheduling policy $\hat{\pi}(x,h_t)$ by replacing the PDSVF V_t^{λ} in (6) with the approximated PDSVF $\hat{V}_t^{\lambda}$. We can also update the PDSVF $V_t^{\lambda}(\tilde{x},h)$ similar to (7), but by approximating it with a concave and increasing function. Our results in [2] indicate that the learned PDSVF converges to an ε-optimal PDSVF, where $\varepsilon > 0$ is a user-defined parameter.

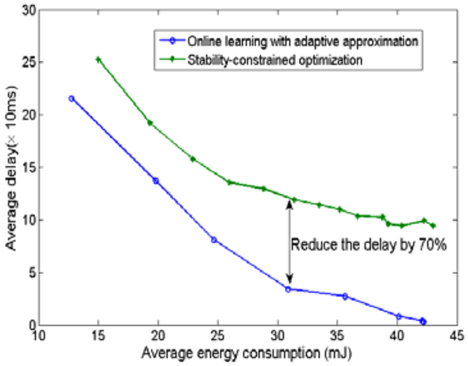


Figure 3. Delay-energy tradeoff comparison in a Markovian channel.

In Figure 3, we compare our proposed approximation-based approach with the stability-constrained optimization method proposed in [10]. Targeting stability leads to poor performance in the low-delay region because it does not directly minimize queuing delays and it ignores the channel and multimedia traffic dynamics. Figure 3 shows the delay-energy tradeoffs of our adaptive approximation-based learning approach and the stability-constrained optimization, from which it is clear that our proposed method achieves lower power consumption in both the high delay region (>150 ms) and the low delay region. In the low delay region, the proposed method can reduce the delay by 70% or more for the same average energy consumption. Other preliminary results, not reported here due to space limitations, have shown improvements of 60% even when the channel is *not* Markovian [2].

5. Conclusion

We provide a high-level summary of the novel reinforcement learning approaches proposed in [1][2][3] to achieve energy-efficient point-to-point transmission of delay-critical multimedia data in unknown wireless environments. Importantly, the proposed solutions can be applied to significantly more complex transmission strategies and network scenarios such as multi-user, cooperative, multi-hop, etc. – see, e.g., [4][5][6][7] for some preliminary results – as well as more complex video traffic.

References

[1] N. Mastronarde and M. van der Schaar, “Joint physical-layer and system-level power management for delay-sensitive wireless communication,” *IEEE Trans. on Mobile Computing*, to appear. Available online: http://dx.doi.org/10.1109/TMC.2012.36

[2] F. Fu and M. van der Scahar, “Structure-aware stochastic control for transmission scheduling,” *IEEE Trans. on Vehicular Technology*, to appear. Available online: http://dx.doi.org/10.1109/TVT.2012.2213850

[3] F. Fu and M. van der Schaar, "Decomposition Principles and Online Learning in Cross-Layer Optimization for Delay-Sensitive Applications", *IEEE Trans. Signal Process.*, vol 58, no. 3, pp. 1401-1415, Feb. 2010.

[4] F. Fu and M. van der Schaar, "Structural Solutions for Dynamic Scheduling in Wireless Multimedia Transmission", *IEEE Transactions Circuits Systems for Video Tech.*, vol. 22, no. 5, pp. 727-739, May 2012.

[5] N. Mastronarde, F. Verde, D. Darsena, A. Scaglione, and M. van der Schaar, “Transmitting important bits and sailing high radio waves: a decentralized cross-layer approach to cooperative video transmission,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1597-1604, Oct. 2012.

[6] H. P. Shiang and M. van der Schaar, "Online Learning in Autonomic Multi-Hop Wireless Networks for Transmitting Mission-Critical Applications," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 5, pp. 728-741, June 2010.

[7] H. P. Shiang and M. van der Schaar, "Quality-Centric TCP-Friendly Congestion Control for Multimedia Transmission," *IEEE Trans. on Multimedia,* vol. 14, no. 3, pp. 896-909, June 2012.

[8] R. S. Sutton, and A. G. Barto, “Reinforcement learning: an introduction,” Cambridge, MA:MIT press, 1998.

[9] M. Chiang, S. H. Low, A. R. Caldbank, and J. C. Doyle, “Layering as optimization decomposition: A mathematical theory of network architectures,” *Proceedings of IEEE*, vol. 95, no. 1, 2007.

[10] L. Georgiadis, M. J. Neely, and L. Tassiulas, ”Resource allocation and cross-layer control in wireless networks”, Foundations and Trends in Networking, vol. 1, no. 1, pp. 1-149, 2006.

[11] D. Rajan, A. Sabharwal, and B. Aazhang, “Delay-bounded packet scheduling of bursty traffic over wireless channels,” *IEEE Trans. on Information Theory*, vol. 50, no. 1, Jan. 2004.

[12] R. Berry and R. G. Gallager, “Communications over fading channels with delay constraints,” *IEEE Trans. Inf. Theory*, vol 48, no. 5, pp. 1135-1149, May 2002.

[13] N. Salodkar, B. Bhorkar, B. Karandikar, and V. S. Borkar, “An on-line learning algorithm for energy efficient delay constrained scheduling over a fading channel,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 732-742, 2008.

[14] D. V. Djonin, Vikram Krishnamurthy, “Q-learning algorithms for constrained Markov decision processes with randomized monotone policies: application to MIMO transmission control,” *IEEE Transactions on Signal Processing* 55(5-2): 2170-2181 (2007)

[15] D. M. Topkis, “Supermodularity and complementarity,” Princeton University Press, Princeton, NJ, 1998.

[16] E. Altman, B. Gaujal, and A. Hordijk, “Multimodularity, convexity, and optimization properties,” *Mathematics of Operations Research*, vol. 25, no. 2, pp. 324-347, May 2000.

Nicholas Mastronarde received his B.S. and M.S. degrees in electrical engineering (EE) from the University of California at Davis in 2005 and 2006, respectively and the Ph.D. degree in EE from the University of California at Los Angeles (UCLA) in 2011. He is currently an Assistant Professor in the Department of Electrical Engineering at the State University of New York at Buffalo.

Mihaela van der Schaar is Chancellor's Professor of Electrical Engineering at the University of California, Los Angeles. She is an IEEE Fellow, the Editor in Chief of IEEE Transactions on Multimedia, a Distinguished Lecturer for IEEE Communications Society, and the founding director of the UCLA Center for Engineering Economics, Learning, and Networks (see netecon.ee.ucla.edu).

An Enriched Multimedia Experience for Wireless Networks in Horizon 2020 and Beyond

David Soldani, David Virette, Giovanni Cordara, Qing Zhang
Huawei European Research Centre, Munich, Germany
{david.soldani, david.virette, giovanni.cordara, qing.zhangqing}@huawei.com

1. Introduction

Looking at horizon 2020 and beyond, we are expecting an interconnected world where virtual and augmented reality will be an integral part of our daily lives. The user/prosumer experience will be enriched through 3D interactions and immersions (virtual world, 3D vision, interactive gaming, augmented reality, collaborative multimedia, etc.), enhanced interfaces for devices (gesture, voice, etc.) and efficient, contextual and personalized search engines [1].

Future network technologies are envisioned to connect people, things, goods, processes and any source of relevant information. As illustrated in Figure 1, future carrier networks will provide seamless, high speed access to ICT resources, and full connectivity; it will be possible to perform big analytics in an open cloud ecosystem with reliable and secure hardware, software and prosumer centric applications [2].

5G mobile networks will implement new technologies and improve performance for direct Device-to-Device (D2D) and massive Machine-to-Machine (M2M) communication, ultra-dense and moving networks, and ultra-reliable communication. Hyper connectivity will be provided to everyone and to every-thing. The radio interface will be redesigned to support new waveforms, multiple accesses, enhanced multi-node coordination, multi-antenna, and multi-hop communications, as well as heterogeneous multi-layer scenarios and flexible spectrum allocation [3].

In this paper, we focus on future mobile terminals and present our view on new multimedia technologies, as the crucial ingredients for the immersive, collaborative, emotional and interactive user experience. As a part of this framework, the paper introduces the new notion of collaborative multimedia, 3D audio and video formats and evolution of the corresponding codecs. Also, this letter provides some insights into augmented reality for several applications, such as gaming, advertising, localization, and orientation. There technologies will demand for more bandwidth consumption, innovative multimedia coding technologies and specific standards. The technical feasibility and business viability of the proposed solutions for mobile terminals are also discussed in the following sessions, paying particular attention to the most promising research trends and related ongoing standardization activities.

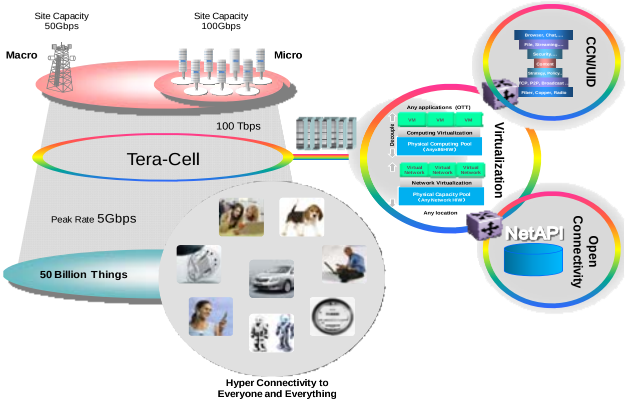


Figure 1. Wireless network and services vision.

2. Future mobile terminals

This section introduces three key enabling technologies, namely collaborative multimedia, 3D multimedia and augmented reality, which will influence the design of the next generation mobile terminals and networks.

Collaborative multimedia

As illustrated in Figure 2, future mobile terminals will support by default multi-sensors or arrays of sensors. In addition to multiple antennas for optimal connectivity, a mixture of multimedia sensors (microphone, speaker, camera, display) will allow the mobile phone to capture, process and render multimedia information in a natural, contextual, immersive and intelligent way. Furthermore, the technology evolution for a direct interaction among terminals, such as *collaborative multimedia*, will likely be one of the most important differentiating factors for an emotional multimedia wireless experience. The key technical ingredients for collaborative multimedia are: *location sensing*, *synchronization*, *content adaptation* and *natural audio/video interfaces*. Social proximity sensing gives a better estimate of location information. In systems of microphone arrays, the synchronization accuracy is a critical parameter for multi-screen rendering, augmented reality, noise reduction and acoustic echo cancellation. Content adaptation may provide a smart content adjustment to different locations, number of devices and various actions. Some use cases worth mentioning are:

- *Collaborative music playing* (stereo or 5.1 chan-nels play back by multiple phones, each phone playing one of the audio channels).

- *Collaborative mixing* (DJ mixing on a tablet, using music coming from the cloud, tablet and every phones locally connected).
- *Collaborative conferencing* (speakers may move around in the room, as their mobile phones include auxiliary microphones to capture the sound).
- *Collaborative augmented reality games* (people jointly interact with the surrounding environment).

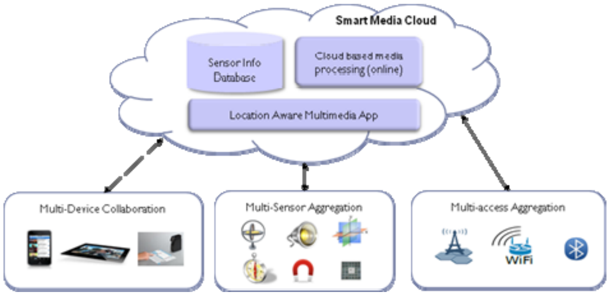


Figure 2. Collaborative multimedia vision.

3D multimedia

In the near future, it is expected that 3D multimedia will be one of the fundamental applications for mobiles. Augmented reality, 3D user interfaces, natural gesture recognition and interaction, immersive multimedia and communications, are examples of such a vast range of applications. As of today, only mono audio format is supported for communications in mobile networks, and 3D video technology is encountering difficulties in penetrating the market, especially due to the current low quality of experience with 3D displays, along with the lack of good 3D video contents. In fact, while today's 3D display technology can recreate the effect of convergence (eyes rotation in opposite directions to maintain binocular vision), other important depth cues, such as accommodation (change of focus), cannot be appropriately reproduced as the resulting image is being displayed on a flat surface. This limitation has several consequences, such as eye strain, fatigue, and temporary diplopia [4]. It is believed that a natural, "real" and not tiring viewing experience is the key enabler for a mass adoption of 3D technologies in the mobile domain. Besides this, we deem that future-proof 3D devices shall support two more features: *full motion parallax* and *multi-viewer fruition*.

Similar observations can be made for 3D audio technologies. The surround sound formats such as 5.1, 7.1 or 22.2 channels have been developed mainly for cinema or home theater applications. Indeed, while a real immersive audio rendering can be obtained for a 3D audio personal experience with earphones through binaural synthesis [5], the use of multi-speaker systems in small devices is still extremely difficult to realize due to space constraints. Furthermore, the integration of multiple transducers to capture and render 3D sound and improve the user experience will of course have a serious impact on the acoustic design of the terminal, the associated signal processing and on transmission of audio content [6]. In mobile applications, it is expected that spatial audio communications will gain momentum with the integration of stereo microphones or microphone arrays (e.g. Huawei Ascend D2 terminal supports a novel enhanced stereo microphone technology).

Augmented reality

Augmented Reality (AR) is considered as one of the most prominent mobile application of the future [7], giving users an unprecedented way to enrich real world with related information. However, current computer vision technologies for augmented reality present severe limitations, especially in latency performance and battery consumption. The former relates to the need of real-time information exchange with data centers (the cloud) for object identification, the latter is due to the continuous local execution of computer vision algorithms for tracking objects. We consider the latency aspect of particular importance to convey a graceful user experience: due to the fully interactive nature of AR applications, the experienced delay shall be even smaller than the one typically acceptable for conversational services.

3. Immersive audio

Spatial audio conferencing has recently gained strong interest from the research community. Virtual meeting room can be simulated through binaural synthesis for a more natural and immersive experience, with improved intelligibility and less fatigue. This spatialization effect is achieved through multichannel audio signals, using stereo audio codecs (parametric stereo coding or spatial audio coding) with improved performance with respect to dual mono operations [8][9]. For mobile applications the next generation of communication codecs, denoted as Enhanced Voice Service (EVS) codec, is currently defined in 3GPP [10]. The new codecs are for VoIP over LTE. Furthermore, the multichannel support will allow a real 3D audio communication with a complete audio scene representation.

3D audio format is also adopted for movie and gaming applications, using fixed and mobile terminals. Real 3D audio contents, e.g. 22.2 or sound field format, such as Higher Order Ambisonic (HOA), are emerging technologies for future broadcast services. Moving Pictures Experts Group (MPEG) has started a new activity for the standardization of a new compression format [11], addressing those new multichannel and sound field formats for intermediate bit rates (form 256 Kb/s to 1.5 Mb/s). Low bit rates targeting mobile applications will be addressed at the later time. The rendering of various loudspeaker systems will be a part of the same research framework, as heterogeneous speaker arrays must of course be considered in such application scenarios.

4. Three dimensional video

Video is the most bandwidth hunger content and 3D video applications will certainly exploit more network capacity. Integral imaging, also referred as Holoscopic imaging [12], is a promising solution to overcome the shortcomings of the current 3D displays technologies. This system makes it possible to capture multiple 2D images through and array of small lenses (lenslet array), giving a different perspective of the 3D scene. Such a multiplicity of perspectives is recreated at the display side, through a re-projection to a specular lenslet array. The stereoscopic effect is achieved thanks to the different views conveyed to the user's eyes, thus enabling real 3D experience and allowing full motion parallax. The state of the art of integral imaging is rapidly progressing [13], and we are expecting that it will be the key technology for mass-market adoption of 3D video. However, this technology will have a significant impact on the network infrastructure, due to the huge amount of data necessary to transmit the elementary images, the different images captured by the lenslets. The 3D holoscopic images will call for dedicated coding schemes, thus reducing the strongly redundant nature of the integral imaging signal. Feasible solutions for a mass adoption of the integral imaging technology are far from being ready. In facts, if we look at the standardization domain as the driving force for next generation commercial products, we can see that integral imaging is not fitting their short-medium term plans. Recently, the two mostly active bodies on the definition of enabling technologies for video coding, Video Coding Experts Group (VCEG) within ITU-T and MPEG within ISO, have defined a joint collaborative project, named JCT-3V, addressing the next generation of 3D video coding algorithms. The most futuristic technology [15], built on top of the High Efficiency Video Coding (HEVC) standard **Error! Reference source not found.**, is not addressing integral imaging, but a multi-camera/multi-depth representation, still relying on the current 3D display approach. The completion of this standard is expected by the end of 2014 and we believe that the standardization work based on real-3D and holoscopic imaging will follow.

5. AR enabling technologies

Augmented reality applications require frequent and very fast communication over the network. This is particularly critical when the object recognition is completely done on the data center side, since frequent upload of captured images is required. A better approach, extensively addressed by the research community, relies on the local extraction of relevant image descriptors. However, this only partially solves the problem, because most of the existing searching and tracking algorithms rely on local features, e.g. SIFT [16] and SURF [17]. Normally, hundreds or even thousands of features are extracted for one image. Under this assumption and the fact that, e.g., a vector of 128 bytes is used for representing one SIFT feature, the overall required bit rate is comparable to those needed for transmitting a compressed version of the raw image. Hence, future technologies of augmented reality need to complement the local features with compression schemes. The research community is already pursuing this direction, which is also addressed by a dedicated ongoing standardization activity carried out in MPEG [18]. The target is to minimize the bit rate of the compressed features extracted in one image, e.g. down to 512 bytes per image, yet maintaining an adequate level of recognition accuracy.

References

[1] B. Cardinaël, et al., "Telcos Vision for Horizon 2020", in *Eurescom mess@ge magazine*, November, 2012.

[2] FI Assembly 2012, "The Future Internet", in *Lecture Notes in Computer Science*, Springer, vol. 7281, 2012.

[3] Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society (METIS), EC FP7 funded project.

[4] T. Shibata, et al. "The zone of comfort: Predicting visual discomfort with stereo displays", *Journal of Vision*, 11(8):11, 1–29, 2011.

[5] R. Nicol, "Binaural Technology", *AES Monograph*, 2010.

[6] T. Sugimoto, K. Matsui, H. Okubo, "A loudspeaker array frame reproducing 22.2 multi-channel sound for super Hi-Vision Flat Panel Display", in Proc. *NAB Broadcast Engineering Conference*, 2012.

[7] Van Krevelen, D. W. F., and R. Poelman, "A survey of augmented reality technologies, applications and limitations", *International Journal of Virtual Reality*, vol. 9, issue 2, 2010.

[8] Y. Lang, D. Virette, C. Faller, "Novel low complexity coherence estimation and synthesis algorithms for parametric stereo coding", in Proc. *EUSIPCO*, pp. 2427 – 2431, 2012.

[9] M. Lutzky, et al. "AAC-ELD V2 - The new state of the art in High Quality Communication Audio Coding", in Proc. 131st AES Convention, October 2011.

[10] K. Järvinen, et al. "Media coding for the next generation mobile system LTE", in *Computer Communications*, Elsevier, Vol. 33, pp1946-1927, 2010.

[11] "Audio for High Efficiency Video Coding", *ISO/IEC JTC1/SC29/WG11/N11865*, January 2011.

[12] G. Lippmann, "La Photographic Integrale", C.R. Acad., Sci. vol. 146, pp. 446-451, 1908.

[13] M. Cho, M. Daneshpanah, I. Moon, and B. Javidi, "Three-dimensional optical sensing and visualization using integral imaging", Proc. *IEEE* 99, 556–575, 2011.

[14] G. Sullivan, et al. "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Trans. On Circuit and Systems for Video Technology, 2012.

[15] White Paper on "State of the Art in 3D Video", ITU-T/ISO/IEC Joint Collaborative Team on 3D Video Coding Extension (JCT-3V) document JCT3V-A1006, Stockholm, Sweden, 2012.

[16] D. Lowe, "Distinctive image features from scale-invariant keypoints", *Int. Journal of Computer Vision*, vol. 60, issue 2, pp. 91–110, 2004.

[17] H. Bay, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features", in Proceedings of *European Conference on Computer Vision (ECCV)*, Graz, Austria, 2006.

[18] M. Bober, G. Cordara, S. Paschalakis, G. Francini, V. Chandrasekhar, K. Iwamoto, "Text of Working Draft 1: Compact Descriptor for Visual Search", ISO/IEC/MPEG document w12928, Stockholm, Sweden, 2012.

David Soldani received a M.Sc. degree with maximum score and *cum laude approbatur* in Electronic Engineering from the University of Florence, Italy, in December 1994; and a D.Sc. degree in Technology with distinction from Aalto University, Finland, in October 2006.

In September 2007, he was qualified for Professorship on Communication Networks from the same university. He has been in the ICT industry for more than 18 years. From 1997 to 2007 he was at Nokia in various technical and research management positions. Prior to joining Nokia, he was a research engineer in Rohde & Schwarz and Sirti SpA, Milan, Italy; and a graduated officer at the Italian Military Navy, Livorno, Italy. He is currently Vice President of the Huawei European Research Centre (ERC) and Head of Central Research Institute (CRI) at ERC, Munich, Germany. In his current role, he is leading the "*research to innovation via standardization on Future Networks and Services"* in Europe. His areas of technical expertise include mobile broadband networks, transport network layer and wireline broadband access technologies, and future network technologies.

David Virette received a Dipl.-Ing. degree in electronics and computer science from the ENSSAT, Lannion, France and an M.Sc. in telecom and signal processing from the University of Rennes, France (2000). From 2000 to 2009, he worked as Researcher with France Télécom R&D, France. Since 2009, he has been working as Senior Research Engineer with Huawei European Research Center (ERC), Munich, Germany. He is now leading the Audio Technologies team in ERC. He has contributed to the standardization of speech and audio coders in MPEG and ITU-T, namely AAC-ELD, MPEG Surround, USAC, G.729.1, G.718, G.711.1 Annex D,E,F and G.722 Annex B,C,D. His main research interests include 3D audio, speech/audio coding and processing.

Giovanni Cordara received a Dipl.-Ing. in Telecommunications from the Polytechnic of Turin, Italy (2000). In 2001 he joined Telecom Italia Lab, where he conducted and coordinated research on computer vision and video coding.

In 2011, he joined the Huawei European Research Center (ERC), Munich, Germany. He is now leading the Video Technologies team in ERC. Since 2003, he has contributed to the MPEG standardization activities; he is currently co-chair of the Ad hoc Group on Visual Search and editor of the emerging "Compact Descriptors for Visual Search" standard (MPEG-7, part 13). His research interests related to computer vision, 3D video processing, 3D reconstruction and video coding.

Qing Zhang received his B.Sc. and M.Sc. from Peking University and Xi'an Jiaotong University, in China, in 1987 and 1996, respectively. Since 1996, he has been in Huawei for 16 years, leading the product development of video conferencing, video gateway, the multimedia product and technology planning, technology research on audio coding and processing. He is currently deputy director of Media technology lab, in charge of the technology planning, multimedia communication research and leads the ARTIST program, focused on the next generation multimedia technology for mobile device. He has been involved in several standardization activities of ITU-T, MPEG and 3GPP; contributed to video telephony, multimedia conference, and audio codec standards. In 2009, he received a national award for science and technology achievement. His research interest includes multimedia communication, collaborative multimedia and audio processing. He owns more than 60 patents.

Towards QoE-Driven Resource Allocation for Multimedia Services in LTE

Krunoslav Ivešić, Lea Skorin-Kapov and Maja Matijašević
University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{krunoslav.ivesic, lea.skorin-kapov, maja.matijasevic}@fer.hr

1. Introduction

Multimedia services are, in general, composed of several media components (e.g., audio, video, data, 3D graphics), the number of which, as well as their properties and resource requirements, may vary over time. For example, a media flow may be started, stopped, or modified to adapt to current network conditions or user preferences. In a wireless network environment, the efficient use of limited radio resources is crucial in meeting the end user Quality of Experience (QoE) expectations regarding multimedia service usage. Although QoE metrics focus on subjective user quality perception, an appropriate mapping to system related characteristics and quantitative network performance parameters is necessary for designing the mechanisms in the network to actually deliver the desired QoE.

Utility functions (UF) have been studied in the context of QoE modeling [1] and QoE-driven optimized resource allocation [2][3], as the mathematical means to specify the relation between user perceived value and the consumption of network resources. QoE can change over time, given the variations in QoE influencing factors (e.g., network conditions, service state, user preferences, usage context, etc.). UFs to be considered for making resource allocation decisions in networks may hence take on different forms during a service's lifetime, for example, due to different active media flows, different delay/loss values, change in user terminal capabilities, etc.

For sessions involving multiple media components (e.g., audio and video), different UFs may correspond to different components, with overall session utility expressed as a weighted combination. Utility-based multimedia adaptation decision-taking has been previously applied in the scope of the MPEG-21 Digital Item Adaptation standard, and further addressed in the scope of multi-modal media streams [4]. A key issue in making QoE-driven multimedia adaptation and resource allocation decisions is consideration of individual user network/device capabilities and service preferences, such as, e.g., indicating relative importance of individual streams (comprising a single session) such as audio and video [5].

In this position paper we present and discuss a solution for QoE-driven resource allocation of multimedia services, based on the negotiation and calculation of a per-session *Media Degradation Path* (MDP). The MDP is specified as an ordered collection of feasible session configurations, ordered according to achievable utility calculated based on UFs. Considering a given network domain, we discuss applicability of the MDP in making admission control decisions. Session MDPs are further used as input for a mathematical model for optimized resource allocation among multiple multimedia sessions, based on maximization of overall system utility and operator's profit subject to resource constraints. With respect to ongoing work, we present a simulator tool we developed to evaluate the proposed model, and discuss how our solution is being studied in the context of a 3GPP-based Evolved Packet System (EPS) network environment supporting delivery over Long Term Evolution (LTE) wireless access.

2. QoE-driven resource allocation

A. Session parameters negotiation and calculation of a Media Degradation Path

We have introduced an enhanced per-session application-level quality matching and optimization functionality into a 3GPP based core network by way of a QoS Matching and Optimization Application Server (QMO AS) [6]. In a 3GPP network scenario, the QMO AS may be invoked as a service enabling function included along a session negotiation signaling path. We focus on a cross-layer approach, whereby we jointly consider parameters collected at different layers (e.g., application level data, network level data).

The optimization process considers user-related parameters, (specified in a *user profile*), such as service/media preferences (e.g., "audio is more important than video"), access network, and device capabilities; service requirements and adaptation capabilities (specified in a *service profile*) (e.g., support for different codecs, resolutions, frame rates, etc.); and the *operator policies*. These parameters represent important QoE influence factors to be taken into account when optimizing QoE.

The *service profile* builds upon utility mappings for a given multimedia service (comprised of multiple media components). It specifies the service resource requirements utilizing UFs as related to service configuration parameters such as, e.g., type of media flows, encodings, resolution, etc. To be used in session negotiation, the service profile may be signaled from an application server to the core network, or it can be retrieved from a service profile repository.

The QMO AS is invoked at session initiation and it gathers the input parameters related to the user profile, the service profile, and the operator policies. Service requirements are matched with signaled (or retrieved) user parameters (preferences, requirements, capabilities) specified in a user profile, access network capabilities, and operator policy to determine feasible service configurations. This is followed by a utility-based optimization process used to determine the optimal service parameters and requested resource allocation (referred to as the *optimal configuration*) for the given service session. The optimal configuration specifies media flow operating parameters (e.g., frame rate, codec), resource requirements (e.g., bit rate, delay, loss), and a utility value that represents a numerical estimation of the configuration's QoE. Besides the optimal configuration, several suboptimal configurations are calculated and ordered by decreasing utility value, thus forming a *Media Degradation Path* (MDP) (see example in Fig. 1).

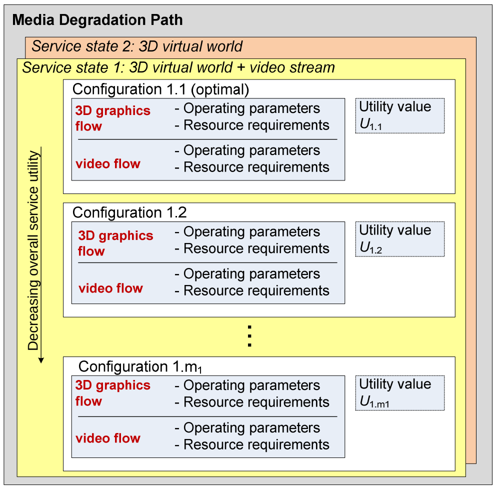


Figure 1. Example MDP structure

The goal of the MDP is to serve as a “recipe” for controlled service adaptation (including degradation, if needed), by achieving maximum utility in light of dynamic conditions. For example, for a user indicating that he/she prefers audio over video for a given audiovisual service, an MDP may be constructed so as to first degrade video quality in case of a decrease in network resources availability, while maintaining high audio quality. Hence, in case of decreased resource availability, the second-best (or the next-best) sub-optimal configuration can be activated, thus preventing unpredictable degradation of a service. Since the media components of the service are not necessarily all active at the same time, the MDP configurations are grouped by the *service state* they pertain to, whereby the service state refers to a set of service components, which are simultaneously active during a given time period.

B. Optimized resource allocation for multiple sessions

Calculated per-session MDPs may be passed on to a resource and admission control entity responsible for making domain-wide resource allocation decisions. (We note that possible dynamic changes in, e.g., user preferences or device being used, may lead to recalculation of the MDP).

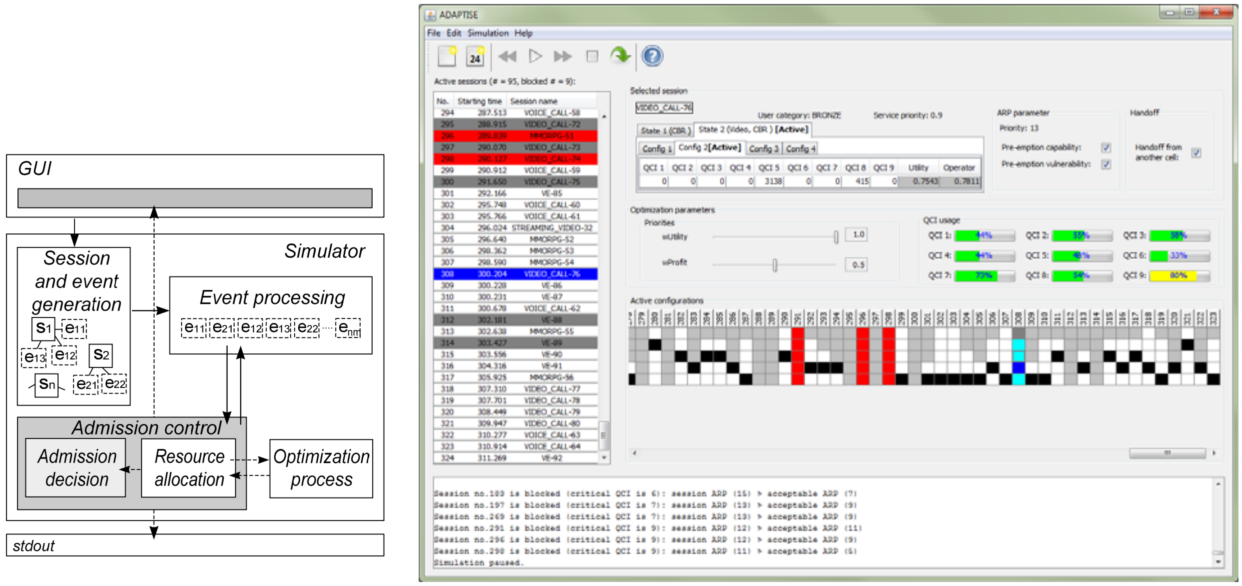
Calculating the optimal resource allocation across multiple sessions has been formulated as a multi-objective optimization problem with the objectives of maximizing the total utility of all active sessions described by their MDPs, along with operator profit [2]. This corresponds to a multi-choice multidimensional 0-1 knapsack problem (MMKP). The problem formulation has been given in the context of the 3GPP Evolved Packet System (EPS), which maps session flows to one of nine standardized QoS Class Identifiers (QCIs). QCIs as standardized by the 3GPP define different types of standardized packet forwarding behavior. It may be noted as well that this is but one possible problem formulation, focusing on discrete optimization and assuming a weighted combination of multiple objectives.

3. Simulation

In order to evaluate the proposed QoE-driven resource allocation model, we have developed a tool called ADAPTISE (short for *ADmission control and resorce Allocatation for adaPtive mulTImedia SErvices*), which simulates arrivals, durations, resource allocation decision-making, and state changes of multimedia sessions [7]. We have considered several multimedia service types (voice and video call, gaming, streaming and virtual environment) and specified the corresponding inter-arrival time and duration distribution parameters. The simulation is event-based, with all events pertaining to sessions (e.g., session start, session state change) generated at the beginning of the simulation and stored in an event queue (Fig. 2a). The events are then processed one by one and the required action regarding resource allocation is conducted.

Constraints with regards to resource availability may be set as simulation parameters. As we assume each media flow to be assigned a certain QCI depending on its traffic requirements, for simplicity reasons we set constraints regarding the aggregate bandwidth that may be allocated to all flows in a given QCI.

Furthermore, given that we are considering a multi-objective optimization problem (involving max. of overall system utility and operator profit), different weight factors may be assigned to each of the objectives.



Active sessions (# = 95, blocked # = 9):

No.	Starting time	Session name
294	287.513	VOICE_CALL-58
295	288.915	VIDEO_CALL-72
296	289.839	MMORPG-51
297	290.070	VIDEO_CALL-73
298	290.127	VIDEO_CALL-74
299	290.912	VOICE_CALL-59
300	291.650	VIDEO_CALL-75
301	292.166	VE-85
302	295.748	VOICE_CALL-60
303	295.766	VOICE_CALL-61
304	296.024	STREAMING_VIDEO-32
305	296.640	MMORPG-52
306	298.362	MMORPG-53
307	298.590	MMORPG-54
308	300.204	VIDEO_CALL-76
309	300.228	VE-86
310	300.231	VE-87
311	300.678	VOICE_CALL-62
312	302.181	VE-88
313	302.638	MMORPG-55
314	303.427	VE-89
315	303.556	VE-90
316	304.316	VE-91
317	305.925	MMORPG-56
318	307.310	VIDEO_CALL-77
319	307.701	VIDEO_CALL-78
320	308.449	VIDEO_CALL-79
321	309.947	VIDEO_CALL-80
322	310.277	VOICE_CALL-63
323	310.914	VOICE_CALL-64
324	311.269	VE-92

Selected session

VIDEO_CALL-76

User category: BRONZE

Service priority: 0.9

State 1 (CBR) | State 2 (Video, CBR) **[Active]**

Config 1 | Config 2**[Active]** | Config 3 | Config 4

QCI 1	QCI 2	QCI 3	QCI 4	QCI 5	QCI 6	QCI 7	QCI 8	QCI 9	Utility	Operator
0	0	0	0	3138	0	0	415	0	0.7543	0.7811

ARP parameter

Priority: 13

Pre-emption capability: ☑

Pre-emption vulnerability: ☑

Handoff

Handoff from another cell: ☑

Optimization parameters

Priorities

wUtility 1.0

wProfit 0.5

QCI usage

QCI 1: 44%
QCI 2: 55%
QCI 3: 58%
QCI 4: 44%
QCI 5: 48%
QCI 6: 33%
QCI 7: 73%
QCI 8: 54%
QCI 9: 80%

Active configurations

```
Session no.183 is blocked (critical QCI is 6): session ARP (15) > acceptable ARP (7)
Session no.197 is blocked (critical QCI is 7): session ARP (13) > acceptable ARP (9)
Session no.269 is blocked (critical QCI is 7): session ARP (13) > acceptable ARP (9)
Session no.291 is blocked (critical QCI is 9): session ARP (12) > acceptable ARP (11)
Session no.296 is blocked (critical QCI is 9): session ARP (12) > acceptable ARP (9)
Session no.298 is blocked (critical QCI is 9): session ARP (11) > acceptable ARP (6)
Simulation paused.
```

(a) Block diagram (b) Graphical user interface

Figure 2. ADAPTISE simulator

At session arrival, the MDP-based admission control mechanism is invoked and the session is either admitted with one of the available configurations from its MDP, or rejected. This option increases admission probability as opposed to the case when a session is either admitted with an optimal configuration or rejected. The admission decision depends on the available resources, service type, user priority and the origin of the session (new or handed off from neighboring cells).

The optimization process responsible for optimal resource (re)allocation is triggered when resource consumption surpasses a predefined threshold (e.g., 95% available resources). The event resulting in an increase in resource consumption includes change of state within an active session, e.g., addition of a new flow to the session.

The optimization problem is solved using heuristics. Resource reallocation results in some session configurations being switched to less resource consuming configurations, thus decreasing utility (while maintaining utility above an acceptable utility threshold), but making it possible to support new flows.

The ADAPTISE graphical user interface is depicted in Figure 2b. The list of sessions and their starting times is displayed in the table (on the left) and updated as new sessions appear. The configurations from the sessions' MDPs are displayed as a matrix of squares (in the middle) where black squares indicate enforced configurations and white squares pertain to other configurations from the regarding session's MDPs. Having clicked on a session in the table or in the matrix, the details about the session appear, displaying the data about states and flows. The resource consumption ("QCI usage") for the nine QCIs is represented by respective gauges.

4. Applicability in LTE

We foresee the proposed solution as being applicable in the 3GPP core network, with the QMO AS included as an AS in the signaling domain, and the multi-session resource allocation algorithm being run on a policy server (e.g., Policy and Charging Rules Function). In order to validate the applicability and performance of the proposed solution assuming services delivered via an LTE network [8], our current work is focusing on testing the radio resource utilization and resulting end user QoE assuming our algorithm being used to shape downlink traffic. For that purpose we are using the publicly available LTE-Sim simulator [9], developed at the Politecnico di Bari, Italy.

The idea is to run simulations by using ADAPTISE in two scenarios: (a) with MDP-based admission control, and (b) with admission control rejecting sessions entirely if the best configuration cannot be enforced, resulting in lower admission probability. By loading simulation traces into LTE-Sim, it will be possible to observe the resulting QoS parameters (e.g., delay and loss, obtained by parsing the LTE-Sim simulation output) and map them to QoE levels by using UFs. The goal is to study the tradeoffs between the increased number of admitted sessions (resulting from MDP-based admission control) and resulting QoE levels.

With regards to the optimization algorithm, we plan to collect simulation traces prior to invocation of the algorithm (showing the network in a congested state), and after the algorithm has been run (resulting in service adaptation to free network resources, while maintain QoE at acceptable levels). By loading these traces into LTE-Sim, with cell bandwidth set to the resource constraints corresponding to those used in the ADAPTISE simulation, it will be possible to compare session QoS parameters before/after optimization invocation and map them to resulting QoE levels.

We note that a current limitation of LTE-Sim is the lack of support for packet scheduling at the base station (eNodeB) account for the QCI assigned to a given flow. Such scheduling algorithms are not standardized by 3GPP and are considered vendor specific.

5. Conclusion

We have presented an idea and an ongoing work towards a solution for a domain wide QoE-driven resource allocation of multimedia services based on utility functions and negotiated per-session MDPs. We introduce the simulation tool to evaluate the proposed model, considering the notion of 3GPP EPS bearer level QoS. We further discuss ongoing work related to validating the proposed approach in an LTE access network by using our simulation results as input for the publicly available LTE-Sim simulator.

References

[1] P. Reichl, B. Tuffin, R. Schatz. "Logarithmic laws in service quality perception: where mi-creconomics meets psychophysica and quality of experience," *Springer Telecomm. Systems*, pp. 1-14, June 2011.

[2] K. Ivešić, M. Matijašević, L. Skorin-Kapov, "Utility Based Model for Optimized Resource Allocation for Adaptive Multimedia Services", *21st International Symposium PIMRC*, pp. 2638-2643, 2010.

[3] M. Shehada, S. Thakolsri, Z. Despotovic, W. Kellerer. "QoE-based Cross-Layer Optimization for Video Delivery in Long Term Evolution Mobile Networks," *14th International Symposium WPMC*, pp.1-5, 2011.

[4] M. Prangle, T. Szkaliczki, and H. Hellwanger, "A Framework for Utility-Based Multimedia Adaptation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 6, pp. 719–728, June 2007.

[5] L. Skorin-Kapov, M. Matijašević. "Modeling of a QoS Matching and Optimization Function for Multimedia Services in the NGN", *12th IFIP/IEEE MMNS*, vol. 5842, Springer, pp. 54-68, Oct. 2009,

[6] L. Skorin-Kapov, M. Mošmondor, O. Dobrijević, M. Matijašević, "Application-level QoS Negotiation and Signaling for Advanced Multimedia Services in the IMS", *IEEE Communication Magazine*, vol. 45, no. 7, pp. 108-116, 2007.

[7] K. Ivešić, M. Matijašević, L. Skorin-Kapov, "Simulation Based Evaluation of Dynamic Resource Allocation for Adaptive Multimedia Services", Proc. of the *7th CNSM*, pp. 1-4, October 2011.

[8] 3GPP, Tech. Specif. Group Radio Access Network, "Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)", 3GPP TS 25.913.

[9] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi and P. Camarda, "Simulating LTE Cellular Systems: an Open Source Framework", *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, February 2011.

Krunoslav Ivešić received a Masters Degree in Electrical Engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing, in Croatia. He is currently working towards his PhD in the area of admission control and resource allocation algorithms for multimedia services next generation wireless networks.

Lea Skorin-Kapov is an Assistant Prof. at the Univ. of Zagreb, Faculty of Electrical Engineering and Computing, in Croatia, where she also received her PhD degree in 2007. She was previously employed as a research engineer at Ericsson Nikola Tesla, Zagreb working in the area of QoS signaling and negotiation. Her current research interests are focused towards modeling QoE and QoE optimization strategies. She is a member of the EU COST Action QUALINET.

Maja Matijašević is a Professor at the University of Zagreb, Faculty of Electrical Engineering and Computing, in Croatia. Her research interests include networked multimedia and QoS in IP-based next generation networks, with particular focus on session negotiation, adaptation, and mobility. She received her Dipl.-Ing, M.S. and Ph.D. degrees in Electrical Engineering from the Univ. of Zagreb and the M.Sc. in Computer Engineering from the Univ. of Louisiana at Lafayette, LA, USA. She is a Senior Member of IEEE and a member of ACM.

Radio Resource Management for Video Traffic in IEEE 802.16 Networks

Marwen Abdennebi [†] *and Yacine Ghamri-Doudane* [‡]
marwen.abdennebi@univ-paris13.fr, yacine.ghamri@ensiie.fr

[†] *Laboratoire de Traitement et de Transport de l'Information (L2TI)*
University of Paris Nord

[‡] *Université Paris-Est, Laboratoire d'Informatique Gaspard Monge*
(LIGM – UMR 8049)
[‡] *École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise*
(ENSIIE)

1. Introduction

Radio resources management includes fundamental functions for the radio interface that ensure a good Quality of Service (QoS) for mobile users. This is especially important for video streaming as it constitutes one of the most exigent services, in terms of bandwidth, latency and bitrate variability.

As the radio spectrum is limited, radio resource management can be seen as a better usage of this bandwidth while keeping a good QoS for users. Hence, the question is: how to use the minimum amount of resources to transmit the maximum amount of traffic while keeping an acceptable QoS. In this purpose, mainly two strategies can be established.

First, the Call Admission Control (CAC) will ensure that the traffic accepted never exceeds a maximum threshold to keep the QoS acceptable. This can be seen as macroscopic resource management, i.e., at the call/connection time level. For video, the difficulty comes from the generated traffic behavior and the heterogeneity of flows. Then, the second step, the allocation inside the radio frames takes place, connection by connection. This allocation must be adapted to the traffic generated by the video sources, packet by packet or at least burst by burst, while ensuring good QoS for users.

Many works have been done in order to enhance video transmission. For example, work in [1] deals with CAC function and propose to adapt CAC parameters for an accurate decision. However, those proposals do not take account of the different types of flows accepted at the same time. Others [2] [3] deal with microscopic level in order to obtain better scheduling but do not ensure that QoS is not affected.

In this paper, we present two of our proposed solutions for intra-cell radio resource management optimization. We focus on IEEE 802.16 [4] as the technology is well implemented on the existent simulators and as it is the nearest one to the LTE/LTE-A 4G radio interface. Our proposed functions mainly consist in enhancing the QoS of users while optimizing the intra-cell bandwidth utilization, i.e., reducing the resources wasting.

Our solutions will then take place at both the macroscopic and the microscopic levels, i.e., at the connection level as well as at the packet/burst level; they are presented respectively in section 2 and section 3. Section 4 then presents the performance evaluation of our proposed solutions, and finally section 5 gives our concluding remarks.

2. Adaptive Call Admission Control

An appropriate CAC is essential to ensure the QoS for accepted sessions. It must be performed jointly with the dimensioning in order to maximize the GoS (Grade of Service) in term of blocking rate for all users. CAC must be based on traffic model to evaluate the bandwidth consumed by all accepted sources.

However, video traffics are highly heterogeneous in term of bandwidth consumption which makes the CAC function imprecise. We then propose to classify transmitted video flows according to the generated bitrates and to adapt the CAC according to the repartition of the traffics among each defined category. Multi-models are then used, and flow measurements are performed by the IEEE 802.16 Base Station (BS) on each video flow to classify the traffic, and hence update the traffic based CAC function.

Our proposal concerned basic AVC video [5] and enhanced SVC video traffics [6] as well. For SVC traffic [7], measurements and the whole process must be done for each layer of accepted flow. The complete process for AVC is described in figure 1.

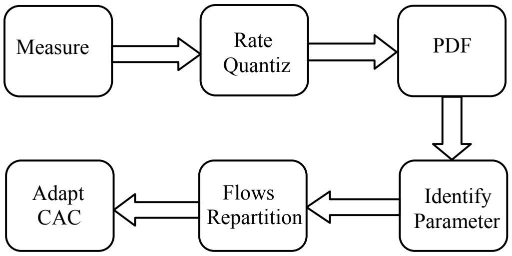


Figure 1. Adaptive CAC process description

Mapping with the appropriate source model is done for each measured flow after rate quantization and parameters identification (using an LMS, least mean squares). The steps of the process are briefly described in the sub-sections below.

A. Measurements and Parameters Identifications

From this step we will get the flows repartition among the defined categories. The measurements are performed by the BS to fit the flow parameters with one of the defined models. Three traffic models are considered, one for high rates videos (games, action movies, …), another for medium rates (movies, shows, …), and the last one for low rates (talk, news, …). As we considered H264 video sources, we used a rate based Markovian model as recommended in [8].

The update and measurement time period is set to some hundreds of ms, according to the GOP length which is in the time scale of the 802.16 rtPS (real time Polling Service) allocation period.

For each flow, after obtaining the quantized bitrate values *i.D*, (*i* from 0 *to M,* where *M.D* is the maximum bitrate), we compute the measured mean value *E[D]*, the variance *Var[D]*, and the correlation function *corr*. Then, the identification can be done with the respective theoretical values given by each model (obtained after MMPP calculations) to find the model that fits with the measured flow.

Figure 3 illustrates the measurements, the quantization and the fitting steps for the mean rate values *E[D]*. This last step took only a few hundreds of seconds (convergence time) which is short enough compared to the session duration.

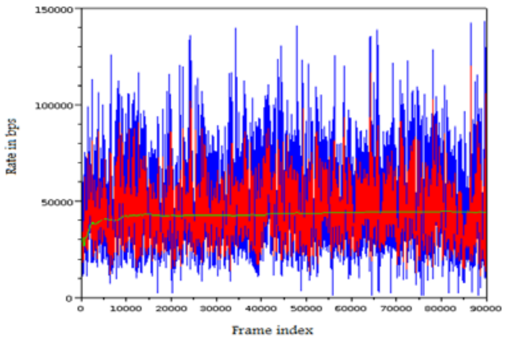


Figure 3. Example of bandwidth (blue), quantization (red) and convergence of the mean value (green) *Vs* the frame number

B. CAC adaptation and update

Once a new repartition obtained, the CAC function is updated according to the QoS policy and the CAC decision will depend on the aggregated traffic behavior. If *N1, N2, N3* represent the number of sources in each category *M1*, *M2* and *M3* respectively, then the CAC function will give the total maximum number *N* of flows that the system can accept.

If η_1, η_2, η_3 represent the relative number (in percent) of sources respectively in each category M1, M2, M3:

$$\begin{cases} N1 = N.\eta_1 \Rightarrow N = \dfrac{N1}{\eta_1} & (1) \\ N = \dfrac{N2}{\eta_2} & (2) \\ N = \dfrac{N3}{\eta_3} & (3) \end{cases} \quad \text{where } N1+N2+N3 = N$$

The mean bandwidth needed is then computed according to the probability $\pi_{i,j,k}$ of each *i.D* repartition:

$$BW = \sum_{i=1}^{N1}\sum_{j=1}^{N2}\sum_{k=1}^{N3}\pi_{i,j,k}.[rate(i)+rate(j)+rate(k)] \qquad (4)$$

The maximum bandwidth BW_{max} (in term of the maximum of equivalent admitted sources) is related to the QoS threshold and to the probability value $\pi_{i,j,k}$ of each state.

$$prob(BW_{\max}) = \sum_{i=1}^{N1}\sum_{j=1}^{N2}\sum_{k=BW_{\max}-(i+j)}^{N3}\pi_{i,j,k}.rate(i,j,k) < Thres \qquad (5)$$

Then, the maximum number of flows that could be accepted in each traffic category can be computed according to probability of each state representing $\{\eta_1, \eta_2, \eta_3\}$. We used a 3D Markov chain that describes the aggregated traffic system and then the probability values $\pi_{i,j,k}$ are deduced using the well known steady state and the conservation state equations. This, combined to (4) and (5), gives the CAC decision for our system.

3. Long-Term Radio Resources Reservation

In this section we present our solution [9] for enhancing radio resource allocation for video streaming. This could be done in addition to the adaptive CAC function, and will result in a better bandwidth usage. The proposal is based on the observation that video image frame sizes are predictable and that their impact on video quality is different depending on the image frame type.

Hence, we proposed a "long-term" radio resource allocation based on image frame size prediction for those "important" pictures. The "long-term" aspect consists in an advance reservation mechanism that guarantees the availability of resources for *I*-frames as they are the reference for encoding the *I, P*, and *B* images in the H264 GOP (Group of Pictures). This will result in a reduced delay and/or reduced drop rates for those important frames and hence in a better QoS.

We applied our proposal to the rtPS reservation mechanism of 802.16 system. This system is based on regular request/grant mechanism that updates the reservation/allocation to fit with the video scenes bitrate variations. However, rtPS suffers bad performance if congestion, i.e., if too many sources request large bandwidth at the same time. Our proposed solution will then give priority to important frames through the long-term reservation mechanism.

Figure 4 illustrates the mechanism of the combined short-term, $R(t)$, and long-term requesting $R^P(t)$, the bandwidth needed BW in the 802.16 MAC frames and the "long-term" bandwidth prediction BW^p (sent at frame time t + 1). In this example, we set the number of 802.16 frames per bandwidth request to two in order to describe more clearly the proposed mechanism.

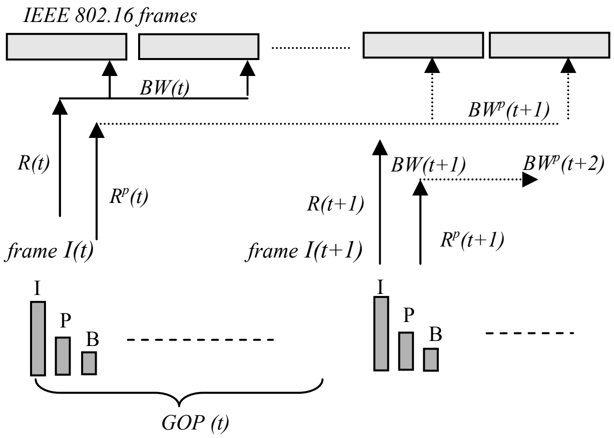


Figure 4. Access and requesting on the 802.16 MAC layer

As illustrated by the figure 4 above, the mechanism is based on *I* image-frames prediction. For each incoming frame, the rtPS request/grant mechanism is used; however, in addition to that, "long-term" reservation/provision is done for expected next I frame. The implications of this proposal are potential errors related to the prediction used. Errors must be computed as they can induce losses or bandwidth wasting. The prediction error $\varepsilon(t)$ for each frame gives the bandwidth difference between BW and BW^p :

$$\left|BW^{p}(t)-BW(t)\right|=\varepsilon(t) \qquad (6)$$

Auto-regressive models [10] are used to predict the next frame size $Y(t+1)$ using p previous frames.

$$Y(t+1)=c+\sum_{i=0}^{p-1}\gamma_i Y(t-i)+\varepsilon(t) \qquad (7)$$

$(Y_k)_{k>0}$ is a discrete stationary process representing the size of the k^{th} I-frame, γ_i the parameter of the model, and c a constant (offset). For those regressive prediction models, the order p can be an important parameter for errors. The higher p is, the more known frames are included to compute the next frame. This can have a positive effect if frames belong to the same scene as well as a negative effect if they do not.

We then propose a prediction combined to scene change detection. It is based on an autoregressive model with an adaptive prediction order p that is adjusted according to the scene changes. The scene changes are easily detected using a threshold for relative rates variations.

4. Performance evaluation

In this section, we evaluate the performances of our 2 proposed solutions, the adaptive CAC and the long-term reservation mechanism. We used the WiMAX module for NS-2 simulator adding the long-term reservation mechanism. H264 video flows downloaded from [11] are injected as traffic traces. We used the *evalvid* tool [12] which provides different encoding and allows quality evaluation (like PSNR, Peak Signal to Noise Ratio). The number of frames per seconds was set to 25 fps, with 10 image-frames per GOP, this for different kinds of video streams, high, medium and low rates. For the 802.16 rtPS parameters, we set the number of 802.16 MAC frames to 5 per bandwidth request, with a MAC frame duration of 20 ms. This leads us to set the duration of the quantization rate to the time scale of 100 ms for the traffic measurements.

A. Adaptive CAC testing

First, we show the benefits of the proposed adaptive CAC. The testing is done changing the incoming flows categories repartition as it could happen in real networks: low rates at the beginning of the evening (talk, news, …), medium and high rates at the end of the evening (movies and soccer games for example).

The 3 categories traffic models considered have the following characteristics. As explained in section 2, each model is Markovian, rate based, and is characterized by its mean arrival and service rates λ and μ (which are exponentially distributed).

Model M1	$\lambda = 0.8$ $\mu = 2.7$	Mean rate = 1.0 Mbps
Model M2	$\lambda = 0.6$ $\mu = 2.1$	Mean rate = 0.7 Mbps
Model M3	$\lambda = 0.5$ $\mu = 1.7$	Mean rate = 0.56 Mbps

We start the testing with different repartition (from 1 to 5) and hence, various values for η_1, η_2, η_3 : {10%, 30%, 60%}, {20%, 30%, 50%}, {30%, 30%, 40%}, {30%, 50%, 20%}, {20%, 60%, 20%}, we compare the adaptive CAC to the case the where the CAC takes into account only the total value *N*.

Figure 5 gives the obtained results in term of admission for each ratio for both adaptive and non adaptive CAC.

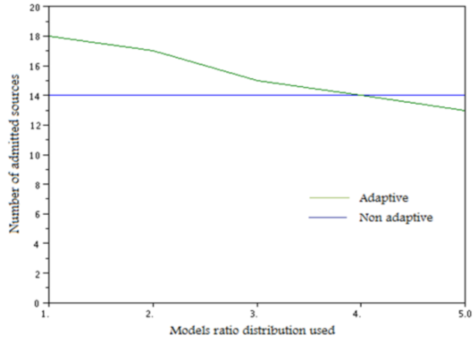


Figure 5. Admission N for adaptive and non adaptive CAC

As we can see, the system was able to admit more (enhanced GoS) sources when possible, without QoS degradation. At the same time, for the last distribution, the CAC system reduced the admission in order to avoid the drop rates and hence the QoS degradation.

B. Long-term reservation mechanism

In this sub-section we show the performance of the long-term reservation mechanism. We demonstrated in [8] that the waste due to prediction will not exceed a mean value of some Kbps thanks to our predictive algorithm that combines order adaptation and scene detection. Figure 6 illustrates the efficiency of this algorithm in term of bandwidth errors. Based on this prediction, the long-term reservation/grant algorithm is able to reduce loss rates for I frames, and improve the PSNR for final user. The figure 7 below shows this improvement compared to the basic rtPS scheme.

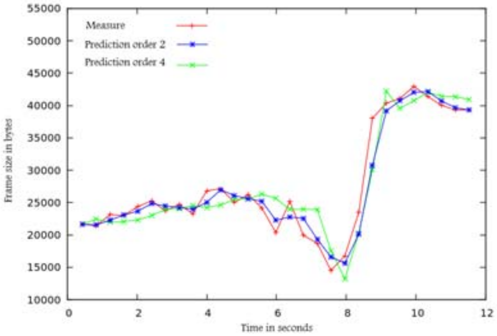


Figure 6. I frames sizes predictions with scene detection

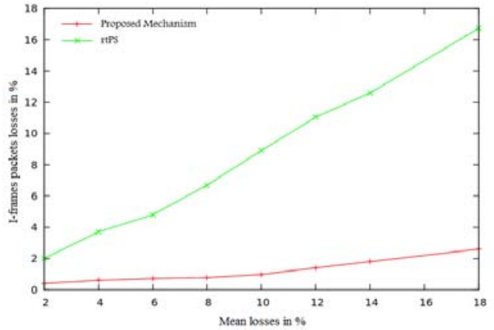


Figure 7. Relative loss rates for *I* frames in percent

5. Conclusion

This paper summarizes our two proposals for intra-cell resource management in broadband radio networks for video traffic. We first proposed an enhanced multi-model CAC function based on measurements and then a prediction based "long-term" bandwidth reservation scheme. Performance evaluation showed the positive impact of the two solutions on both GoS and QoS. We were particularly interested in IEEE 802.16 interface; we are now working on extending this work to LTE/LTE-A technologies.

References

[1] Al Masri et al, "Proposal of a Novel Bandwidth Management Framework for IEEE 802.16 Based on Aggregation", *NTMS*, May 2009.

[2] Y. Le, Y. Wu, D. Zhang, "An Improved Scheduling Algorithm for rtPS Services in IEEE 802.16", *VTC Spring*, 2009.

[3] Bianco, C. F. Chiasserini, E. Leonardi, "Bandwidth Allocation for Video Streaming in WiMAX Networks", *IEEE VTC 2009*, Barcelona, April 2009.

[4] IEEE Std 802.16e, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems", 2009.

[5] M. Abdennebi, Y. Ghamri-Doudane and Y. Li, "Adaptive CAC with Traffic Flows Classification for IEEE 802.16 Networks", *IEEE-IFIP Wireless Days 2010*, Venice, Italy, October 2010.

[6] M. Abdennebi, Y. Ghamri-Doudane, "Adaptive CAC for SVC Video Traffic in IEEE 802.16 Networks", to appear in *IEEE GLOBECOM'12,* USA, December 2012.

[7] P. Li, H. Zhang; B. Zhao, S. Rangarajan, "Scalable Video Multicast With Adaptive Modulation and Coding in Broadband Wireless Data Systems", *IEEE/ACM Transactions on Networking*, February 2012.

[8] H. Koumaraset al, "A Markov Modified Model of H.264 VBR Video Traffic", *IST Mobile Summit* 2006

[9] M. Abdennebi, Y. Ghamri-Doudane, "Long – Term Radio Resource Reservation in IEEE 802.16 rtPS for Video Traffic", Global Information and Infrastructure Symposium, GIIS 2011, Vietnam, August 2011.

[10] L. Lanfranchi and B. Bing, "MPEG-4 Bandwidth Prediction for Broadband Cable Networks", *IEEE Transactions on Broadcasting*, Fall 2008.

[11] Martin Reisslein, Lina Karam and Patrick Seeling, "H.264/AVC and SVC Video Trace Library: A Quick Reference Guide", http://trace.eas.asu.edu, May 2009.

[12] C. Ke, C. Shieh, W. Hwang, A. Ziviani, "An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission", *Journal of Information Science and Engineering*, 2008.

Marwen ABDENNEBI received his M.Sc. Degree in Telecommunications from University of Carthage, Tunisia (2002) and his PhD Degree in Computer Science (2008) from the University of Versailles, France. He worked as an application engineer at STMicroelectronics, later as an Associate Researcher at the CNRS PR*i*SM Laboratory and ENSIIE Evry. He is currently Assistant Professor at the University of Paris Nord (France) and researcher in the L2TI laboratory. His research interests are in broadband radio B3G networks, WiMAX and LTE/LTE-A.

Yacine GHAMRI-DOUDANE received an engineering degree in computer science from the National Institute of Computer Science (INI), Algeria (1998), an M.S. degree from the National Institute of Applied Sciences (INSA), France, a Ph.D. degree in computer networks from the Pierre & Marie Curie University, Paris 6, France and a Habilitation to Direct Research in computer science from Université Paris-Est (1999, 2003 and 2010, respectively). Yacine is currently Associate Professor at the Ecole Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise (ENSIIE), a major French post-graduate school located in Evry and member of the Computer Science Laboratory of the Gaspard Monge Institut at Marne-la-Vallée. His research interests include Wireless Sensor Networks, Vehicular Networks, QoS in WLAN/WMAN, TCP and Multimedia over Wireless, Mobility Management in 4G Mobile Networks, Management of Wireless/Mobile Networks.

Watermarking Technologies in Wireless Multimedia Sensor Networks

Lei Shu[1], *Yuanfang Chen*[2], *Zhangbing Zhou*[3]
[1] *College of Electronic Information and Computer, Guangdong University of Petrochemical Technology, China,*
email: lei.shu@live.ie
[2] *School of Software Technology, Dalian University of Technology, China,*
email: yuanfang_chen@ieee.org
[3] *China University of Geoscience (Beijing), China and Telecom SudParis, France,*
email: zbzhou@cugb.edu.cn

1. Introduction

Multimedia (e.g., image, audio and video) have become more and more important due to the ubiquitous proliferation of multimedia applications over Wireless Sensor Networks (WSNs) as shown in Fig. 1 [1]-[4]. Multimedia streaming in Wireless Multimedia Sensor Networks (WMSNs) is an important research issue.

Many approaches have been developed for facilitating multimedia data transmission in WMSNs, e.g., [5], [6], and [7]. However, many multimedia applications such as audio surveillance and human acoustic health monitoring require security protections for multimedia data streaming over WMSNs. Multimedia data integrity and transmission quality cannot be guaranteed in an insecure wireless multimedia streaming due to the existence of malicious nodes in the network, which cause the network to be insecure and the security of the multimedia data streaming cannot be guaranteed.

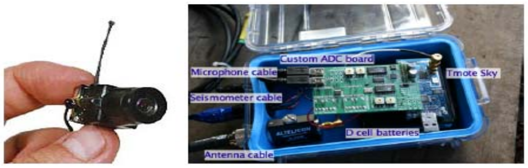
Furthermore, due to the unreliable wireless communication feature (e.g. error-prone wireless channels) and resource constraints (e.g. energy, processing capacity) of the sensor network, it is hard for real-time, imperceptible and energy-efficient finding of malicious nodes. To solve this, some researchers proposed watermarking-based schemes.

2. Watermarking based schemes

Watermarking technique has been widely used to assert an image data authentication over wired networks. However, resource constraints in small sensor nodes and the feature of error-prone wireless channels result in fundamental challenges for developing effective watermarking schemes in WMSNs.

The related work can be classified into two categories: quality-aware schemes and resource-aware schemes.

(a) Quality-aware schemes: The quality includes: transmission quality and watermarking quality. In article [8], H. Wang proposes a communication-resource-aware and adaptive watermarking scheme for multimedia authentication in WMSNs. The author keeps the transmission quality for the watermark, watermarked multimedia authenticity and multimedia data (image), by embedding watermark with adaptive coding redundancies and by unequally allocating network resources to protect the image and multimedia packets with the watermark information.



(a) Wireless camera sensor node (b) Wireless seismic and acoustic sensor node

(c) The applications of WMSNs

Fig. 1 Wireless Multimedia Sensor Networks (WMSNs)

In article [9], W. Wang et al. propose a quality-driven scheme to optimize stream authentication and Unequal Error Protection (UEP). The scheme provides digital image authentication and image transmission quality optimization. Moreover, the scheme is energy-efficient for WMSNs by using an authentication-aware wireless network resource allocation scheme (this scheme also reduces the image distortion in transmission).

For secure audio transmissions with high quality, in article [10], H. Wang and W. Wang et al. propose a scheme that dynamically determines the range of middle sub-band components for embedding the watermark with minimum quality distortion, based on psycho-acoustic models and adaptive sub-band thresholds. In addition, through unequal network resource allocation approach, the proposed scheme protects both middle sub-bands and high sub-bands, which include the important audio components. Moreover, the scheme is energy-efficient watermarking approach for audio transmissions.

For enduring underwater channel impairment and mitigating packet loss due to node failure and intrinsic underwater acoustic channel characteristic in Underwater Multimedia Sensor Networks (UMSNs), the article [11] presents a performance evaluation about error concealment and error correction algorithms for quality-aware image transmission over UMSNs. And, different combinations of multipath transport, watermarking-based error concealment (EC), forward error correction (FEC) and adaptive retransmission mechanism are also evaluated. In addition, two image quality assessment metrics are proposed to obtain the predicted quality of the image depending on the channel and node failures. From the comparative performance evaluations, the article shows that the EC scheme reconstructs the distorted image while avoiding a retransmission and the correlative delay.

(b) Resource-aware schemes: The resource includes: all constrained resources in WMSNs. "Resource-aware" is important for WMSNs-based scheme. Compared with traditional security techniques, watermarking schemes are usually light weight and do not require extensive computing and energy resources. Thus the watermark-based authentication can be an attractive option for wireless sensor applications. The articles [12] [13] [14] [15] are resource-aware, based on watermarking technique. Research work in [8] [10] also considered the resource-aware aspect.

3. The future challenges

Multimedia transmission may require multiple routing paths in the WMSNs, and the transmitted packages on different routing paths may not arrive at the destination in the designed order. Furthermore, a part of the image can be highly damaged due the lost of packages during the transmission. The security authentication issue for multimedia data in WMSNs is novel, and is proposed in recent years and gets more and more attention. Some challenges still exist in the topic:

- How to embed/transmit/protect/extract the watermark efficiently and robustly in low-cost sensors?
- How to transmit authenticated image/audio/video (multimedia data with watermark) with high energy efficiency and the transmission quality will not be affected by watermarking?
- How to find the balance between the transmission quality and the channel resource of WMSNs?
- How to find the balance between the multimedia-self quality and the high reliability of a watermark-embedded authentication [13]?

4. Conclusion

Watermarking is an old technology that has been studied and developed for copyright protection in image for many years. The normal assumed working environment of previous proposed methods did not consider the wireless communication in small size sensors, in which the transmission is not reliable and packages can be easily lost. In this paper, we briefly studied and summarized the existing methods that proposed for using watermarking technologies in wireless multimedia sensor networks, and try to discuss and discover some new research issues that we can further enhance the related applications.

Acknowledgement

Lei Shu's research work in this is supported by the Guangdong University of Petrochemical Technology's inner project.

References

[1] E. Gurses, O.B. Akan, "Multimedia Communication in Wireless Sensor Networks", In *Ann Telecommun*, 60(7-8), pp.799-827, 2005.

[2] I.F. Akyildiz, T. Melodia, K.R. Chowdhury. "A Survey on Wireless Multimedia Sensor Networks", In *Computer Networks*, 51(4), pp.921-960, 2007.

[3] S. Misra, M. Reisslein, G. Xue, "A Survey of Multimedia Streaming in Wireless Sensor Networks", In *IEEE Communication Surveys & Tutorials*, 10(4), pp.18-39, 2008.

[4] L. Shu, T. Hara, S. Nishio, Y. Chen, M. Hauswirth, "The New Challenge: Mobile Multimedia Sensor Networks" In *International Journal of Multimedia Intelligence and Security*, 2(2), pp.107-119, 2011.

[5] D.G. Costa, L. Affonso Guedes, "A Survey on Multimedia-Based Cross Layer Optimization in

Visual Sensor Networks", In *Sensors*, 11, pp.5439-5468, 2011.

[6] L. Shu, Y. Zhang, Z. Zhou, M. Hauswirth, Z. Yu, G. Hynes, "Transmitting and Gathering Streaming Data in Wireless Multimedia Sensor Networks within Expected Network Lifetime", in *ACM/Springer Mobile Networks and Applications*, 13(3-4), pp.306-322, 2008.

[7] L. Shu, M. Hauswirth, Y. Zhang, J. Ma, G. Min, Y. Wang, "Cross Layer Optimization on Data Gathering in Wireless Multimedia Sensor Networks within Expected Network Lifetime", in *Springer Journal of Universal Computer Science (JUCS)*, 16(10), pp.1343-1367, 2010.

[8] H. Wang, "Communication-resource-aware adaptive watermarking for multimedia authentication in wireless multimedia sensor networks", In *The Journal of Supercomputing*, pp. 1–15, 2010.

[9] W. Wang, D. Peng, H. Wang, H. Sharif, and H. Chen, "A multimedia quality-driven network resource management architecture for wireless sensor networks with stream authentication", in *IEEE Transactions on Multimedia*, 12(5), pp. 439–447, 2010.

[10] H. Wang, W. Wang, M. Chen and X. Yao, "Quality-driven secure audio transmissions in wireless multimedia sensor networks", In *Multimedia Tools and Applications*, pp. 1–17, 2011.

[11] P. Sarisaray-Boluk, V. Gungor, S. Baydere, and A. Harmanci, "Quality aware image transmission over underwater multimedia sensor networks", In *Ad Hoc Networks*, 9(7), pp. 1287–1301, 2011.

[12] H. Wang, D. Peng, W. Wang, H. Sharif, and H. Chen. Energy-aware adaptive watermarking for real-time image delivery in wireless sensor networks. In *IEEE Proceedings of ICC 2008*. pp. 1479–1483.

[13] W. Zhang, Y. Liu, S. Das, and P. De. Secure data aggregation in wireless sensor networks: a watermark based authentication supportive approach. In *Pervasive and Mobile Computing*, 4(5), pp. 658–680, 2008.

[14] S. Sultana, E. Bertino, and M. Shehab, "A provenance based mechanism to identify malicious packet dropping adversaries in sensor networks", In *Proceeding of the 31st IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW),* pp. 332–338, 2011.

[15] P. Boluk, S. Baydere, and A. Harmanci, "Robust image transmission over wireless sensor networks", In *Mobile Networks and Applications*, 16(2), pp. 149–170, 2011.

Lei Shu received the B.Sc. degree in Computer Science from South Central University for Nationalities, China, in 2002 and the M.Sc. degree in Computer Engineering from Kyung Hee University, Korea, in 2005 and the Ph.D. degree in Digital Enterprise Research Institute, from National University of Ireland, Galway, Ireland, in 2010. He is currently a full professor in College of Electronic Information and Computer, Guangdong University of Petrochemical Technology, China, also the vice-director of the Guangdong Petrochemical Equipment Fault Diagnosis (PEFD) Key Laboratory. He was a Specially Assigned Research Fellow in Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, Japan. He has published over 100 papers in related conferences, journals, and books. He had been awarded the MASS 2009 IEEE TCs Travel Grant and the Outstanding Leadership Award of EUC 2009 as Publicity Chair, the Globecom 2010 Best Paper Award. He has served as editor in chief for IEEE COMSOFT E-Letter, and editors for Wiley, European Transactions on Telecommunications, IET Communications, KSII Transactions on Internet and Information Systems (TIIS), Journal of Communications, Inderscience, International Journal of Sensor Networks, Wiley, Wireless Communications and Mobile Computing. He has served as more than 30 various Co-Chair for international conferences/workshops, e.g., IWCMC, ICC, ISCC; TPC members of more than 100 conferences, e.g., ICC, Globecom, ICCCN, WCNC, ISCC. His research interests include wireless sensor network, sensor network middleware, multimedia communication, and security. He is a member of IEEE and IEEE ComSoc.

Yuanfang Chen is a Ph.D. Candidate in Dalian University of Technology, China (from 2008); the expected graduation time: 2012. She received the M.S. degree in Dalian University of Technology, China, 2007. She is an assistant researcher in Illinois Institute of Technology

(Xiang-Yang Li), U.S.A., from 2009 to 2010. She had been awarded the MASS 2009 IEEE Travel Grant, IWCMC 2009 and MSN 2010 Invited Paper. She has served as volunteer of Mobicom & Mobihoc 2010 and IEA-AIE 2012. She has been invited as the TPC member of ICCNT 2011, GPC 2011 and ICCIT 2012; and the Publicity Co-Chair of the International Symposium on Mobile and Wireless Network Security 2011. She has served as reviewer of journals and conferences, e.g., Ad Hoc & Sensor Wireless Networks. She is a student member of IEEE and ACM. Her research interests include wireless sensor networks, cloud computing, game theory, Markov Decision Process, shadow price theory and sensor network middleware. She is a student member of China Computer Federation (CCF). Email: yuanfang_chen@ieee.org.

Zhangbing Zhou is an associate professor at the school of information engineering, China University of Geosciences (Beijing), China, and an adjunct associate professor at the computer science department, TELECOM SudParis, France. He received his PhD from Digital Enterprise Research Institute, National University of Ireland, Galway. After receiving his Master of Engineering from the Institute of Automation, Chinese Academy of Sciences in Beijing, he worked as a software engineer at Huawei Technologies Co. Ltd. for one year, and was served as a Member of Technical Staff and a team leader at Bell Labs China, Lucent Technologies (now Alcatel-Lucent) for five years. His research interests include process-aware information system, service-oriented computing, cloud computing, and sensor network middleware. He serves as a track chair, a program committee member, and a reviewer at many international conferences and journals and has been participating in several national and European research projects. Email: zbzhou@cugb.edu.cn.

INDUSTRIAL COLUMN: SPECIAL ISSUE ON "Quality of Service in the Smart Grid"

Quality of Service in the Smart Grid

Guest Editors: Florin Ciucu (T-Labs / TU Berlin) and Yik-Chung Wu (University of Hong Kong)
{florin@net.t-labs.tu-berlin.de, ycwu@eee.hku.hk}

Over the recent years there has been a substantial interest from the networking community to contribute their (external) expertise to the development of the next generation of the electric grid, i.e., the Smart Grid. While communication networks and the electric grid differ in several ways, they also share fundamental aspects which have been extensively studied within the networking community. One is the principle of Quality of Service (QoS) which broadly refers to the functioning of a system while entailing some forms of performance guarantees. In the particular networking context, QoS has been subject to more than a century of extensive research, from the Blocking Probability Formula of Erlang for dimensioning the telephone network to complex mechanisms designed for the seamless functioning of the Internet.

Given the necessity of the Smart Grid to encompass QoS mechanisms, the goal of this issue of the E-Letter is to bring related expertise from the networking community. Through a program consisting of six papers, contributed by leading networking researchers, the concrete goal of this issue is to convey a strong message that fundamental QoS networking research can be carried out and applied to the development of the Smart Grid. In particular, the papers transpose networking QoS concepts and mechanisms such as load characterization, scheduling, flow control, energy management systems, decentralized information processing, and energy trading, in the context of the Smart Grid.

The first paper, titled "*On Load Elasticity*", is co-authored by Keshav and Rosenberg from the University of Waterloo, Canada. The authors formalize the notion of elasticity, as a characteristic of some electric loads, by making subtle analogies and discussing the differences with the notion of elasticity previously introduced in economics and communication networks. The proper conceptualization of elasticity for electric loads is particularly important for the understanding of the demand response mechanism, which is instrumental for reducing the gap between the peak and average energy demand while ensuring the smart grid's reliability.

The second paper, titled "*Utility Maximizing PMU Data Rate Allocation under End-to-End Delay Constraints*", is co-authored by Dán and Vuković from KTH, Sweden. The authors formulate and solve a rate allocation problem for Phasor Measurement Units (PMUs) subject to end-to-end delay constraints. The interesting aspect of their solution is that it relies on the rate-controlled priority queueing discipline, which was proposed as an efficient service discipline in packet networks for providing both rate and delay guarantees, while decoupling the two. The overall problem addressed in this paper is particularly important since PMUs have been increasingly used for accurate monitoring of the state of power systems, whereas the predictable communication between PMUs and control centers is needed for many power system applications.

The third paper, titled "*Distributed Algorithms in Resistive Network Optimal Power Flow*", is co-authored by Tan, Cai, and Lou from City University of Hong Kong, Hong Kong. The authors address the Optimal Power Flow (OFP) problem in a resistive network which can accommodate energy load from renewable sources. While OFP is a classically hard non-linear and non-convex optimization problem, the authors propose an iterative fixed-point analysis which lends itself to a computationally efficient distributed algorithm to compute the global optimal solution. This is particularly relevant for the optimal power allocation in the Smart Grid, whereby energy, voltage, and networking resources need to be allocated on a fast time-scale and in a distributed manner.

The fourth paper is "*Customer Energy Management System in Smart Grid*" by Yang and Li from The University of Hong Kong. This paper discusses the features and challenges of energy management system (EMS) in the customer domain, and proposes an innovative framework, Pervasive Service-Oriented Networks (PERSON), to address the challenges of heterogeneity, distributive and dynamic natures of smart grid. An implementation of an EMS system based on the proposed PERSON is also reported in this article.

The fifth paper is "*Decentralized Information Processing and Energy Management for Smart Grid*" by Li and Scaglione from University of California, Davis. The authors discuss various issues on wide-area monitoring and data analysis. In particular, gossiping algorithms are proposed to reduce network usage and processing latency during state estimation of power system. Furthermore, ideas of demand side management from the view of power switching, and distributed load scheduling within a neighborhood are surveyed.

Last but not least, the sixth paper is "*Toward Evolved Energy Trading Mechanisms in the Smart Grid*" by Wang and Saad from University of Miami. This paper develops game-theoretic models for managing the energy trading processes. This paper provides insights into the potential of game-theoretic approaches in smart grid, and shows a promising future for consumers to control their energy usage and trading of their energy.

As a final remark, we believe that the six contributions included in this issue illustrate the potential of extending QoS research from communication networks to the Smart Grid. Moreover, we hope that this line of research to be adopted on a large scale, given both the vision and challenges of the Smart Grid.

Florin Ciucu was educated at the Faculty of Mathematics, University of Bucharest (B.Sc. in Informatics, 1998), George Mason University (M.Sc. in Computer Science, 2001), and University of Virginia (Ph.D. in Computer Science, 2007). Between 2007 and 2008 he was a Postdoctoral Fellow in the Electrical and Computer Engineering Department at the University of Toronto. Currently he is a Senior Research Scientist at Telekom Innovation Laboratories (T-Labs) and TU Berlin. His research interests are in the stochastic analysis of communication networks and the smart grid, resource allocation, and randomized algorithms for large systems. He is a recipient of the ACM Sigmetrics 2005 Best Student Paper Award.

Yik-Chung Wu received the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from the University of Hong Kong (HKU). He received the Ph.D. degree from Texas A&M University, College Station, in 2005. From August 2005 to August 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with HKU, currently as an Associate Professor. He was a visiting scholar at Princeton University in summer 2011. His research interests are in general area of communication systems. He is currently serving as an editor for the IEEE Transactions on Communications, IEEE Communications Letters, and Journal of Communications and Networks.

On Load Elasticity

S. Keshav and C. Rosenberg
University of Waterloo, Canada
{keshav, cath}@uwaterloo.ca

1. Introduction

Grid operators today meet the twin goals of system reliability and revenue maximization by provisioning the system conservatively at close to the peak load and charging customers for the resulting high costs of the necessary infrastructure. This maximizes revenues while maintaining reliability. Most grids are regulated monopolies or have nearly monopolistic market share in their home market, so their operators have had no incentive to reduce capital spending that can be recouped from a large and captive customer base.

Unfortunately, this business-as-usual approach does not take some externalities into account. For instance, coal, which was used to generate two-thirds of all electricity in the US in 2008, causes air pollution, radioactive emissions, and an enormous carbon footprint. With the looming danger from anthropogenic climate change, there is tremendous political pressure to decommission coal plants, reducing baseload generation capacity. Similarly, widespread fears of nuclear proliferation are causing some countries to hold back on the deployment of additional nuclear power plants (as in the United States) or even decommission existing plants (as in Germany). In view of these externalities, the inherent inefficiency of uncontrolled peak loads, and past experience in demand management, *demand response* has emerged as a critical feature of the smart grid [1].

Demand response (DR) refers to the use of pricing to cause electricity consumers to intentionally modify the time at which they consume electricity, their peak demand level, or their total electricity consumption [2]. It can be achieved either through incentive-based programs, where consumers are paid to participate in a program where utilities directly control their load, or market-based programs, where customers respond to price signals that reflect overall system load. In either case, DR reduces the peak-to-average demand ratio, increases the power factor of generators, and allows generators to defer capacity increases. Moreover, it decreases generation costs by reducing the size of spinning reserves and, in some cases, the use of expensive energy sources [2]. Finally, it can increase system reliability–averting load-shedding or blackouts–by reducing load when system stability is in jeopardy.

The underlying assumption made by any DR scheme is that load is *elastic*, that is, some loads can be time-shifted with no loss of utility. Although this is a critical assumption, we are not aware of prior work that analytically examines load elasticity; hence this communication.

2. The concept of load elasticity

It is illustrative to first study the concept of elasticity both in economics and in the context of the Internet.

Economic concept of elasticity

The economic definition of demand elasticity measures the degree to which the total revenue (the product of price and demand) is affected by a change in price [3]. Demand is said to be elastic at a particular price point when a decrease in price increases the total revenue. For example, with elastic demand, a 1% decrease in price at a particular price point would increase demand by more than 1% so that the total revenue increases. Otherwise, it is said to be inelastic (see Figure 1). A customer with elastic demand at a particular price point can be thought to be responsive to a price signal, in contrast to a customer with inelastic demand, who does not reduce demand proportional to the percentage increase in price.

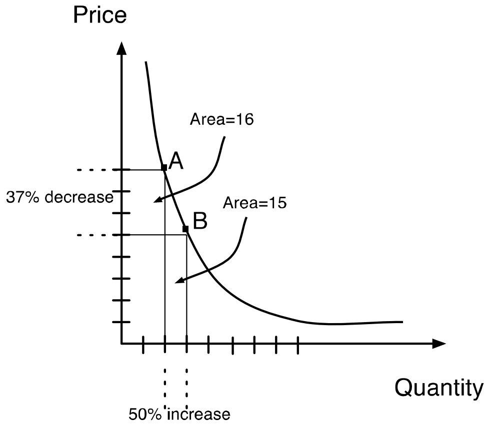


Figure 1: Demand Elasticity. A 37% decrease in price increases demand by 50%, but the total revenue at point B is lower than at point A, so demand is said to be *inelastic* at point A.

Network formulation

The widely-accepted definition of elasticity in a communication network, due to Shenker [4], is in terms of the utility to a distributed application (i.e., an application that has been implemented as two or more parts with a network connection between them) as a function of the connection's *bandwidth* (the network equivalent of power that is measured in bits per second rather than watts) available to it. If this utility exhibits a diminishing marginal increase as a function of bandwidth, then the application is said to be elastic (see Figure 2).

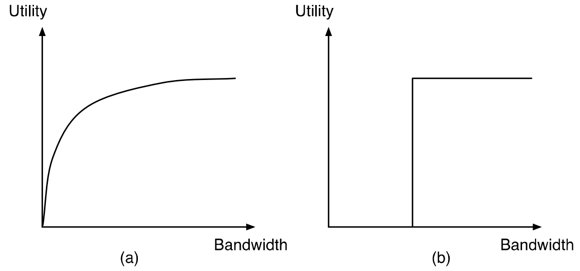


Figure 2: Elastic (a) and inelastic (b) applications

It is instructive to contrast an elastic application with an inelastic one. An inelastic application gets no utility at all when its connection's bandwidth decreases below some threshold and no additional increase in utility when this bandwidth exceeds this threshold. Thus, for such an application, the marginal increase in utility as a function of bandwidth is zero everywhere, except at the threshold value itself.

Note that the Internet definition of elasticity does not consider the duration of the connection. This obscures the fact that the connection must eventually transfer a data item from source to destination. In other words, although the bandwidth associated with a connection may vary over time, a certain number of bits must eventually be transferred over the connection. In the electrical grid, this corresponds to a load that must receive a certain amount of energy, though the power may vary over time. The Internet definition of elasticity must therefore be augmented with the constraint that for an elastic application the area under the bandwidth vs. time curve, which is the size of the data item, is conserved.

Note also that the Internet formulation of elasticity, unlike the economic formulation, does not take pricing into account.

Elastic electrical loads

We now consider how to apply these concepts of elasticity to electrical loads. Intuitively, we think of an electrical load as being elastic if it can be modified in some way, such as, for example, in response to a price or congestion signal, without overly reducing the comfort of the consumer [5]. This allows the system operator to manipulate demand to achieve some system objective such as reducing peak demand or transmission line congestion, without sacrificing customer comfort, or perhaps compensating for the reduction in comfort with a payment. Note the inherent three-way trade-off between user comfort, payments to the user, and the system objective (in some cases utilities may mandate load reduction without payments). For example, an electric vehicle (EV) owner may be insensitive to its charging rate as long as the EV is charged before some loose deadline. Therefore, reducing its charging rate can help to achieve the system objective of reducing peak load, which makes the EV's load elastic. In contrast, the power given to a refrigerator today can be neither diminished nor time shifted. We draw upon this intuition to develop a definition of elasticity next.

3. Quantifying elasticity

We define a *load profile* π to be a continuous function of time that represents an appliance's load as a function of time. We define the utility to a customer of a load profile $U_C(\pi)$ to be the benefit to the customer from a particular load profile. This generalizes the Internet concept of the utility as a function of the bandwidth to the utility as a function of the bandwidth over a time period.

A nominal load profile π^* is *inelastic* if $U_C(\pi) = 0$ for all $\pi^* \neq \pi$. In other words, the slightest change in the profile causes the customer's utility to drop to zero – for example, a load due to a television that would be damaged if there a drop in either voltage or drawn current. In practice, we expect no load to be completely inelastic, because all devices have some built-in margins to deal with short-term fluctuations.

In contrast, a nominal load profile π^* is said to be *purely elastic* when there exists a set Π of other profiles such that the customer's utility from all members of Π is equal to $U_C(\pi^*)$. That is, the customer is indifferent to all profiles in Π.

We can generalize this further. Assume that a nominal load profile π^* with utility $U_C(\pi^*) = a$ is such that there exists at least one other load profile π' such that $U_C(\pi') = a - \varepsilon$. Then, we call the load profile ε-*elastic*.

4. Load types

We now classify load types by their degree of elasticity, drawing on a classification of connection types in data networks such as the Internet.

In the Internet, packets are sent on a connection from the source to a destination. Generally speaking, the quality of a connection diminishes and the utility of the connection to the application user is reduced as the number of bits per second that the network provisions for that connection (also called its bit rate) decreases. Three types of dependency of utility on the bitrate are well known [6]:

- <u>Constant Bit Rate</u> (CBR): Such applications are inelastic with respect to the bit rate. Their marginal increase in utility with respect to the bit rate is zero everywhere except at a particular threshold value. In other words, the source receives zero utility unless the network allocates enough resources to carry at least the threshold bitrate.
- <u>Variable Bit Rate</u> (VBR): Such applications generate traffic in bursts, rather than in a smooth stream. A VBR source is typically modeled as having an intrinsic long-term bit rate (called its average rate) with occasional bursts of limited duration at a rate as high as a specified peak rate. An application receives zero utility if the network allocates resources less than the average rate to the connection, and its utility increases and then saturates as the network increases its allocated resources to the peak rate and then beyond.
- <u>Available Bit Rate</u> (ABR): Such applications have a non-zero marginal gain in utility everywhere. Recall that there is an implicit assumption that the area under the bandwidth profile is conserved. A typical example is a file transfer where a source wants to send some number of bits to a recipient, the sooner the better.

This inspires us to classify loads as Fixed Power, Variable Power, and Available Power loads.

- <u>Fixed Power</u> (FP): Such loads are inelastic with respect to their load profile. Their utility is zero everywhere except when served using their nominal load profile.
- <u>Variable Power</u> (VP): Such loads are ε-*elastic*. That it, they have a preferred profile, but their utility does not change much for profiles 'close to' that profile.
- <u>Available Power</u> (AP): This refers to purely elastic loads whose utility does not change despite certain changes in the load profile. An example of such a change could be (a) the area under the profile curve is conserved and (b) the demand is satisfied before a given deadline (these implicitly define Π).

Traces show that a consumer's aggregate electrical load can be roughly partitioned into two portions. The base load is the load from always-on devices such as set-top boxes, safety lighting etc. This is typically fairly low and can be modeled as Fixed Power loads. Demand sharply increases when heavy-load devices such as air-conditioners, refrigerators, electric ovens, and baseboard heaters are turned on. These loads can be modeled as AP (or as VP under certain conditions). Finally, new loads, such as EV, could be modeled as VP or AP.

4. Conclusion

This paper discusses and formalizes the notion of load elasticity in the electrical grid, and compares it with the notion of elasticity used in economics and in the Internet. We use these concepts to describe electrical loads as falling into three natural categories: Fixed Power, Variable Power, and Available Power, that mimic the well-known categories of CBR, VBR, and ABR, respectively, in the field of computer networking.

References

[1] Ipakchi, Ali, and Farrokh Albuyeh. "Grid of the future." *Power and Energy Magazine, IEEE* 7, no. 2,pp. 52-62, 2009.

[2] US Department of Energy, "Benefits of demand response in electricity markets and recommendations for achieving them," Report to Congress, 2006.

[3] Samuelson, P.A., "Economics," *McGraw Hill*, 11th Ed. 1980.

[4] Shenker, S., "Fundamental design issues for the future Internet." *IEEE J. Selected Areas in Communications*, v.13, no. 7, pp. 1176-1188, 1995.

[5] Srikantha, P., C. Rosenberg, and S. Keshav, "An analysis of peak demand reductions due to elasticity of domestic appliances." In *Proc. of the 3rd International Conference on Future Energy Systems*, p. 28, 2012.

[6] Keshav, S., "An Engineering Approach to Computer Networking," *Addison-Wesley*, 1997.

Srinivasan Keshav is a Professor, Canada Research Chair, and Cisco Chair in Smart Grid in the Cheriton School of Computer Science at the University of Waterloo. He received a PhD in Computer Science from UC Berkeley in 1991.

Catherine Rosenberg is a Fellow of the IEEE, Professor, and Canada Research Chair in the Department of Electrical and Computer Engineering at the University of Waterloo. She received a Doctorat en Sciences from Université de Paris XI Orsay, France.

Utility-based PMU Data Rate Allocation under End-to-end Delay Constraints

György Dán and Ognjen Vuković
School of Electrical Engineering
KTH Royal Institute of Technology, Stockholm, Sweden
{gyuri,vukovic}@ee.kth.se

1. Introduction

The proper operation of the electric power system requires that generation and demand be balanced continuously respecting the capacity constraints of the power transmission and distribution infrastructure. Maintaining the balance is ever more challenging due to the increasing penetration of solar and wind energy, which are highly intermittent. Economic operation requires that the system should operate close to its capacity limits, but large interconnected power systems under high stress can exhibit oscillations that can lead to a disconnection of the power system into islands, with blackouts as a consequence.

To monitor the state of the power system, operators require accurate, frequent and time synchronized measurements of the state of the power system. Phasor Measurement Units (PMUs) are increasingly used to periodically measure the instantaneous state of the power system, such as voltage and current phasors (i.e., amplitude and phase angle). The frequency at which measurements can be taken is selectable, from 1/s up to twice the mains frequency of the power system (i.e., up to 100Hz or 120Hz). The measurement frequency might increase in the future if applications monitoring higher frequency harmonics become wide spread. PMUs rely on the time information provided by the global positioning system (GPS) to attach time stamps with sub-millisecond precision to measured data [1][8]. The data collected by PMUs at the substations need to be delivered from the substations to one or more control centers or to distributed controllers.

The frequency at which PMU measurements are needed, the location where the data should be delivered to, and the latency at which the data should be available for processing vary significantly depending on the particular application [2][3][4][5]. In general, high frequency measurement data might not be needed all the time and thus dimensioning the communication infrastructure for peak bandwidth would be inefficient and costly. Instead, the communication infrastructure could be dimensioned so that it can accommodate high frequency measurement data from a subset of the PMUs. In this case the rate of the PMUs should be allocated adaptively as a function of their importance.

The communication infrastructure of most modern power systems, which is based on Optical Ground Wire (OPGW) installations along the power transmission lines, and SDH/SONET based time division multiplexing (TDM), does not support such adaptive rate allocation. Instead, it provides predictable end-to-end delays through dedicated channels, but low bandwidth use efficiency and little flexibility in terms of the connections that can be established. IP over switched Ethernet or over MPLS could provide the required flexibility and efficient use of the available bandwidth, but providing delay guarantees in IP networks proved to be difficult in general. There is thus a need for a solution that would allow rate adaptation under delay constraints over packet switched networks with high utilization, designed for the communication needs of power systems.

In this paper, we formulate the problem of adaptive PMU rate allocation subject to end-to-end delay constraints as a utility maximization problem. We show that for a shape preserving scheduling discipline the problem can be converted into a network utility maximization problem, and provide a solution to the problem. We use the IEEE 118 bus benchmark power system to obtain numerical results.

2. Adaptive Rate Allocation Problem

We model the communication infrastructure as a directed graph $G=(V,E)$, where every vertex is a switch and there is an edge e_{ij} between two vertices v_i and v_j if there is a link connecting v_i to v_j. We denote the bandwidth of edge e by c_e. We assume that the switches implement non-preemptive priority queuing, and PMU data are assigned highest priority. We denote by P_{max} the maximum packet size in the network.

We denote the set of PMUs by $S=\{1,\ldots,N_S\}$. We denote the vertex to which PMU s is attached by v_s, and by $L(s)$ the path in the network established to the destination of the data generated by PMU s, e.g., to a control center. We denote by $r_s \geq 0$ the rate at which PMU s generates measurement data and by D_s the end-to-end delay budget of the data generated by PMU s. We assume that the rate r_s is chosen from a closed interval and that the utility of the PMU data is a concave increasing continuous function $U_s(r_s)$ of its rate. We denote by P_s the packet size used by PMU s, which is determined by the number of digital and analog values sent, and is typically constant.

Utility Maximal Rate Allocation

Given the above notation we can formulate the problem of utility maximal rate allocation in the framework of network utility maximization [7] as

$$\max \sum_s U_s(r_s) \qquad (1)$$

$$s.t.$$

$$\sum_{s:e\in L(s)} r_s \le c_e \qquad \forall e \in E \qquad (2)$$

$$\sum_{e\in L(s)} d_e(r_1,\dots,r_S) \le D_s \qquad \forall s \in S \qquad (3)$$

The worst case queuing delay d_e at edge e is a function of the rates and the arrival process of the messages belonging to the flows traversing the edge. Although the arrival process of the data from PMU s is deterministic (i.e., 1 message per sampling period) upon entering switch v_s, this is not necessarily the case upon arrival to the second switch on the path due to multiplexing with other PMU data and possibly with background traffic.

3. Solution under Rate-controlled Priority Queuing

In the following we consider that every switch implements rate-controlled priority queuing (RCPQ). This makes it possible to bound the worst case per-hop delay. Under the RCPQ discipline [6] every switch restores the message inter-arrival time of a flow of messages to the inter-arrival time at the origin of the flow. While such a queuing discipline can be impractical for Internet traffic, in the case of PMU data delivery the intended inter-arrival time of messages is known from the sampling frequency. Furthermore, every message carries a timestamp with sub-millisecond accuracy, which facilitates the implementation of this queuing discipline.

As shown in [6], under the RCPQ discipline the worst-case per hop delay of a message on link e is bounded by the smallest d_e that satisfies

$$\sum_{s:e\in L(s)} \left\lceil \frac{d_e}{t_s} \right\rceil \frac{P_s}{c_e} + \frac{P_{\max}}{c_e} \le d_e. \qquad (4)$$

Let us consider first the case that for all flows the message inter-arrival time satisfies $t_s \ge d_e$. In this case (4) is only a function of the number of flows traversing the link, but not that of the rate of the individual flows. This allows us to simplify (4) and get a bound of the worst case delay at link e

$$\sum_{s:e\in L(s)} \frac{P_s}{c_e} + \frac{P_{\max}}{c_e} \le d_e \qquad (5)$$

We can substitute the bound in (5) into (1)-(3) to obtain the following optimization problem

$$\max \sum_s U_s(r_s) \qquad (6)$$

$$s.t.$$

$$\sum_{s:e\in L(s)} P_s + P_{\max} \le c_e d_e \qquad \forall e \in E \qquad (7)$$

$$\sum_{e\in L(s)} d_e \le D_s \qquad \forall s \in S \qquad (8)$$

Observe that (7) implies (2) because $r_s=P_s/t_s$ and $t_s \ge d_e$. The problem can be solved in two steps. First, find the smallest possible link delays d_e for every link using (7). If the per link delays d_e satisfy the end-to-end delay constraints D_s in (8), then the rates r_s that maximize (6) can be found by observing that by assumption $t_s \ge \max_{e\in L(s)} d_e$. Rate allocation is not necessary in this case, but dynamic routing could be used to decrease t_s for a particular PMU, and thereby to allow its rate to be increased.

Without the assumption $t_s \ge d_e$ we can bound the left hand side of (4) by

$$\sum_{s:e\in L(s)} \left\lceil \frac{d_e}{t_s} \right\rceil \frac{P_s}{c_e} + \frac{P_{\max}}{c_e} \le \sum_{s:e\in L(s)} \left(\frac{d_e}{t_s} + 1 \right) \frac{P_s}{c_e} + \frac{P_{\max}}{c_e},$$

which is tight if the message inter-arrival times $t_s << d_e$. By substituting $t_s=P_s/r_s$, and solving for d_e we obtain the bound for the per hop delay

$$\frac{\sum_{s:e\in L(s)} P_s + P_{\max}}{c_e - \sum_{s:e\in L(s)} r_s} \le d_e. \qquad (9)$$

Clearly, the constraint (9) is stricter than (7), but it holds for higher data rates. We can use this bound to formulate the following optimization problem

$$\max \sum_s U_s(r_s) \qquad (10)$$

$$s.t.$$

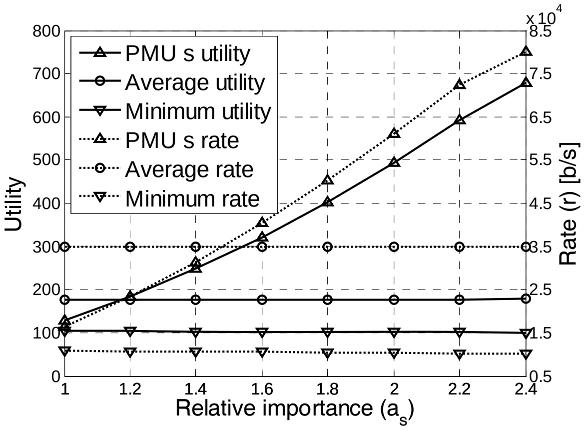
$$\sum_{s:e\in L(s)} r_s \le c_e - \frac{\sum_{s:e\in L(s)} P_s + P_{\max}}{d_e} \qquad \forall e \in E \qquad (11)$$

$$\sum_{e\in L(s)} d_e \le D_s \qquad \forall s \in S \qquad (12)$$

The objective function (10) is strictly concave and the domain of the optimization is convex and compact, thus there is a global optimum and it is attained if there is a feasible solution to (11) and (12). The standard dual decomposition of the problem as described in [7] cannot be applied directly because the rate constraints in (11) are coupled through (12). We can nevertheless use the projected subgradient method to find the global optimum.

4. Numerical Results

We use the IEEE 118 bus system to illustrate the proposed adaptive rate allocation algorithm. We consider that a PMU is installed at every substation, the number of phasors per message and thus the message size is given by the connectivity of the substation. All PMU data are delivered to the control center located at the bus with highest degree via a shortest path in the network topology which resembles the topology of the power system. All communication links have a capacity of 1Mbps, except for the link that connects the control center to the substation with highest degree,



which has tenfold capacity. The maximum packet size $P_{max}=1500$ bytes, and the end-to-end delay constraint D_s is set to 500ms for all PMUs. We use $a_s\sqrt{r_s}$ as the utility function of PMU s, and use a_s to set the relative importance of an individual PMU.

Figure 1 shows the change of the rate and that of the utility as the relative importance a_s is increases for a single PMU to prioritize its data. The results show that through adapting a_s the rate and the utility of PMU s increases monotonically, while the rates and utilities of the other PMUs decrease slowly.

Conclusion

In this work we considered the problem of adapting the PMU data rate subject to link capacity and end-to-end delay constraints. We formulated a network utility optimization problem and discussed the solution to two special cases of the problem. We illustrated the solutions on the IEEE 118 bus benchmark power system. The feasibility of the proposed approach for PMU rate adaptation depends on whether or not it is possible to map end user requirements into utility functions, and whether a solution can be obtained for discrete sets of PMU data rates with low computational complexity. Addressing these questions will be subject of our future work.

References

[17] "IEEE Standard C37.118 for Synchrophasors for Power Systems," IEEE C37.118-2005

[18] K.E. Martin and J.R. Carroll, "Phasing in the Technology", IEEE Power and Energy Magazine, vol.6, no.5, pp.24-33., 2008

[19] J.H. Chow and S.G. Ghiocel, "An Adaptive Wide-Area Power System Damping Controller using Synchrophasor Data", in Control and Optimization Methods for Electric Smart Grids (ed. A. Chakrabortty), Springer, 2012

[20] W. Yao, L. Jiang, J.Y. Wen, S.J. Cheng and Q.H. Wu, "An adaptive wide-area damping controller based on generalized predictive control and model identification", IEEE Power and Energy Society General Meeting, July 2009

[21] Z. Hu and J.V. Milanovic, "The Effectiveness of WAM Based Adaptive Supervisory Controller for Global Stabilization of Power Systems," *in Proc. of IEEE PowerTech, pp. 1652-1659, July 2007.*

[22] H. Zhang and D. Ferrari, "Rate-controlled static priority queueing," *in Proc. of IEEE Infocom, 1993*

[23] F. P. Kelly, "Charging and rate control for elastic traffic," European Transactions on Telecommunications, vol. 8, pp. 33–37, 1997.

[8] IEC/TR 61850-90-5, "Communication networks and systems for power utility automation - Part 90-5: Use of IEC 61850 to transmit synchrophasor information according to IEEE C37.118", May 2012

György Dán received the M.Sc. degree in computer engineering from the Budapest University of Technology and Economics, Hungary in 1999 and the M.Sc. degree in business administration from the Corvinus University of Budapest, Hungary in 2003. He worked as a consultant in the field of access networks, streaming media and videoconferencing 1999-2001. He received his Ph.D. in Telecommunications in 2006 from KTH, Royal Institute of Technology, Stockholm, Sweden, where he is currently an associate professor. He was a visiting researcher at the Swedish Institute of Computer Science in 2008 and a Fulbright research scholar at University of Illinois Urbana-Champaign in 2012-2013. His research interests include the design and analysis of content management systems, game theoretical models of networked systems, and cyber-physical system security in power systems.

Ognjen Vuković is a PhD student in the Laboratory of Communication Networks at KTH Royal Institute of Technology in Stockholm, Sweden. In 2010, he received his M.Sc. degree in Telecommunications, System engineering and Radio Communications, from the Faculty of Electrical Engineering, University of Belgrade. His research interests include power system communication technologies, communication security and availability, and resource management for networked systems.

Distributed Algorithms in Resistive Network Optimal Power Flow

Chee Wei Tan[1], *Desmond W. H. Cai*[2] *and Xin Lou*[1]
City University of Hong Kong, Hong Kong[1]
California Institute of Technology, CA 91125 USA[2]

1. Introduction

The smart grid network will have hundreds of millions of active endpoints that need to be coordinated at all levels of functionalities from generation, transmission to inter-communication. This interconnected system of hundreds of millions of distributed energy resources introduce rapid, large and random fluctuations in power supply and demand, and thus requires fast and efficient computational algorithms for optimizing the power flow in the network. From its scale, network connectivity in the smart grid will pose new challenges to power flow optimization. Unlike in a traditional regulated system that is centrally controlled by a utility at a slow timescale, power flow in smart grid network requires energy, voltage support and its networking resources to be optimized in a decentralized manner to meet real-time performance requirements.

The Optimal Power Flow (OPF) problem is a classical nonlinear nonconvex optimization problem that minimizes the power generation costs and transmission loss in a power network subject to physical constraints governed by Kirchhoff's and Ohm's law [1]-[2]. There is a huge body of work on solving OPF since Carpentier's first formulation in 1962 [1]. To overcome the nonlinearity, the majority of the work uses linearization [3] and approximation techniques [4]-[5] to simplify the OPF. In this paper, we use the bus injection model of OPF that focuses on the voltages and power injection at each node of the network. We study the OPF problem in a purely resistive network, where only the real power exists. This kind of network can be practically important when there are renewable generations integrated into the smart grid, e.g., solar cells that produce direct current power in microgrids.

A source of computational difficulty in OPF is the nonconvexity and nonlinear coupling of power, current and voltages to satisfy the power balance equations. Nonconvex problems are generally hard to solve [6]. An exhaustive search for the global optimal solution can be prohibitively expensive. Weak relaxation and approximation methods can lead to poor solutions. Difficult as it seems, the authors in [7] showed recently that the Lagrange duality gap between the OPF problem and its convex dual can be zero in a radial network or under some mild conditions on the system model, and this was numerically verified to be true for a number of practical IEEE power networks. The implication is that the OPF can be optimally solved by a reformulation-relaxation technique using semidefinite programming (SDP) [6]. In the purely resistive network case, this relaxed problem is in fact exact when there is load oversatisfaction [7]-[8].

Different from prior work on algorithm design for solving OPF, e.g., in [4]-[5], we leverage the zero duality gap condition and dual decomposition to design decentralized algorithms to solve the OPF. Each bus (either the generator bus or the demand bus) exchanges local information with its one-hop neighbors. The uniqueness characterization provides an interesting perspective on the convergence proof of local algorithms to the global optimal solution.

Overall, the contributions of the paper are as follows:

1. We characterize the uniqueness of the resistive network OPF solution using the Poincare-Hopf Index Theorem for radial network topologies.
2. We solve the resistive network OPF problem using dual decomposition and iterative fixed-point analysis. Computationally fast convergent local algorithms with low complexity are proposed to compute the global optimal solution of the resistive network OPF.

2. SYSTEM MODEL AND PROBLEM FORMULATION

Let us consider a resistive power network with a set of buses $\mathcal{N} = \{1, 2, \ldots, N\}$ and a set of transmission lines $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. We assume that each bus is either a generation bus or a demand bus. A demand bus i can model the aggregate of users (loads) in a distribution network. For each bus i, we use $\mathbf{\Omega}_i$ to represent the set of buses connecting to bus i and $|\mathbf{\Omega}_i| \geq 1$. Moreover, we assume that the line admittance satisfies $Y_{ij} = Y_{ji} \in \mathbb{R}_+$ if $(i, j) \in \mathcal{E}$; and $Y_{ij} = Y_{ji} = 0$, otherwise. We use $\mathbf{V}$ and $\mathbf{I}$ to denote the voltage $(V_i)_{i\in\mathcal{N}}$ and current vector $(I_i)_{i\in\mathcal{N}}$, respectively.

In this resistive network OPF, we consider nodal power and voltage constraints given by $V_i I_i \leq \overline{p}_i$ and $V_i \in [\underline{V}_i, \overline{V}_i], \forall i \in \mathcal{N}$, respectively. If bus i is a generation bus, then $\overline{p}_i$ represents the generator capacity and $\overline{p}_i > 0$. If bus i is a demand bus, then $\overline{p}_i < 0$ and this constraint corresponds to the minimum demand that has to be satisfied at bus i. Since we consider the power stored at each demand bus, we assume that demand can be over-satisfied. Therefore, a demand bus i not only absorbs $|\overline{p}_i|$ amount of power, but it also absorbs additional power to charge the power storage device attached to it. Moreover, by Ohm's Law and Kirchhoff's current law, we have **I** = **Y V**, i.e.,

$$\begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{pmatrix} = \begin{pmatrix} \sum_{j\in\Omega_1} Y_{1j} & \cdots & -Y_{1N} \\ -Y_{21} & \cdots & -Y_{2N} \\ \vdots & \ddots & \vdots \\ -Y_{N1} & \cdots & \sum_{j\in\Omega_N} Y_{Nj} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{pmatrix}, \quad (1)$$

where **Y** is the system admittance matrix. Using the relationship between **Y** and **I**, the optimal power flow problem contains only one variable **V** and is formulated as:

$$\begin{array}{ll} \text{minimize} & \mathbf{V}^T\mathbf{Y}\mathbf{V} \\ \text{subject to} & \mathbf{V}^T\mathbf{Y}_i\mathbf{V} \le \overline{p}_i \quad \forall i \in \mathcal{N}, \\ & \underline{\mathbf{V}} \le \mathbf{V} \le \overline{\mathbf{V}}, \qquad\qquad (2) \\ & \mathbf{V}^T\mathbf{Y}_{ij}\mathbf{V} \le c_{ij} \quad \forall (i,j) \in \mathcal{E}, \\ \text{variables:} & \mathbf{V}, \end{array}$$

where $\mathbf{Y}_i = \frac{1}{2}(\mathbf{E}_i\mathbf{Y}+\mathbf{Y}\mathbf{E}_i)$ and $\mathbf{Y}_{ij} = \mathbf{Y}_{ij}(\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T$, where $\mathbf{e}_i$ is the standard basis vector in $\mathbb{R}^n$ and $\mathbf{E}_i = \mathbf{e}_i\,\mathbf{e}_i^T \in \mathbb{R}^{n\times n}$ [7].

3. UNIQUENESS CHARACTERIZATION

In this section, we study the uniqueness of the solution to (2). Our approach is to leverage the zero duality gap property in (2) and the Poincare-Hopf Index Theorem in [9] together with the nonnegativity associated with the variables and problem parameters of (2). The uniqueness characterization of (2) has implications on how local algorithms with low complexity can be designed to solve (2) in Section 4.

We first use the Lagrange dual decomposition to relax (2). If we consider the partial Lagrange minimization (after eliminating nodal and line capacity constraints), the problem is still nonconvex in **V**. However, since the objective is smooth and constraints satisfy the linearity constraint qualification, the optimal solution to the partial Lagrangian minimization problem satisfies the Karush-Kuhn Tucker (KKT) conditions. By rewriting the KKT conditions, it can be shown that the optimal solution satisfies:

$$\mathbf{V} = \max\{\underline{\mathbf{V}}, \min\{\overline{\mathbf{V}}, \mathbf{B}(\boldsymbol{\lambda},\boldsymbol{\mu})\mathbf{V}\}\}, \quad (3)$$

where the nonnegative matrix **B** has entries in terms of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$:

$$B_{ij} = \begin{cases} \dfrac{2Y_{ij} + \lambda_i Y_{ij} + \lambda_j Y_{ij} + 2\mu_{ij} Y_{ij}}{2\left((1+\lambda_i)\sum_{j\in\Omega_i} Y_{ij} + \sum_{(i,j)\in\mathcal{E}} \mu_{ij} Y_{ij}\right)}, & \forall (i,j) \in \mathcal{E} \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where the dual variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ correspond to nodal and line capacity constraints, respectively. Suppose that the optimal dual variables $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ of (2) are given (recall that they exist due to zero duality gap), then the optimal voltage vector $\mathbf{V}^*$ must satisfy

$$\mathbf{V}^* = \max\{\underline{\mathbf{V}}, \min\{\overline{\mathbf{V}}, \mathbf{B}(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\mathbf{V}^*\}\}.$$

Next, by applying the differential topology results in [9], we have shown in [10] about the uniqueness of the OPF under simple topologies. Now, we propose the following lemma on the more general case:

Lemma 1: The optimal solution of (2) is unique for distribution networks, i.e., the line-network and the radial network.

For example, let us consider a simple 2-bus line network as shown in Figure 1(a), where a generator is attached to bus 1 and a load is attached to bus 2. Also, $\overline{p}_1 > 0$ and $\overline{p}_2 < 0$. Consider the Perron-Frobenius eigenvalue ρ (**B**), it is easy to see that ρ (**B**) > 1 must hold if $\mathbf{V}^*$ is not trivial (not corresponds to the all-one vector). Thus, $\mathbf{V}^*$ is not in the interior of the box constraint, i.e., at least one of the box constraints is binding. It can be checked that the determinant of the principal submatrix (a scalar) of the hessian is given by $(1+\lambda_i+\mu_{12})Y_{12}$, i = 1 or 2, which is always positive. Then we applied [9] to conclude that **V** has a unique optimal solution. To illustrate, Figure 2(b) plots the solution space over V_1 and V_2 in the box constraint [1, 2] for fixed dual variables, and there is only one optimal solution, i.e., $V_1 = 2V$, $V_2 = 1.5V$.

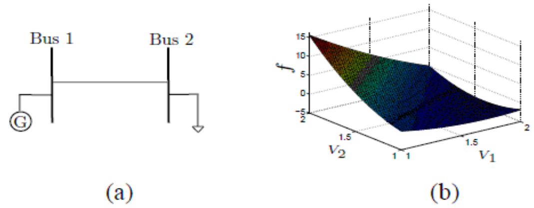


Fig. 1. (a) In this example, two connected buses (bus 1 and bus 2) are respectively attached by a generator and a load. (b) The value of *f* by varying V_1 and V_2 for fixed dual variables.

Remark 1: There are several implications for the uniqueness of $\mathbf{V}^*$. First, from (3), it implies that the Lagrange dual function is smooth at the optimality of (2). Second, the smoothness of the Lagrange dual at optimality implies that optimal dual variables are unique and stable, and thus can be suitably used as power and line prices in pricing schemes. Third, it enlarges the space of designing simple local algorithms with low complexity to solve (2), and we address this in the following section.

4. DISTRIBUTED ALGORITHMS

In this section, we explain how computationally fast local algorithms can be designed to solve (2) as proposed in [10].

Local Algorithms

We first solve the partial Lagrangian minimization problem for given {$\boldsymbol{\lambda}$, $\boldsymbol{\mu}$}, and then use a projected gradient method in [11] to update the dual variables iteratively. We propose the following fixed point algorithm that computes the fixed point $\mathbf{V}$ in (3) for a given set of feasible dual variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$.

```
Algorithm 1:
Compute voltage $\mathbf{V}$ :
    $V_i(k+1) = \max\left\{ \underline{V}_i, \min\left\{ \overline{V}_i, \sum_{j \in \Omega_i} B_{ij} V_j(k) \right\} \right\},$    (5)
for all $i \in \mathcal{N}$, where $B_{ij}$ is given in (4) for all $i, j$.
```

Theorem 1: Suppose (2) has a unique optimal solution. Then, given any $\mathbf{V}(0)$ which satisfies $\underline{\mathbf{V}} \leq \mathbf{V}(0) \leq \overline{\mathbf{V}}$, $\mathbf{V}(k)$ in Algorithm 1 converges to the unique optimal solution of (2).

Then, to obtain $\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}$, projected gradient method can be applied to update λ_i and μ_{ij}, $\forall\, i \in \mathcal{N}, \forall\ (i, j) \in \mathcal{E}$. Our local algorithms is run in a distributed manner using message passing to transmit and receive iterates of the voltage and dual variables with neighboring one-hop nodes. This is illustrated in Figure 2 for the IEEE 5-bus system.

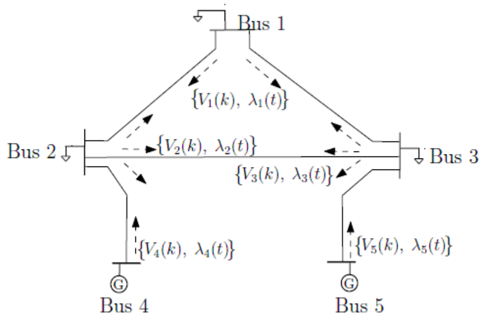


Fig. 2. Message passing in the IEEE 5-bus system.

In Figure 3, we show the iteration result of our algorithms in the IEEE 5-bus system (i.e., Figure 2) in low and high line capacity cases. From Figures 3, we observe that our algorithms have fast convergence time. From Figures 3, we observe that our local algorithms have fast convergence time. Moreover, the high capacity case converges much faster because it only needs to update λ_i for each node i.

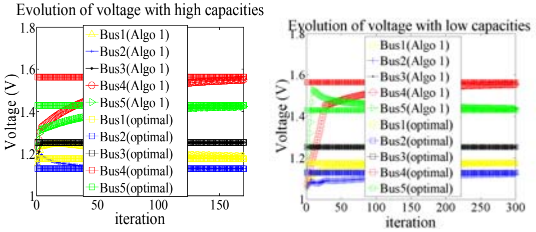


Fig. 3. Illustration of the convergence of our local algorithms in the IEEE 5-bus system. The left figure shows the voltage convergence in the high capacity lines scenario. The right figure shows the voltage convergence in the low capacity lines scenario.

Further Discussions

We have proposed the design of a distributed algorithm by leveraging dual decomposition and the solution uniqueness object model to that of a background model associated with the network topology. This intriguingly simple fixed-point algorithm converges very fast empirically. We will study more general fixed-point extensions of (5), i.e., $V_i(k+1) = \mathcal{P}\{\sum_{j \in \Omega_i} B_{ij} V_j(k)\}$ where $\mathcal{P}\{\cdot\}$ is a general projection operator onto a nonlinear feasible set as well as its asynchronous iteration. This approach can potentially enlarge the design space of algorithms with global behavior. It also highlights the role of network inter-connectivity (the nonnegative matrix $\mathbf{B}$ captures the network connectivity), and provide insights to designing algorithms with good convergence properties that can be much faster and have lower complexity than the state-of-the-arts SDP approach in [7], [12]. In addition, it is also interesting to explore other forms of decomposition such as primal decomposition to this problem which can lead to different protocol implementation requirement for the resistive network OPF problem.

ACKNOWLEDGEMENT

The authors acknowledge helpful discussions with Steven Low at the California Institute of Technology.

REFERENCES

[1] J. Carpentier. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 3(8):431–447, 1962. In French.

[2] M. Huneault and F. D. Galiana. A survey of the optimal power flowliterature. *IEEE Trans. on Power Systems*, 6(2):762–770, 1991.

[3] R. D. Christie, B. F. Wollenberg, and I. Wangensteen. Transmission management in the deregulated environment. *Proc. of IEEE*, 88(2):170–195, 2000.

[4] A. J. Conejo and J. A. Aguado. Multi-area coordinated decentralized DC optimal power flow. *IEEE Trans. on Power Systems*, 13(4):1272–1278, 1998.

[5] P. N. Biskas, A. G. Bakirtzis, N. I. Macheras, and N. K. Pasialis. A decentralized implementation of DC optimal power flow on a network of computers. *IEEE Trans. on Power Systems*, 20(1):25–33, 2005.

[6] S. Boyd and L. Vanderberghe. *Convex Optimization*. Cambridge University Press, 2004.

[7] J. Lavaei and S. H. Low. Zero duality gap in optimal power flow problem. *IEEE Trans. on Power Systems*, 27(1):92–107, 2012.

[8] J. Lavaei, A. Rantzer, and S. H. Low. Power flow optimization using positive quadratic programming. *Proc. of 18th IFAC World Congress*, 2011.

[9] A. Simsek, A. E. Ozdaglar, and D. Acemoglu. Uniqueness of generalized equilibrium for box constrained problems and applications. *Proc. Of Allerton*, 2005.

[10] C. W. Tan, D. W. H. Cai, and X. Lou. DC optimal power flow:
Uniqueness and algorithms. *IEEE SmartGridComm 2012 Workshop - Architectures and Modeling for the Smart Grid (Invited Paper)*, 2012.

[11] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont,
[12] S. Bose, D. F. Gayme, S. H. Low, and K. M. Chandy. Optimal power flow over tree networks. *Proc. of Allerton*, 2011.

Chee Wei Tan (M'08) received the M.A. and Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2006 and 2008, respectively. Currently, He is an Assistant Professor at City University Hong Kong.

Desmond W. H. Cai (S'10) received the B.Sc. degree in electrical and computer engineering (with highest distinction) from Cornell University, Ithaca, NY, in 2009. He is currently working towards the Ph.D. degree in electrical engineering at the California Institute of Technology, Pasadena.

Xin Lou received the B.Eng. degree in communication engineering from Sichuan University, China, in 2009. He is currently working towards the Ph.D. degree at the City University of Hong Kong.

Customer Energy Management System in Smart Grid

Guang-Hua Yang and Victor O.K. Li
The University of Hong Kong, Hong Kong, China
{ghyang, vli}@eee.hku.hk

1. Introduction

Recently, Smart Grid (SG) has attracted much interest from governments, power companies, and research institutes [1], [2], [3]. Compared to the traditional electrical grids, by employing advanced information technologies, SG can achieve higher efficiency, better reliability, stability and sustainability, lower total infrastructure investment, better exploitation of renewable energies, lower power consumption and lower greenhouse gas emission. An energy management system (EMS) monitors, controls and optimizes the performance of energy generation, transmission, distribution, and consumption. EMS is an important building block of an SG, and plays a key role in achieving the advantages of an SG.

The Conceptual Reference Model of SG proposed by NIST [4] divides an SG into seven domains, specifically, customers, markets, service providers, operations, bulk generation, transmission, and distribution. Among the seven domains, EMS for energy generation, transmission, and distribution have been studied for decades. Many models, standards, protocols and systems have been proposed, implemented, and deployed in practical systems. However, EMS in the consumer domain is largely neglected in existing studies.

In this work, we focus on the EMS in the customer domain (for simplicity, we use EMS to denote EMS in the customer domain in the rest of the paper). The desired features and design issues are discussed, the characteristics are highlighted and the challenges are identified. An innovative framework, namely, Pervasive Service-Oriented Networks (PERSON), is proposed to address the major challenges of heterogeneity, distributed system and dynamicity. As far as we know, this is the first framework for the customer EMS. The effectiveness of PERSON and customer EMS are demonstrated in Demand Response (DR) [5] application.

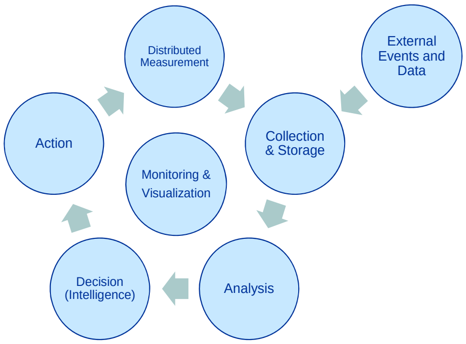


Figure 1 EMS control flow chart.

2. Customer energy management system

2.1 Features of customer EMS

Existing EMSs in the customer domain are generally simple monitoring and control systems, which ignore the challenges of an SG and thus only have limited capability and limited intelligence. To support the advantages of SGs, an EMS is required to have some basic features:

- Supports various existing actors as well as emerging actors.
- Continuously monitors the energy consumption at different granularities, such as home level and appliance level.
- Continuously monitors environmental parameters, such as temperature and humidity, which can be exploited for context-based intelligent control.
- Supports automatic and manual control of the actors.
- Supports the integration of renewable power sources, such as solar and wind.
- Interacts with actors in other domains to realize advanced features, such as DR.

Besides the basic features, some desired advanced features include:

- Intelligence and efficiency: ability to achieve optimized performance under dynamic situations.
- User friendliness: plug-and-play with self-configuration capability.
- High reliability and durability: robustness and self-healing capability after system failures.

- Low cost and low power consumption.

2.2 EMS control flow

The features listed above are from the perspective of a user. From the perspective of system control, EMS can be abstracted as a closed-loop control system as shown in Fig.1. The control flow is composed of several building blocks.

- Distributed measurement: Measurements include not only power consumption/generation data, but also environmental parameters, and users' preferences and behaviors. The measurements are conducted on different actors in a distributed manner.
- Collection & storage: Measurements are exchanged among actors, and stored in some aggregator for further analysis and processing.
- Analysis: The collected and stored data are analyzed to obtain historical and statistical information. In addition, the data is processed to obtain an up-to-date view of the system.
- Decision (intelligence): Based on the results of the analysis, energy management decisions are made via some intelligent algorithms.
- Action: Control decision is delivered to some actuators, such as switches and breakers, for execution.
- Besides the main closed-loop control flow, some building blocks are also necessary:
- Monitoring & visualization: This provides an interface and a user-friendly way to monitor and control the system.
- External events and data: This includes users' or utilities' control decisions, price of electricity, weather, natural disasters, etc.

2.3 Energy flow, information flow and challenges

If we take a closer look at the control flow, it is basically energy flow management utilizing information flow.

Energy flow

In the customer domain, energy is distributed through an electricity network in a tree-like structure. In general, the power meter sits at the root, while AC outlets, lights and appliances are at the leaves. The energy flow can be measured by energy monitoring devices, such as meters and power gauges; and can be manipulated by controllers, such as breakers and switches. Although the electricity network is a homogeneous network in terms of the way energy is distributed, monitored and controlled, the energy flow is quite dynamic in terms of quantity and quality. The dynamicity is due to the variations in supplies and demands, dynamic user behaviors, and continuously changing environments. In an SG, increasing usage of renewable energy sources, such as wind turbine and solar panels, makes the problem even more challenging.

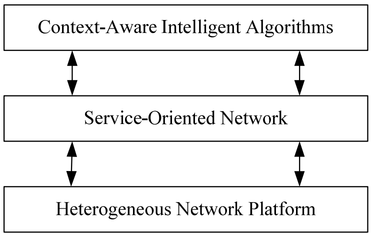


Figure 2. Three-layer structure of PERSON.

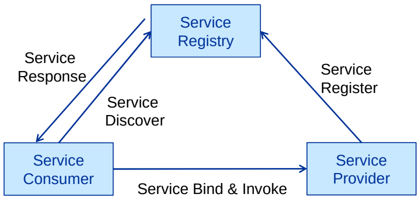


Figure 3. Relationships between actors in an SON.

Information flow

The information utilized in an EMS includes real-time measurements, historical data, external events, and control decisions. The information is exchanged among distributed actors through some kind of communication channels. The communication channels form a communication network, which is generally a heterogeneous and distributed network. Heterogeneity is unavoidable because different actors may follow different communication protocols, use different media (wire or wireless) and have different communication capabilities. The distributed nature is due to actors being physically dispersed throughout the space. The heterogeneity and distributed nature impede the formation of connected and efficient information flows, and thus are considered as two other major challenges of an EMS.

Thus, to effectively realize efficient energy management, the three major challenges are heterogeneity, distributed system, and dynamicity.

3. Pervasive Service-Oriented Network Model

3.1 Overview

To address the three major challenges, we propose a three-layer general framework to seamlessly integrate diverse kinds of actors and networks into a unified

Pervasive Service-Oriented Network. The principle is to decompose the complexities into different layers and to have loose coupling among different layers. Fig.2 shows the three-layer structure of PERSON.

3.2 Heterogeneous network platform (HNP)

The objective of HNP is to build a homogeneous communication infrastructure for the information flow. HNP provides simple APIs to the upper layer for information exchanging. The Upper layer does not care how the information is delivered. The underlying communication protocols, media and communication capabilities are transparent to the upper layer. An implementation of such HNP can be found in [6].

3.3 Service-oriented network (SON)

The basic idea of SON is to achieve interoperability, modularity and reusability by abstracting the functions provided by the actors into services [7]. Actors in an SON have different roles:

- ***Service provider*** The service provider creates a service and registers its interface and access information to the service registry. Each provider decides which service to offer, how to balance security and easy accessibility and how to price the service.
- ***Service consumer*** The service consumer discovers the services available in the network, locates services in the service registry and then binds to the service provider to invoke the services. A service consumer can access multiple services from multiple service providers at the same time.

For a certain application, such as EM, a suite of services needs to be defined. Meanwhile, mechanisms also need to be defined for service creation, registration, discovery, binding and invoking. Fig. 3 shows the relationships between different actors in an SON.

3.4 Context-aware intelligent algorithm

In this layer, the services provided by SON are exploited for intelligent control purposes. The interoperability and service reusability provided by SON facilitate the development of proactive and context-aware intelligence to address the challenge of dynamicity. With the context-aware intelligent algorithms, the overall performance of a system is optimized. Better service availability, more convenience, and higher efficiency are achieved.

4. EMS IMPLEMENTATION

4.1 EMS structure

Fig.4 shows the structure of an EMS in PERSON framework. It is composed of two major parts: Heterogeneous Home Area Network (HHAN) and Data and Service Center (DSC).

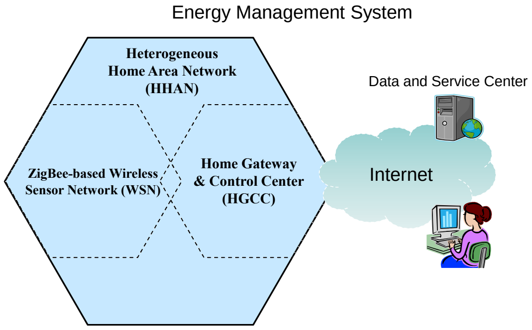


Figure 4. EMS structure.

Heterogeneous Home Area Network

HHAN realizes HNP, the first layer of PERSON. It provides the basic infrastructure of an EMS. HHAN is composed of a ZigBee-based Wireless Sensor Network (WSN) and a Home Gateway and Control Center (HGCC). ZigBee-based WSN is composed of diverse EM actors, such as power meter, sensors, switches and in-home display, to measure, control and visualize the energy flow, while HGCC is used to host local intelligence, provide interface between users and the EMS, and support communications among EMSs and DSC.

Data and Service Center

DSC is an actor outside the customer domain (generally located at the service provider domain). With DSC, customers can obtain broader and deeper awareness leading to enhanced potential for energy savings. A more important value of DSC is that it enables a tighter customer-grid collaboration, which not only facilitates power savings, but also provides more EM-related capabilities. These capabilities are necessary for realizing the advantages promised by an SG.

EMS services

To support the basic and desired features of an EMS, four kinds of services, specifically, Measurement, Control, Display, and Action, are defined. In addition, an XML-based mechanism is developed to support the creation, registration, discovery and consumption of the services. In the implementation, HGCC serves as the service registry of the SON.

4.2 Demand response application

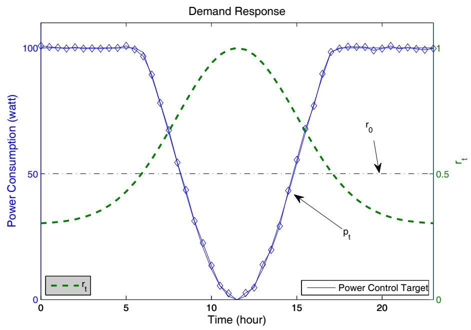


Figure 5. Demand response with EMS.

Demand response [5] is one major driving force of an SG. The basic idea of DR is to manage customer load in response to power supply constraints. By employing automatic load management, DR can realize peak shaving through reducing power load at peak times, or shifting power load from peak times to times with lower demand.

An effective DR imposes some functional requirements on the underlying EMS, such as capabilities of real-time load monitoring, two-way data communication between the demand side and utilities, data processing and demand side load control. All these functional requirements are satisfied by our EMS. Exploiting the EMS for DR application is a good demonstration of the system capabilities.

The power supply limit and instantaneous total power load are known at DSC. We define the instantaneous load level r_t as $r_t = \frac{\text{Total power load}}{\text{Power supply limit}}$. DSC distributes r_t to HGCCs. Given r_t, HGCC calculates the load control target p' and places the load control orders to the controllable power loads.

In Fig. 5, r_t, p' and the load measurement results p_t are shown. We can find that p_t matches p' very well. In addition, p_t and p' show a good response to the varying r_t as expected. The results demonstrate the capabilities of the EMS in the context of DR, and the effectiveness of the PERSON framework for EMS.

4. Conclusions

In this paper, we study customer EMS in SG. The special challenges of heterogeneity, distributed system and dynamicity are identified. PERSON, a novel framework, is proposed to address the three challenges. Further, we implement a cost-effective EMS prototype to realize PERSON and show its effectiveness.

References

[1] American Recovery and Reinvestment Act of 2009, P. L. 111-5, USA.

[2] "Smart Grid Policy", [Docket No. PL09-4-000], Federal Energy Regulatory Commission, USA, July 16, 2009.

[3] "Title XIII - Smart Grid, Sec. 1301, Statement of Policy on Modernization of Electricity Grid", Energy Independence and Security Act of 2007 (EISA), USA.

[4] National Institute of Standards and Technology (NIST), "NIST Framework and Roadmap for Smart Grid Interoperability Standards," Release 1.0, Office of the National Coordinator for Smart Grid Interoperability, USA, January 2010.

[5] "Harnessing the Power of Demand - How ISOs and RTOs Are Integrating Demand Response into Wholesale Electricity Markets", Markets Committee of the ISO/RTO Council, October 16, 2007.

[6] V. O.K. Li, C. Li, Q. Liu, G.-H. Yang, Z. Zhao, K.-C. Leung, "A Heterogeneous Peer-to-Peer Network Testbed", in the First International Conference on Ubiquitous and Future Networks (ICUFN 2009), Hong Kong, June 7-9, 2009.

[7] Bell, Michael, "Introduction to Service-Oriented Modeling," Service-Oriented Modeling: Service Analysis, Design, and Architecture, Wiley & Sons. pp. 3. ISBN 978-0-470-14111-3, 2008.

Guang-Hua Yang received his Ph.D. degree from the Department of Electrical and Electronic Engineering in the University of Hong Kong in 2006. From 2006, he served as post-doctoral fellow, research associate, and project manager in the University of Hong Kong. Dr. Yang's research interests are in the general areas of communications, networking and multimedia.

Victor O.K. Li received SB, SM, EE and ScD degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1977, 1979, 1980, and 1981, respectively. He is Chair Professor of Information Engineering and Head, Department of Electrical and Electronic Engineering at the University of Hong Kong, Hong Kong. He also served as Managing Director of Versitech Ltd., the technology transfer and commercial arm of the University. Previously, he was Professor of Electrical Engineering at the University of Southern California (USC), Los Angeles, California, USA, and Director of the USC Communication Sciences Institute. He also serves on various corporate boards. Sought by government, industry, and academic organizations, he has lectured and consulted extensively around the world. His research interest is in the core technologies and applications of information technologies. He is a Fellow of the IEEE, the IAE, and the HKIE.

Decentralized Information Processing and Energy Management for Smart Grid

Xiao Li and Anna Scaglione
University of California, Davis, USA
{eceli, ascaglione}@ucdavis.edu

1. Introduction

Infrastructures for cyber-physical systems become increasingly networked over wide areas and comprise a variety of sensing modalities, such as the Smart Grid. Traditionally, power systems use the Supervisory Control and Data Acquisition (SCADA) model for data processing and management in the transmission grid. Sensor data in the transmission grid, such as bus injections and power flows, phasor measurements, and circuit breaker statuses are queried by a collection of software in Energy Management Systems (EMS) at the control centers [1].

Generally speaking, the periodicity of SCADA systems is about a few seconds. However, today the control actions are centralized and relatively slow, precisely due to bottlenecks in accessing the data from distributed databases and processing them at the application servers, as well as the huge computational burdens imposed on the application side. Therefore, instead of real-time controls and maintenance, the way that failures and instability are averted and managed today is usually by setting wide margins of operations, using redundancy that allows local controls to isolate individual components and activate a path on standby.

Ongoing market deregulation, progressive penetration of renewable resources, demand response possible malicious physical or cyber attacks are exacerbating the need to provide reliable access to information, as well as efficient processing schemes for data analytics. The need for databases is expanding in part due to Wide Area Measurement Systems (WAMS), a name coined in recent years to refer to the enhanced infrastructure for data acquisition and system control under study, and in part due to the advanced metering infrastructure (AMI) that is expected to be deployed in the power distribution network to enhance the development of Home Energy Management Systems (HEMS). WAMS are going to be used for wide-area monitoring and analysis, complementing existing SCADA/EMS systems and, possibly, for wide area control. Furthermore, AMI in the distribution grid together with SCADA/WAMS data on the transmission grid produce an avalanche of information, upon which the trend of developing decentralized architectures for data management becomes inevitable [1][2].

In particular, future smart grid intends to modernize the transmission and distribution networks by integrating more intelligence into the cyber-infrastructure, which poses both great opportunities and challenges to the existing legacy systems. The main goals are to harness computation power and communication technologies to realize wide-area monitoring and distributed control for transmission networks, and to facilitate the integration of renewable sources by demand side management for distribution networks. While it is possible to attain these practical goals by aggregating local information hierarchically and making decisions centrally, we contend that several benefits could be gained if the associated data support platform and more importantly, the information processing modules could be engineered differently from traditional architectures.

In this letter, we introduce the vision of network signal processing for smart grid and summarize briefly some of our developed tools and results for both transmission and distribution grid.

2. Wide-Area Monitoring in Transmission Grids

Monitoring in transmission grids is usually by obtaining the state of power systems via state estimation from SCADA/WAMS data. Most of the existing methods for state estimation requires each distributed control area to solve for a local state, relying on redundant local measurements, and then refine the local estimates in a hierarchical manner by leveraging on the tie-line structure and tuning the estimates on boundary buses with those from neighboring areas (see e.g., [3]-[10]).

Although these methods do alleviate the computational burdens at control centers, they rely on aggregation trees that require coordination and depend on the power grid topology. This is a limitation in reconfiguring the system, if failures or attacks call for it. To make the future smart grid more resilient in terms of information access and processing, we propose a decentralized architecture for state estimation via gossiping, as discussed in [11].

Gossiping Algorithms.

Gossip-based algorithms are very appealing for their intrinsic resilience to communication failures. They also provide a gradual degradation of the estimate as communications resources become scarcer, and continue to work even if the network topology changes. This is true for both deterministic and randomized gossip protocols. Furthermore, gossip-based algorithms improve the scalability compared to centralized schemes, because they eliminate the overhead of central coordination. This is well suited for smart transmission grid that needs to be resilient to attacks.

Proposed Decentralized Monitoring Architecture.
Since many EMS applications have common queries for the state, state estimation can be embedded as a fundamental data analytics for the data support layer. The individual databases push and pull information from/to peer databases to compute the state via gossiping, giving each database an online copy of the power system state that can be used for diagnostics and anomaly detections. Furthermore, since the power measurements across the whole power grid are closely coupled with each other through the state, the wide-area state awareness can help encode, compress and store the measurements efficiently in the database [11].

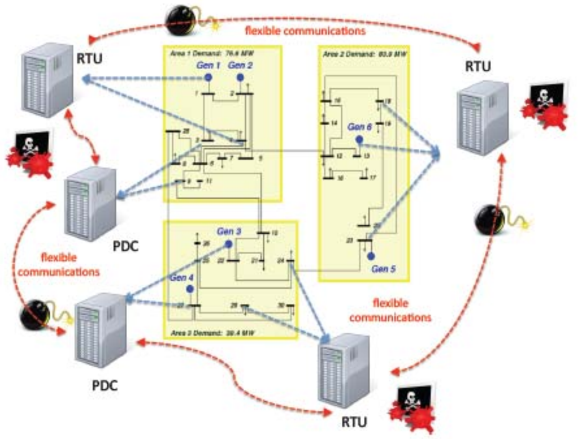


Fig.1 – *Decentralized Architecture for State Estimation*

This consideration leads to the architecture in Fig. 1, resembling a peer-to-peer (P2P) model among the distributed databases, where the data are acquired locally in Remote Terminal Units (RTU) or Phasor Data Concentrators (PDC). The state estimation function is embedded as a stored procedure in the database management system, which keeps running at a lower level than that of the client applications. The system becomes a State-Aware (SA) Decentralized Database Systems (DDBS), with other applications as clients accessing state information and stored data.

Potential Benefits.
The most obvious benefits of embedding the state estimation as a stored procedure in the databases are
1) each distributed area achieves wide area state awareness by local communications and processing;
2) reduced network usage and processing latency for recurrent applications that query state information;
3) improved security and access control by designing a trust model to evaluate trustworthy peers for gossiping.
While the SA-DDBS is considered for the transmission grid, our models are actually valid for both the transmission and the distribution network. In fact, the AMI database system has fewer legacy systems in place and, therefore, possibly more suitable to be designed differently. In the following, we discuss the main goal of demand response in distribution grids by having coordinated home energy management systems (Co-HEMS) to schedule and control electric loads in different households.

3. Demand Side Management in Distribution Grids

With the penetration of smart appliances and electric vehicles in households, there is a trend to control these loads intelligently for emergency or market purposes. Many options have been proposed for demand side management or demand response techniques, starting from the initial *load curtailment model* to *price-based load control models*. Details of demand response and demand side management techniques can be found in our magazine paper [13] and the references therein, hence we do not repeat the enumeration of literature due to space limitation.

Most load curtailment programs focus on large entities such as industrial plants and factories that can accommodate enough electricity in case of emergency, while there are much more degrees of freedom from individual households in distribution grids. On the other hand, the problems with price-based control models are mainly due to the uncertainty in how customers respond to variable real-time prices, i.e., an unknown feedback behavior. If this feedback loop is somehow opened or its responses are based on settings that can be learned by a control center, the stated problems will no longer exist (or will be mitigated).

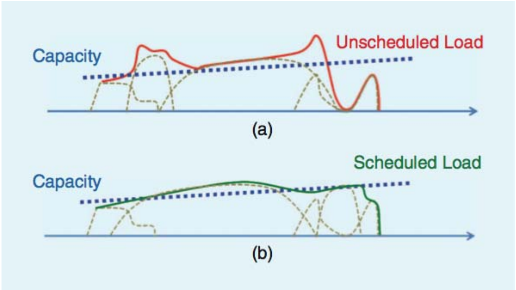


Fig.2 – *Load Scheduling*

The key idea we proposed is to unbundle the load from certain appliances and view each contribution like a set of LEGO pieces that can be reassembled to follow a desired load profile, by delaying appropriately the power delivery to each individual appliance (see Fig. 2).

From Packet Switching to Power Switching.
In our model, each arriving smart appliance has an associated parameter vector that determines uniquely the time evolution of the load contribution when that appliance is turned on. A simple example is that of EVs, for which the vector is a two-dimensional vector, representing the charging rate and the fraction of battery charge needed by the car upon its arrival.

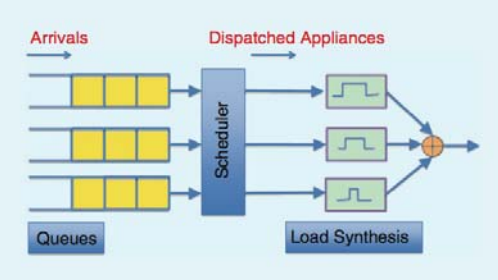


Fig.3 – *Digital Direct Load Scheduling (DDLS)*

What makes our model scalable and practical though, is the quantization step: the continuous load injection parameters are mapped onto a quantized request, with Q quantization levels. These quantizations will bundle different requests by appliances into a set of discrete load classes. Hence, the scheduling design consists of managing the departures from a set of Q queues, with a FIFO discipline, as in Fig. 3. For example, with EVs, the charge duration has a maximum of 8 hours and can be quantized into 15-minute intervals, thus requiring 5 bits to communicate and resulting in 32 different queues where EVs can be bundled to wait for energy.

The DDLS model described above is a solution to the load scheduling problem at the aggregator level from a centralized perspective. More details can be found in the article [12]. As an example, a typical configuration of the architecture is depicted in Fig. 4.

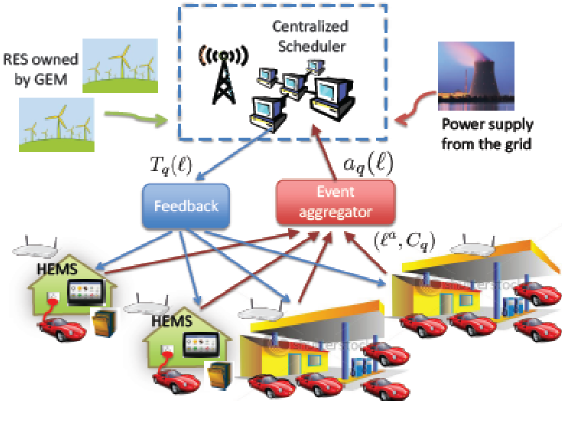


Fig.4 – *Centralized Load Scheduling at an Aggregator*

In the following, we introduce a fully distributed load scheduling scheme that is coordinated among different residential units.

Distributed and Coordinated Load Scheduling.
Instead of managing smart appliances centrally at aggregators, there is extensive literature emerging on Home Energy Management Systems (HEMS) [14][15]. In these works, researchers look into finding optimal designs for the software and hardware suited for residential use that would schedule smart loads in an automated fashion. These HEMS units receive requests from their owners specifying the appliances they plan to use in the near future and their preferences.

We have recently proposed a distributed scheme for load scheduling problems by collaborating with neighbor HEMS [16]. The software runs an optimization that plans the use of these appliances, based on their power consumption, job deadlines and other customer specified factors, taking into account the dynamic price made available to the unit from its associated aggregator or retailer, as shown in Fig. 5.

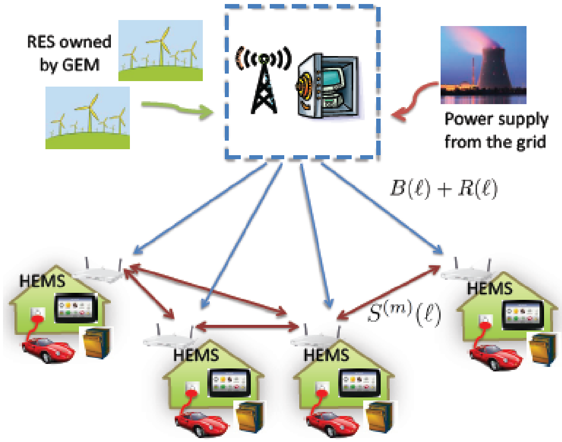


Fig.5 – *Distributed Scheduling by Coordinated HEMS*

The above coordinated HEMS architecture manages the home energy scheduling of multiple residential units in order to reduce the real-time market cost of the retailer by minimizing collectively the deviation of the aggregate load from a certain generation profile. As shown in [16], the proposed coordinated HEMS can effectively achieve real-time power balancing, mimicking the benefits of the centralized DDLS in a fully distributed manner.

4. Conclusion

In this paper, we surveyed some results on distributed network processing for smart grid for wide-area monitoring in the transmission grid and scalable demand response in the distribution grid, which are two important aspects in future power grids. In fact, there are many more control problems and decision-making processes in power systems that will require distributed implementation due to the expanding scale of power systems and increasing penetration of smart appliances and renewable generation. This line of research can help contribute to pave the way for a greener future for the grid.

5. Acknowledgement

The works surveyed in this paper are also contributed by Miss Mahnoosh Alizadeh, Dr. Tsung-Hui Chang and Dr. Zhifang Wang, who have graciously agreed to have their results discussed in this overview.

References

[1] F. F. Wu, K. Moslehi and A. Bose, “Power System Control Centers: Past, Present and Future”, (invited paper), Proc. of the IEEE, 93(11):1890-1908, Nov. 2005.

[2] S. Kirti, Z. Wang, A. Scaglione and R. J. Thomas, “On the Communication Architecture for the Control Problem in Power Networks,” HICSS-40, Jan 2007.

[3] T. Yang, H. Sun, and A. Bose, “Transition to a Two-Level Linear State Estimator : Part I & II,” IEEE Trans. Power Syst., no. 99, pp. 1–1, 2011.

[4] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles, “A Multilevel State Estimation Paradigm for Smart Grids,” Proceedings of the IEEE, no. 99, pp. 1–25, 2011.

[5] D. Falcao, F. Wu, and L. Murphy, “Parallel and Distributed State Estimation,” IEEE Trans. Power Syst., vol. 10, no. 2, pp. 724–730, 1995.

[6] S. Lin, “A Distributed State Estimator for Electric Power Systems,” IEEE Trans. Power Syst., vol. 7, no. 2, pp. 551–557, 1992.

[7] A. Phadke, J. Thorp, and K. Karimi, “State estimation with phasor measurements,” Power Engineering Review, IEEE, no. 2, pp. 48–48, 1986.

[8] R. Ebrahimian and R. Baldick, “State Estimation Distributed Processing,” IEEE Trans. Power Syst., vol. 15, no. 4, pp. 1240–1246, 2000.

[9] L. Zhao and A. Abur, “Multi-area State Estimation using Synchronized Phasor Measurements,” IEEE Trans. Power Syst., vol. 20, no. 2, pp. 611–617, 2005.

[10] T. Van Cutsem, J. Horward, and M. Ribbens-Pavella, “A Two-Level Static State Estimator for Electric Power Systems,” IEEE Trans. Power App. Syst., no. 8, pp. 3722–3732, 1981.

[11] Xiao Li, Z. Wang and A. Scaglione, “Decentralized Data Processing and Management in Smart Grid via Gossiping”, IEEE SAM 2012, New Jersey, USA.

[12] M. Alizadeh, A. Scaglione, and R.J. Thomas “From Packet to Power Switching: Digital Direct Load Scheduling”, Selected Areas in Communications, IEEE Journal on, 30(6):1027 -1036, 2012.

[13] M. Alizadeh, X. Li, Z. Wang, A. Scaglione, and R. Melton, “Demand-Side Management in the Smart Grid: Information Processing for the Power Switch,” Signal Processing Magazine, IEEE, 29(5):55--67, 2012.

[14] D.-M. Han and J.-H. Lim, “Design and implementation of smart home energy management systems based on zigbee,” IEEE Trans. Consumer Elect., vol. 56, no. 3, pp. 1417 –1425, 2010.

[15] A. Mohsenian-Rad and A. Leon-Garcia, “Optimal residential load control with price prediction in real-time electricity pricing environments,” IEEE Trans. Smart Grid, vol. 1, no. 2, pp. 120 –133, Sept. 2010.

[16] Tsung-Hui Chang, Mahnoosh Alizadeh, and Anna Scaglione, "Real-time power balancing via decentralized coordinated home energy scheduling", submitted to IEEE Trans. Smart Grid (3rd round review)

Xiao Li has been pursuing the Ph.D. degree at the University of California, Davis, since 2009. He received the B.Eng. degree (2007) from Sun Yat-Sen (Zhongshan) University, China, and the M.Phil. degree (2009) from the University of Hong Kong. His research interests are in compressive signal processing for big data, distributed optimizations and adaptive learning algorithms, with applications in communications, networking, and power systems.

Anna Scaglione received the Laurea (M.Sc. degree) in 1995 and the Ph.D. degree in 1999 from the University of Rome, “La Sapienza.” In 2001, she was at Cornell University, Ithaca, New York, where she became an associate professor in 2006; prior to joining Cornell she was an assistant professor at the University of New Mexico from 2000 to 2001. She is currently a full professor in the Department of Electrical and Computer Engineering at the University of California, Davis, where she joined in 2008. She has held several editorial posts and is currently the editor in chief of IEEE Signal Processing Letters and guest editor of IEEE Journal on Selected Areas in Communications, for a series of special issues on the smart grid. She received the 2000 IEEE Signal Processing Transactions Best Paper Award, and the 2002 NSF Career Award, and she was also the co-recipient of the Ellersick Best Paper Award (MILCOM 2005). Her expertise is in the broad area of signal processing for communication systems and networks. Her current research is on communication and wireless networks, sensors systems for monitoring, control and energy management, and network science. She is an IEEE Fellow.

Toward Evolved Energy Trading Mechanisms in the Smart Grid

Yunpeng Wang and Walid Saad
Electrical and Computer Engineering Department, University of Miami, Coral Gables, FL 33146
{y.wang68@umiami.edu, walid@miami.edu}

1. Introduction

The electric power grid is undergoing a period of unprecedented changes. The introduction of advanced communication and control technologies has transformed the conventional, hierarchical electric grid, into a smarter, large-scale, and stochastic *smart grid* system. On the one hand, the introduction of renewable energy sources and micro-grids has changed the nature of energy production and distribution; on the other hand, the smart grid provides the consumers and utility companies with the ability to better manage and optimize their energy usage and production via new techniques such as demand-side management. In this respect, numerous recent works have studied smart grid from different perspectives ranging from energy generation to smart metering, control, and security [1].

One key challenge in the smart grid is the modeling, analysis, and design of the *energy trading mechanisms* that arise between the various smart grid players: substations, renewable energy sources, electric cars, consumers, among others. In conventional power systems, energy trading involved only a limited number of buyers/sellers (e.g., substations/production units/customers) and was governed by a hierarchical, centralized market design [2,3]. In contrast, in the smart grid, the ability of consumers to control their energy usage coupled with the rapid deployment of renewable energy sources provides new opportunities for energy trading. In particular, the smart grid will encompass distributed, localized energy trading markets in which various types of participating players (micro-grids, consumers, substations) can dynamically decide on how and when to manage, trade, or store their energy [1].

To seize these opportunities, a number of technical challenges must be addressed that arise from: a) the heterogeneity in the type and capabilities of the smart grid's elements that can include micro-grids, smart meter-enabled consumers, and traditional substations, b) the grid density, c) the constraints on the underlying power system and communication infrastructure, and d) the need for efficient pricing mechanisms. Consequently, moving from the conventional, hierarchical power systems toward smarter power grids mandates a paradigm shift towards distributed energy trading mechanisms that can efficiently address the aforementioned challenges. In particular, there is a need for novel tools to: a) properly model the interactions between heterogeneous players, with possibly different objectives and capabilities and b) develop novel distributed and self-organizing algorithms to optimize the overall efficiency of the energy trading and pricing processes. To this end, *game theory* provides a solid mathematical framework suitable for modeling and analyzing the complex energy trading mechanisms foreseen in future energy markets [4-8].

The main goal of this short letter is to provide insights on how to develop game-theoretic models for managing the energy trading processes in emerging smart grid markets. While we acknowledge that the road toward fully distributed and dynamic energy trading mechanisms is contingent upon providing an advanced power and communication infrastructure, in this short letter, we restrict our attention to the economical aspects of energy trading. In particular, we focus on modeling and analyzing how a number of grid elements can decide on whether to sell or buy energy, given the demand/generation and associated tradeoffs in terms of pricing, revenues, and costs. A more comprehensive survey on the applications of game theory in the smart grid is found in [5].

2. System model

Consider a number of smart grid nodes, including micro-grids (e.g., wind/solar farms, or PHEVs), substations, and consumers with storage elements that can trade energy in a locally formed market. Depending on the generation and demand, each node must decide on whether to sell, buy, or store energy. These decisions can be time-varying, however, for simplicity, in this basic model, we consider the energy trading market at a fixed time period during which the grid is clearly split into N sellers and K buyers.

The need for the buyers to acquire energy so as to meet their demand coupled with the potential revenues that the sellers can reap from selling their energy surplus give rise to an interesting energy trading market. Here, the buyers and sellers must interact so as to decide on the amount of energy they are willing to exchange as well as the ensuing trading price. The decision of each seller or buyer depends, not only on the price and energy needs/surplus, but also on the physical capabilities of the node. For example, a node having a storage device would want to participate in energy trading while also minimizing the amount of charging and discharging of the storage device.

To capture the interactions resulting from such energy trading market, we propose a game-theoretic formulation composed of *two interdependent games*: a noncooperative game in which each buyer or seller decides on the quantity that it wishes to trade, depending on the associated benefits and costs and an auction mechanism that determines the resulting trading parameters (price, quantities). Clearly, as the quantities brought forward by the buyers and sellers determine the price, there is a close coupling between these two games.

For illustrative purposes, here, we consider a special case in which only the sellers can dynamically change the quantities that they out for sale. The buyers are thus considered to have fixed quantities and trading bids. Hence, a noncooperative strategic game between the sellers is formulated. The strategy of each seller is to decide on the *maximum* quantity $a_i \in A_i$ that it wishes to sell, so as to maximize the following utility function [6]:

$$U_i(a_i, \mathbf{a}_{-i}) = (\bar{p} - s_i)Q_i(\mathbf{a}) - f_i(Q_i(\mathbf{a})) \qquad (1)$$

where $\mathbf{a}_{-i}$ is the vector of maximum quantities set by all sellers other than *i*, $Q_i(\mathbf{a})$ is the total energy sold by seller *i*, s_i is a reservation price below which *i* is not interested in trading, $f_i(\square)$ is a cost function dependent on the type of seller *i*, and $\bar{p}$ is the trading price.

The trading price $\bar{p}$ and the associated quantities $Q_i(\mathbf{a})$ for all *i* are the result of a *double auction mechanism* formulated between the *N* sellers and *K* buyers. In a market with multiple sellers and multiple buyers, the double auction is a suitable framework for determining the trading prices, the number of buyers and sellers that will get involved in the trade, and the associated quantities. Given: a) the maximum quantities that each seller (buyer) is willing to put into the market (buy) and b) the reservation prices (bids) of the sellers (buyers), a double auction allows to efficiently determine the resulting price and traded quantities.

To solve the double auction, the sellers (buyers) are arranged in an increasing (decreasing) order of their reservation price (bids). Then, the supply (solid line) and demand (dashed line) curves are plotted, as illustrated in Figure 1. The intersection of these two curves can be used to determine the trading price and the quantities. The details on how this is determined can be found in [6]. Once the trading price and quantities are determined from the double auction, the next step is to find a solution for the proposed game.

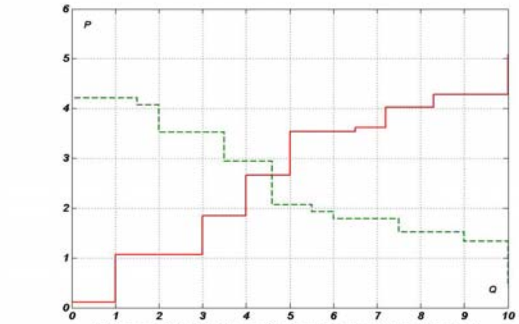


Figure 1 A Multi-unit Double Auction Market

3. Proposed Game Solution and Basic Result

One popular solution concept to solve a noncooperative game is through the concept of *Nash equilibrium (NE)*. The Nash equilibrium is a state of the game in which no player can improve its utility by unilaterally changing its strategy, given that the other players maintain their current strategies [8]. Mathematically, a vector of strategies $\boldsymbol{a}^*$ constitutes a Nash equilibrium, if and only if, it satisfies:

$$U_i(a_i^*, \mathbf{a}_{-i}^*) \geq U_i(a_i, \mathbf{a}_{-i}^*) \quad \forall i, a_i \in A_i .$$

The existence of an NE is not always guaranteed. For the proposed game, we observe that, although the strategy sets are convex, the utility function can include discontinuities due to the trading price resulting from underlying double auction. Therefore, the classical NE existence results cannot be applied [8]. However, in our recent work [7], we have shown that, despite the discontinuity of the utility function; an NE is guaranteed to exist, although it may not be unique.

To find one such NE, we can develop an algorithm based on the idea of a *best response*. Essentially, the best response of any player *i* is the strategy choice that maximizes its utility in (1), given the current vector of strategies $\mathbf{a}_{-i}$ of the other players.

In Figure 2, we show a simulation result highlighting the performance resulting from the proposed best response-based algorithm, for different numbers of sellers. Performance is measured in terms of the average utility per seller. Figure 2 shows that the proposed game-theoretic approach yields significant gains relative to a classical greedy-based energy trading algorithm. Our results in [6] also show a reasonable convergence time.

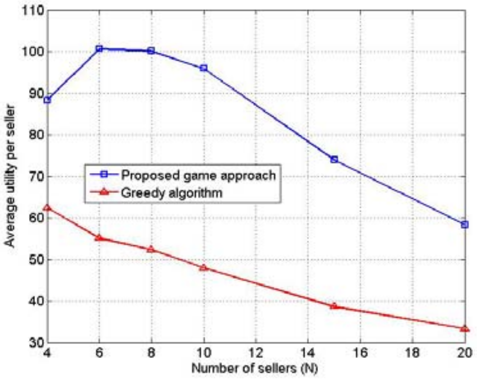


Figure 2. Performance of proposed approach.

4 Conclusions and Future Work

In this short letter, we developed a basic game theoretic model, combining auction theory with noncooperative game, to gain some insights on the potential of local energy trading in the smart grid. This basic model showcases the potential of game-theoretic approaches for enabling evolved energy trading mechanisms in the smart grid. However, this letter only scratched the surface of an emerging area in the smart grid in which many future works can be foreseen.

First, one key direction is to study scenarios in which both buyers and sellers are strategic. Second, it is of interest to define novel game-theoretic equilibrium concepts that can connect multiple, co-existing games. Third, in energy markets, a certain hierarchy might be imposed on the energy trading processes. For example, the main grid substations can take the role of leaders in the market, in which case a Stackelberg formulation such as in [9] is appropriate. Fourth, in the presence of time-varying demand and supply, one must incorporate dynamics into the developed game via a differential/dynamic game formulation [8]. Finally, one major open research problem is to study how the process of energy trading evolves under practical constraints on the underlying power and communication infrastructure.

References

[1] X. Fang; S. Misra, G. Xue, D. Yang, “Smart Grid — The New and Improved Power Grid: A Survey,” *IEEE Communications Surveys & Tutorials*, vol.14, no.4, pp.944-980, 2012.

[2] R. Buyya, and S. Vazhkudai, "Compute Power Market: towards a market-oriented grid," Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on, pp.574-581, 2001.

[3] D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*, Wiley, 2004.

[4] A. Mohsenian-Rad, V.W.S. Wong, J. Jatskevich, R. Schober and A. Leon-Garcia, “Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid,” *IEEE Trans. On Smart Grid* , vol.1, no.3, pp.320-331, Dec. 2010

[5] W. Saad, Z. Han, H.V. Poor and T. Basar, “Game Theoretic Methods for the Smart Grid,” *IEEE Signal Processing Magazine, Special Issue on the Smart Grid,* vol. 29, no. 5, pp. 86 – 105, Sep. 2012.

[6] W. Saad, Z. Han, H.V. Poor, and T. Basar, “A noncooperative game for double auction-based energy trading between PHEVs and distribution grids,” *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Brussels, Belgium Oct. 2011

[7] Y. Wang, W. Saad, Z. Han, H.V. Poor, and T. Basar, “A game-theoretic model for energy trading in the smart grid,” in preparation, 2012.

[8] T. Basar and G. 1. Olsder, “Dynamic Noncooperative Game Theory”. Philadelphia, PA, USA: SIAM Series in Classics Applied Mathematics.

[9] W. Tushar, W. Saad, H.V. Poor and D. Smith, “Economics of Electric Vehicle Charging: A Game Theoretic Approach”, *IEEE Trans. Smart Grid*, to appear, 2012.

Yunpeng Wang received his bachelor degree from North China Electric Power University, Beijing, in 2009 and the master degree from Stevens Institute of Technology, in 2012, both in electrical engineering. He is now pursuing his Ph.D. degree at the Electrical & Computer Engineering Department, University of Miami. His primary research interests revolve around the smart grid, particularly control and design of micro-grids, with emphasis on game theory applied to power systems.

Walid Saad received his B.E. degree in Electrical Engineering from the Lebanese University, in 2004, his M.E. in Computer and Communications Engineering from the American University of Beirut (AUB) in 2007, and his Ph.D from the University of Oslo in 2010. Currently, he is an Assistant Professor at the Electrical and Computer Engineering Department at the University of Miami. Prior to joining UM, he has held several research positions at institutions such as Princeton University and the University of Illinois at Urbana-Champaign.

His research interests include wireless networks, game theory, wireless security, and smart grids. He has over 60 international conference and journal publications in these areas. He was the author/co-author of the papers that received the Best Paper Award at IEEE/ICST WiOpt, in June 2009, at the 5th International Conference on Internet Monitoring and Protection (ICIMP) in May 2010, and IEEE WCNC in 2012.

CALL FOR PAPERS

IEEE EUROCON 2013 - International Conference on Computer as a Tool

IEEE EUROCON Special Session: From QoS to QoE and Back

Purpose of this special session:

In recent years, there has been a marked increase in the use of multimedia applications over the Internet. Nowadays video streaming accounts for a large fraction of all bandwidth used, and Internet Telephony and IP-based teleconferencing systems have become part of our daily lives.

The way users perceive the quality of these services is critical both for the users themselves and for the content and service providers. In turn, the quality of these services depends strongly on the underlying transport network and its performance. It is therefore very important to understand how the network-level QoS (Quality of Service) is related to the services' Perceptual Quality and QoE (Quality of Experience), and conversely, how can the latter be exploited in order to improve the performance of the networks and services that run on them.

This Special Session will focus on the relationship between QoS, Perceptual Quality and QoE in general. We welcome submissions on the following topics:

- QoS to QoE mappings
- QoE-driven Network Management
- Cross-layer optimization
- Network QoS and QoE modeling
- Integration of QoS and QoE models
- Standardization issues

Author instructions:

All submissions to this Special Session should follow the regular submission procedure and will be subjected to peer review coordinated by the Special Session Chairs and TPC Chairs. The paper format is the same as for the general conference. The papers must be submitted through the ConfTool, under the respective Special Session track. See paper format for more details.

Important dates:

- Manuscript submission: **February 1st, 2013**
- Notification of acceptance: **March 15th, 2013**
- Camera ready papers: **April 15th, 2013**

Organizers:

- Martin Varela (VTT Technical Research Centre of Finland, Finland) (Martin.Varela@vtt.fi)
- Periklis Chatzimisios (Alexander TEI of Thessaloniki, Greece) (pchatzimisios@ieee.org)

MMTC OFFICERS

E-LETTER BOARD MEMBERS