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

MMTC Communications – Review

Vol. 9, No. 2, April 2018

IEEE COMMUNICATIONS SOCIETY

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the April 2018 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises four reviews that cover multiple facets of multimedia communication research including wireless multimedia, video streaming networks, 3D structures, and QoE learning and prediction. These reviews are briefly introduced below.

The first paper, published in IEEE Transactions on Wireless Communications and edited by Cong Shen, studies the network utility maximization problem for wireless networks.

The second paper is published in IEEE Transactions on Multimedia and edited by Michael Zink. It presents a comparison of the performance of dynamic adaptive video streaming over TCP/IP and ICN.

The third paper, published in IEEE International Conference on Computer Vision and edited by Jun Zhou, describes an algorithm for recovery of dense 3D structure from videos.

The fourth paper is published in IEEE International Conference on Multimedia and Expo and edited by Lifeng Sun. Motivated by the success of deep neural networks in feature representation and learning, the authors of this paper develop an end-to-end and unified deep learning-based framework (termed as DeepQoE) for QoE learning and prediction.

All the authors, nominators, reviewers, editors, and others who contribute to the release of this issue deserve appreciation with thanks.

IEEE ComSoc MMTC Communications – Review Directors

Pradeep K. Atrey
State University of New York at Albany, USA
Email: patrey@albany.edu

Qing Yang
University of North Texas, USA
Email: qing.yang@unt.edu

Wei Wang
San Diego State University, USA
Email: wwang@mail.sdsu.edu

Jun Wu
Tongji University, China
Email: wujun@tongji.edu.cn

Utility Maximization for Real-Time Wireless Networks

A short review for "Joint rate control and scheduling for real-time wireless networks"
Edited by Cong Shen

S. Zuo, I-H. Hou, T. Liu, A. Swami, and P. Basu, "Joint Rate Control and Scheduling for Real-Time Wireless Networks," IEEE Transactions on Wireless Communications, vol. 16, no. 7, Jul, 2017, pp. 4562 – 4570.

Real-time applications for wireless networks, including multimedia streaming, augmented reality/virtual reality (AR/VR), and Internet of Things (IoT), are expected to dominate future network traffic demands. These applications require stringent performance guarantees, including per-packet deadlines and per-flow delivery ratio that cannot be fully captured by traditional quality of service (QoS) metrics. As a result, directly applying standard congestion control and packet scheduling policies for real-time applications may lead to poor network-wide performance as well as violations to the performance requirements of individual applications.

This paper studied the network utility maximization (NUM) problem for wireless networks, where an access point (AP) serves multiple *real-time* wireless flows. It extended the analytical model in a previous study [1], which jointly considered the per-packet deadline, the per-flow delivery ratio requirement, and the unreliable and heterogeneous wireless channels. This paper further considered that each flow can dynamically adjust its traffic load by, for example, adjusting the video resolution in the application of multimedia streaming. Assuming that each flow obtains some utility based on its traffic load, this paper aimed to find tractable policies that maximize the network-wide total utility while respecting the per-packet deadline and the per-flow delivery ratio requirement of each individual flow.

The paper first formulated the problem of maximizing network-wide total utility as a linear programming (LP) problem. Unfortunately, the LP problem involves *exponentially* many constraints, i.e., one set of constraints for each subset of flows due to the per-packet deadline [1]. As a result, standard techniques for solving LP problems are intractable for this NUM problem in real-time wireless networks.

Surprisingly, even though the LP problem involves exponentially many constraints, the paper demonstrated that there exist tractable policies that not only can achieve near-optimal performance, but can also be implemented in a *fully distributed* fashion. This surprising result was derived by leveraging two important theoretical insights: First, the paper proved that the NUM problem in real-time wireless networks is equivalent to a *submodular minimization problem*, which can be exactly solved by a centralized polynomial-time algorithm [2]. Second, the paper utilized the *Lyapunov function* to derive the distributed policies. It showed that the Lyapunov function nicely decomposes the original NUM problem into multiple sub-problems, one for each of the flow and the AP.

Based on these theoretical insights, the paper proposed two tractable policies. In both policies, the AP is only in charge of scheduling packets, and each flow determines its own traffic load based on its utility function and its observation of the network congestion. The first policy is a fully distributed policy that requires virtually no coordination among the AP and the flows. The second policy aims to achieve a faster convergence speed by requiring the AP to periodically broadcast a single value indicating the overall network congestion. The paper proved that both policies achieve near-optimal performance in terms of the network-wide total utility while satisfying all per-packet deadline and per-flow delivery ratio requirements.

The analytical results of the paper were further evaluated by extensive ns-2 simulations. The paper compared the proposed policies against state-of-the-art packet scheduling and congestion control algorithms under a variety of settings. Simulation results demonstrated that the two proposed policies can significantly outperform existing algorithms. The simulation results further highlighted that the performance requirements of emerging real-time applications cannot be properly addressed by algorithms aiming for traditional network utility maximization problems.

References

[1] I. H. Hou, V. Borkar and P. R. Kumar, "A Theory of QoS for Wireless," *IEEE INFOCOM 2009*, Rio de Janeiro, 2009, pp. 486-494.
[2] Satoru Iwata and James B. Orlin, "A Simple Combinatorial Algorithm for Submodular Function Minimization," *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1230-1237, 2009.

Cong Shen received his B.S. and M.S. degrees, in 2002 and 2004 respectively, from the Department of Electronic Engineering, Tsinghua University, China. He obtained the Ph.D. degree from the Electrical Engineering Department, UCLA, in 2009. From 2009 to 2014, He worked for Qualcomm Research in San Diego, CA. In 2015, he returned to academia and joined University of Science and Technology of China (USTC) as the 100 Talents Program Professor in the School of Information Science and Technology. His general research interests are in the area of communication theory, wireless networks, and information theory. He is a Senior Member of IEEE and serves as an editor for the IEEE Transactions on Wireless Communications.

Comparison of ICN and TCP/IP for Video Streaming

A short review for "Dynamic Adaptive Video Streaming:

Towards a systematic comparison of ICN and TCP/IP"

Edited by Michael Zink

J. Samain and G. Carofiglio and L. Muscariello and M. Papalini and M. Sardara and M. Tortelli and D. Rossi, "Dynamic Adaptive Video Streaming: Towards a Systematic Comparison of ICN and TCP/IP," IEEE Transactions on Multimedia, vol. 19, vo. 10, Oct. 2017.

This article tackles a timely and important topic by comparing the performance of dynamic adaptive Video streaming over TCP/IP and ICN [1]. With "Over-the-Top" video streaming being the Internet's killer application of today, it is important to focus on efficient and well performing approaches for the distribution and streaming of videos. Very initial streaming approaches where based on UDP, since the protocol has better support for real time applications. The introduction and follow-on success of Dynamic Adaptive Streaming over HTTP (DASH) [2] lead to the fact that virtually all video streaming in today's Internet uses TCP as transport protocol. Information Centric Networking (ICN) introduces a new approach by shifting from the host oriented paradigm of IP to a content oriented system. Beyond research testbeds ICN is not deployed in the Internet but might see some adoption in the not so distant future. Obviously, such adoption will only happen if it is shown that ICN provides significant benefits compared to TCP/IP.

This article tries to shed more light on the potential performance gains of ICN over TCP/IP in the case of video streaming. The authors compare present results from an experimental evaluation that compares several state-of-the-art DASH controllers (PANDA, AdapTech, and BOLA) on top of an ICN versus a TCP/IP protocol stack. This evaluation is performed on the basis of a set of tools created by the authors and made available as open source software. In addition, the authors used a variety of scenarios to perform a comprehensive evaluation. These scenarios include several videos (up to 4k), network conditions (DASH profiles, emulated WiFi and LTE profiles), and different ICN approaches (basic NDN [3], loss detection, load balancing). The outcome of the evaluation demonstrates the benefits of using ICN for video streaming. In addition, potential pitfalls are also identified.

The article starts with a problem statement that includes a brief description of ICN and its most prominent implementation Named Data Networking (NDN), the introduction of NDN's pull-based transport, and a description of NDN's in-network control. The article focuses on the most beneficial of these building blocks with respect to adaptive streaming. This is followed by a background section that gives a comprehensive overview on related work. The authors focus in particular on the state-of-the-art of Dynamic Adaptive Streaming (DAS) algorithms, which results in a survey that allows readers to easily identify the benefits and drawbacks of the individual DAS approaches in the case of TCP/IP and ICN. The methodology for performing the comparison between TCP/IP and ICN is described in much detail and starts with a description of the video streaming architecture that encompasses the following components: client and server, congestion control, bandwidth estimation, in-network loss recovery, and multi-path support. In addition, the different scenarios used for the evaluation are described. This includes a description of the video content and the network scenarios. As a prequel to the actual evaluation, the authors perform a "calibration" of the DAS algorithms for TCP/IP and ICP/NDN stacks. The goal of this calibration is to identify which QoE parameters for each individual DAS algorithm are most significantly impacted by bandwidth and delay variations. Initial results show that, in a simple scenario where client and server are directly connected over a wired link, the behavior of DAS algorithms running on top of ICP/NDN is almost similar to the same algorithms running on top of TCP/IP. This is not surprising since both ICP [4] and TCP are similar congestion control protocols. While there are no differences between different transport protocol scenarios, the usage of a particular DAS algorithm has an impact on the obtained results. Using a set of key performance indicators (KPI) demonstrates the strength and weaknesses of the

DAS algorithms used in BOLA, PANDA, and AdapTech.

The results of the evaluation show the benefits of ICN features like in-network loss recovery, granularities of the bandwidth estimation, and in-network load balancing.

For the first feature, the results show that intelligent use of ICN in-network loss recovery results in almost similar performance if compared to an TCP/IP-based approach.

Second, it is shown that the fine-grained bandwidth estimation of NDN leads to more bandwidth exploration than in the TCP case. In the case of NDN each chunk represents a distinct sample, while in the case of TCP a complete video segment download constitutes a sample. This result demonstrates that the instantaneous available bandwidth is better observed with fine-grained measurements. Unfortunately, this improvement in bandwidth estimation is not necessarily reflected in an improvement in QoE for each of the DASH algorithms investigated. It should be pointed out that recent approaches [5, 6], which perform fine-grained bandwidth measurements for TCP-based DASH players show similar performance.

Finally, the performance of load balancing in the case of multi-homed wireless devices is evaluated. Here, the authors apply a probabilistic load balancing scheme [7], which schedules new Interest requests with a probability that is inversely proportional to the pending interests of that face. This approach based on interests is compared to one that is based on segment-level requests. Results from an ns-3 based evaluation show that the Interest-based approach results in use of the aggregate bandwidth. Again, the resulting QoE is highly dependent on the DASH algorithm.

The evaluation is concluded by a qualitative and quantitative summary of the different approaches mentioned above. This summary clearly shows that a careful selection of settings for DAS over ICP/NDN is required to achieve significant gains in QoE compared to a DAS over TCP/IP approach.

In summary, this paper makes significant contributions in the area of Dynamic Adaptive Streaming over Information Centric Networks. The results presented in the evaluation section clearly show the benefits and drawbacks of such an approach. In addition, this work could serve as a basis for the investigation of more realistic scenarios in terms of network topology and content diversity. For example, how do the TCP/IP and ICP/NDN approaches compare in a more realistic video streaming system? Continuing research in this area could build on the results and tools presented in this paper. This is supported by the fact that the authors made the tools and instructions to reproduce their results publicly available.[1]

References

[1] Van Jacobson, Diana K. Smetters, James D. Thornton, Michael F. Plass, Nicholas H. Briggs, and Rebecca L. Braynard. 2009. Networking named content. In Proceedings of the 5th international conference on Emerging networking experiments and technologies (CoNEXT '09). ACM, New York, NY, USA, 1-12. DOI=http://dx.doi.org/10.1145/1658939.1658941

[2] I. Sodagar, "The MPEG-DASH Standard for Multimedia Streaming Over the Internet," in *IEEE MultiMedia*, vol. 18, no. 4, pp. 62-67, April 2011. doi: 10.1109/MMUL.2011.71

[3] Lixia Zhang, Alexander Afanasyev, Jeffrey Burke, Van Jacobson, kc claffy, Patrick Crowley, Christos Papadopoulos, Lan Wang, and Beichuan Zhang. 2014. Named data networking. SIGCOMM Comput. Commun. Rev. 44, 3 (July 2014), 66-73.

[4] G. Carofiglio, M. Gallo, L. Muscariello, M. Papalini and Sen Wang, "Optimal multipath congestion control and request forwarding in Information-Centric Networks," *2013 21st IEEE International Conference on Network Protocols (ICNP)*, Goettingen, 2013, pp. 1-10. doi: 10.1109/ICNP.2013.6733576

[5] K. Spiteri, R. Urgaonkar and R. K. Sitaraman, "BOLA: Near-optimal bitrate adaptation for online videos," *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, CA, 2016, pp. 1-9. doi: 10.1109/INFOCOM.2016.7524428

[6] Cong Wang, Divyashri Bhat, Amr Rizk, and Michael Zink. 2017. Design and Analysis of QoE-Aware Quality Adaptation for DASH: A Spectrum-Based Approach. ACM Trans. Multimedia Comput. Commun. Appl. 13, 3s, Article 45 (July 2017), 24 pages. DOI: https://doi.org/10.1145/3092839

[7] D. Bhat, C. Wang, A. Rizk and M. Zink, "A load balancing approach for adaptive bitrate streaming

[1] https://newnet.telecom-paristech.fr/index.php/icn-das/

in Information Centric networks," 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Turin, 2015, pp. 1-6.

Michael Zink is currently an Associate Professor in the Electrical and Computer Engineering Department at the University of Massachusetts Amherst. His research interests are in cyber-physical systems, multimedia distribution, and Future Internet Architectures. In the area of multimedia streaming his work has focused on Available Bitrate Streaming and Quality of Experience. He is Co-Director of the NSF Engineering Research Center for Collaborative Adaptive Sensing of the Atmosphere, where his research focuses on closed-loop sensing system for severe weather detection and warning. As PI for several GENI projects he contributed to the creation of one of an “at-scale” Future Internet Testbed.

Prof. Zink is a Senior Member of the IEEE, a recipient of an NSF CAREER Award, and received the DASH-IF Excellence in DASH Award for his work on quality adaptation for DASH. He received his M.S. and Ph.D. in Electrical Engineering from Darmstadt University of Technology.

An Algorithm for Recovery of Dense 3D Structure from Videos

A short review for "Monocular Dense 3D Reconstruction of a Complex Dynamic Scene from Two Perspective Frames"

Edited by Jun Zhou

Suryansh Kumar, Yuchao Dai, and Hongdong Li, " Monocular Dense 3D Reconstruction of a Complex Dynamic Scene from Two Perspective Frames," The IEEE International Conference on Computer Vision, pages 4649-4657, 2017.

Reconstructing a dense 3D model from an unconstrained scenario is a very challenging task. This applies to the scenarios when two images are captured by an arbitrary-moving camera on complex, dynamic and non-rigid scenes, for example, for motion analysis or robotic vision. In these applications, sophisticated dense reconstruction algorithms are required in order to deal with a combination of rigid and non-rigid motions, for which classic methods of rigid structure-from-motion are not applicable [1].

A feasible solution to this 3D reconstruction problem is treating the scene as a collection of small rigid pieces. The movement of each piece can be handled by a structure-from-motion model. Then the 3D model of the scene can be reconstructed by combining different pieces [2,3]. The difficulty in this solution is that no robust dense object segmentation and correspondence method is available, as they are normally built on a weak assumption that multiple rigid motions are available, which is not the case when objects are deformable.

The paper from Kumar et al. addresses the drawbacks of existing methods on this topic. This paper has two assumptions: 1) The transformation between two frames are locally piecewise-rigid and globally as rigid as possible; 2) The 3D scene to be reconstructed is piece-wise smooth in both frames. Then a globally coherent dense 3D reconstruction of the complex scene can be achieved through three key steps.

The first step divides an image into superpixels, each of which is considered as a rigidly moving object. The super-pixels are associated with several key parameters, e.g., boundary pixels, an anchor point which is the centroid, surface normal and a rigid moving model, all in 3D space. With the 3D superpixels, a K-NN graph can be constructed by considering the Euclidean distance between their anchor points. The correspondences between superpixels in two frames can be produced by an optical flow method [4]. Then the transformation between two superpixels can be estimated using a plan-induced homography [1].

In a larger scale, it is difficult to follow the rigid assumption. Therefore, the authors propose an as-rigid-as-possible constraint as part of the second key step, which enforces smooth motion between local neighbors and preserves local structure between an anchor node and its K nearest neighbors. At the same time, the superpixels shall follow corresponding planar reprojection error in 2D image space, which measures the average dissimilarity of all correspondences. Another important constraint in this step is 3D continuity, i.e., reconstructed surface shall be continuous and smooth. These three conditions are formulated into three energy functions, which are combined so they can be achieved simultaneously.

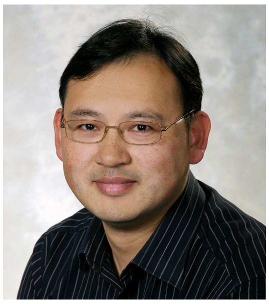
In the last key step, the combined energy function is optimized. Key parameters are solved one after another iteratively until converge. This step also introduces the advantage that the initial dense superpixel correspondences does not need to be perfect, as it can be refined in the optimization step.

In the experiments, four datasets were used to verify the effectiveness of the proposed method. These datasets include scenes extracted from movies, noisy real-world and synthetic scenes, YouTube videos, and some commonly used non-rigid deformation data. Each image in the dataset was segmented into 1000-2000 superpixels using the classic SLIC tool [5]. This method was compared with five baseline monocular dynamic reconstruction methods. The proposed method has shown significant improvement over the alternatives.

The key idea conveyed in this paper is that a complex task like 3D reconstruction in dynamic scenes is not unsolvable. If a divide and conquer strategy is adopted, the complex task can be split into simpler ones that will be easily solved by classic models. Then these divided tasks can be combined with constraints added to achieve the desired goal. Such strategy is much easier and feasible than building a sophisticated model from the beginning.

References:

[1] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge university press, 2003.
[2] R. Ranftl, V. Vineet, Q. Chen, and V. Koltun. Dense monocular depth estimation in complex dynamic scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4058–4066, 2016.
[3] C. Russell, R. Yu, and L. Agapito. Video pop-up: Monocular 3d reconstruction of dynamic scenes. In European Conference on Computer Vision, pages 583–598, 2014.
[4] C. Bailer, B. Taetz, and D. Stricker. Flow fields: Dense correspondence fields for highly accurate large displacement optical flow estimation. In Proceedings of the IEEE International Conference on Computer Vision, pages 4015–4023, 2015
[5] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. S¨usstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(11):2274–2282, 2012.



Jun Zhou received the B.S. degree in computer science and the B.E. degree in international business from Nanjing University of Science and Technology, China, in 1996 and 1998, respectively. He received the M.S. degree in computer science from Concordia University, Canada, in 2002, and the Ph.D. degree in computing science from University of Alberta, Canada, in 2006.

He is now a senior lecturer in the School of Information and Communication Technology in Griffith University. Prior to this appointment, he had been a research fellow in the Australian National University, and a researcher at NICTA, Australia. His research interests are in spectral imaging, pattern recognition, computer vision, and their applications to and environmental informatics and remote sensing.

DeepQoE: When Video QoE Assessment meets Machine Learning

A short review for "DeepQoE: A Unified Framework for Learning to Predict Video QoE"

Edited by Lifeng Sun

Huaizheng Zhang, Han Hu, Guanyu Gao, Yonggang Wen and Kyle Guan, "DeepQoE: A Unified Framework for Learning to Predict Video QoE," 2018 IEEE International Conference on Multimedia and Expo (ICME'18), July 23-27, 2018, Dan Diego, USA.

Recent years have witnessed the explosive growth of multimedia services, which demand for proper metrics to measure the service quality and performance. According to the Cisco Visual Networking Index, the IP video traffic will grow threefold from 2016 to 2021 and take up to 82% of all the consumer Internet traffic [1]. These multimedia applications, including video capturing, compression, transmission, and rendering, rely on QoE (quality of experience) metrics to guide system design, deployment and operation. Therefore, research of QoE plays the fundamental role in multimedia applications.

Although many existing model-based approaches aim to explicitly derive various QoE metrics, they cannot reflect the subjective sentiments. Besides the traditional signal quality metrics, video QoE investigates the effects of contexts and humans to measure the subjective sentiment [2]. It depends on three types of inter-correlated factors, including system factors, context factors and human factors [3]. The prevalent model-based approaches [4-5] commonly choose several factors to create testing samples, recruit participates to score samples under different settings, and derive a close-form QoE model. They have the following drawbacks: 1) these models rely on hand-crafted features and data representation. Therefore, results are confined to specific datasets, and do not generalize well; and 2) subjective tests are time and labor intensive. It is difficult or even impossible to construct a large-scale dataset with representativeness and coverage guarantee.

Motivated by the success of deep neural networks in feature representation and learning, authors develop an end-to-end and unified deep learning-based framework (termed as ***DeepQoE***) for QoE learning and prediction [6]. It consists of three layers, including feature preprocessing, representation learning, and QoE prediction. Feature preprocessing phase utilizes various deep learning-based techniques to extract general features from different datasets or types of data. For example, GloVe, C3D, embedding layer, and dense layer are employed to extract features for text, video, categorical information, and continuous values. Based on the outputs of feature preprocessing phase, representation learning phase concatenates them to a vector, and passes it to two fully connected layers to produce a better representation. QoE prediction phase directly takes the representation as the input for classification or regression tasks.

Decoupled from specific datasets and hand-crafted features, the proposed framework is extensible to a variety of datasets with diverse data modalities and types. This is because the data preprocessing phase is built on pre-trained feature extraction modules. We can utilize the best-fitting modules to preprocess different types of data (e.g., text, video, continuous data, and categorial information) and generate feature representation with customized dimensions. For instance, in consideration of the fast evolving in the natural language processing area, we can replace GloVe with more advanced models if needed. In addition, other information, such as bio-related signals, can also be incorporated into this framework.

In addition to excellent performance gain in QoE prediction, the proposed framework has good generalization and representation learning capabilities. In the small dataset, the performance is comparable to these non-deep-learning based algorithms; while in the large dataset, the performance improvement is significant compared to the best baseline algorithm. Furthermore, using the representation learned from the proposed framework, all the models provide much better performance than using the original features.

Using the open-source tool released by authors [7], this framework can greatly facilitate the QoE related research. Some potential applications include:

- ***Continuous QoE Prediction***: Most of QoE research focuses on predicting QoE of a short video clip (several seconds). Nowadays, video streaming services adopt the DASH standard, which splits a long video into multiple small-duration chunks. We can apply the framework to the QoE modeling of each chunk. Another important issue is how to model the overall QoE of a long video, which depends upon a sequence of inter-correlated QoE of each chunk.

- ***QoE Constrained Video Transmission***: Current video transmission methods (both model based and model free) rely on the explicit forms of QoE metrics. For instance, deep reinforcement learning technique shows a great performance improvement in bitrate selection for DASH based video streaming services, in which the QoE metric is a linear combination of several metrics. We can replace the reward function with the proposed framework.

- ***Emerging Multimedia Applications***: Augmented reality and vidtual reality are touted as the next generation video streaming service. Quite few efforts are devoted for QoE modeling of these services. We can extend the framework to these services.

References

[1]. *Cisco Visual Networking Index: Forecast and Methodology, 2016-2021, White Paper*. Access on 9 April 2018. [Online]. Available: https://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf

[2]. P. Le Callet *et al.*, “Qualinet white paper on definitions of quality of experience (2012),” Eur. Netw. Qual. Exp. Multimedia syst. Services (COST Action IC 1003), v1.2, Lausanne, Siwtzerland, Mar. 2013.

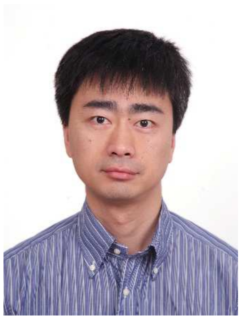
[3]. T. Zhao, Q. Liu, and C.W. Chen, “QoE in video transmission: A user experience-driven strategy,” IEEE Communications Surveys & Tutorials, 19(1), pp.285-302. 2017

[4]. W. Zhang, Y. Wen, Z. Chen and A. Khisti, “QoE-driven cache management for HTTP adaptive bit rate streaming over wireless networks,” IEEE Transactions on Multimedia, 15(6), pp.1431-1445, 2013.

[5]. H. Mao, R. Netravali and M. Alizadeh, “Neural adaptive video streaming with pensieve,” In Proceedings of the Conference of the ACM Special Interest Group on Data Communication (pp. 197-210) 2017. ACM.

[6]. H. Zhang, H. Hu, G. Gao, Y. Wen and K. Guan, “DeepQoE: A Unified Framework for Learning to Predict Video QoE,” 2018 IEEE International Conference on Multimedia and Expo (ICME), July 23-27, 2018, San Diego, USA. [Online: https://arxiv.org/abs/1804.03481]

[7]. http://zhanghuaizheng.info/DeepQoE/



Lifeng Sun received the B.S. and Ph.D. from National University of Defense Technology China, in 1995 and 2000, respectively. He joined the Department of Computer Science and Technology (CST), Tsinghua University (THU), Beijing, China, in 2001. Currently, he is a full Professor in CST of THU. His research interests include networked multimedia, video streaming, 3D/multi-view video coding, multimedia content analysis, multimedia cloud computing and social media. He has authored or coauthored about 100 high quality papers, one chapter of a book. He got the Annual Best Paper Award of IEEE TCSVT (2010), Best Paper Award of ACM Multimedia (2012) and Best Student Paper Award of Multimedia Modeling (2015), IEEE BigMM (2017).

Paper Nomination Policy

Following the direction of MMTC, the Communications – Review platform aims at providing research exchange, which includes examining systems, applications, services and techniques where multiple media are used to deliver results. Multimedia includes, but is not restricted to, voice, video, image, music, data and executable code. The scope covers not only the underlying networking systems, but also visual, gesture, signal and other aspects of communication. Any HIGH QUALITY paper published in Communications Society journals/magazine, MMTC sponsored conferences, IEEE proceedings, or other distinguished journals/conferences within the last two years is eligible for nomination.

Nomination Procedure

Paper nominations have to be emailed to Review Board Directors: Pradeep K. Atrey (patrey@ albany.edu), Qing Yang (qing.yang@unt.edu), Wei Wang (wwang@mail.sdsu.edu), and Jun Wu (wujun@tongji.edu.cn). The nomination should include the complete reference of the paper, author information, a brief supporting statement (maximum one page) highlighting the contribution, the nominator information, and an electronic copy of the paper, when possible.

Review Process

Members of the IEEE MMTC Review Board will review each nominated paper. In order to avoid potential conflict of interest, guest editors external to the Board will review nominated papers co-authored by a Review Board member. The reviewers' names will be kept confidential. If two reviewers agree that the paper is of Review quality, a board editor will be assigned to complete the review (partially based on the nomination supporting document) for publication. The review result will be final (no multiple nomination of the same paper). Nominators external to the board will be acknowledged in the review.

Best Paper Award

Accepted papers in the Communications – Review are eligible for the Best Paper Award competition if they meet the election criteria (set by the MMTC Award Board). For more details, please refer to http://mmc.committees.comsoc.org/.

MMTC examines systems, applications, services and techniques in which two or more media are used in the same session. These media include, but are not restricted to, voice, video, image, music, data, and executable code. The scope of the committee includes conversational, presentational, and transactional applications and the underlying networking systems to support them.