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

MMTC Communications – Review

Vol. 11, No. 1, February 2020

IEEE COMMUNICATIONS SOCIETY

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the February 2020 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises three reviews that cover multiple facets of multimedia communication research including trust assessment in vehicular social network, video-based depth perception, and weakly labeled sound event detection. These reviews are briefly introduced below.

The first paper, published in the IEEE Transactions on Industrial Informatics and edited by Dr. Qin Wang, explores virtual to physical network mappings in elastic virtualized cloud computing environments.

The second paper, published in the IEEE Transactions on Image Processing and edited by Dr. Debashis Sen, describes an adversarial learning approach for person re-identification, i.e., to automatically associate individuals across disjoint camera views in surveillance systems.

The third paper is edited by Dr. Ying Zhang and has been published within the IEEE Transactions on Multimedia. It introduces a visually interpretable image diagnosis network for medical deep learning methods in order to make the output classification results more explainable and thus more trustworthy in clinical settings.

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

IEEE ComSoc MMTC Communications – Review Directors

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

Roger Zimmermann
National University of Singapore, Singapore
Email: rogerz@comp.nus.edu.sg

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

Zhou Su
Shanghai University, China
Email: zhousu@ieee.org

A Dynamic and Quality of Service Driven Mapping Strategy

A short review for "Dynamic Embedding and Quality of Service Driven Adjustment for Cloud Networks"
(Edited by Qin Wang)

H. Cao, S. Wu, G. S. Augla, Q. Wang, L. Yang and H. Zhu, " Dynamic Embedding and Quality of Service Driven Adjustment for Cloud Networks," early access in IEEE Transactions on Industrial Informatics, vol. 16, no. 2, August 2019.

In cloud computing environment, Internet service providers (SPs) can provide elastic virtualized node and link resources [1]. SPs can outsource their virtualized resources as isolated virtual networks (VNs) to multiple customized end users. This scheme is beneficial to sharing physical resources (e.g. CPU, memory, storage, communication bandwidth) and providing flexible vertical network services. A number of technical issues still stand in the way of cloud computing based virtualization implementation [2]. One key technical issue is how to efficiently embed VNs, having various resource and Quality of Service (QoS) demands, onto the underlying physical network (PN), having finite physical resources. The technical issue is called as virtual network embedding (VNE). Till 2019, multiple mapping algorithms [3][4] have been studied.

However, most of prior mapping algorithms are static, including the reinforcement learning based algorithm [5]. In addition, all requested VNs are usually set to be known in advance. With respect to each VN, its embedding results cannot be updated and adjusted once the VN mapping is executed by a certain selected static mapping algorithm. While in the actual cloud computing environment, each VN request is usually unknown in advance. Each VN is dynamically requested. Therefore, dynamic VN mapping algorithms are essential to be developed. In recent years, the dynamic VN mapping algorithms have been studied. However, none considers guaranteeing the concrete QoS performance [1][3] of the accepted VN after completing its initial VN mapping. That is to say, the issue of dynamically re-optimize the initial VN mapping has not been fully investigated yet in the academic community.

In this paper, the authors jointly investigate the dynamic VN embedding and guarantee the QoS performance of each accepted VN. For instance, upon a VN is requested, the authors firstly adopt their proposed dynamic algorithm to embed the VN, fulfilling all virtual resource demands. Then, the authors will check whether the QoS demand of the VN is guaranteed. If the QoS demand of the VN is not guaranteed, a re-embedding scheme of the dynamic algorithm will be triggered and driven. Certain virtual elements (node and link) will be re-adjusted and re-embedded. The proposed dynamic embedding algorithm guarantees flexible VN assignment and VN QoS performance and improves the physical resource utilization in the long run.

Thus, the authors' major contribution is to propose a dynamic heuristic mapping algorithm in order to process online VN requests. With completing one certain VN initial mapping, the authors will check the VN QoS demand. If the QoS demand of the accepted VN is not guaranteed, the re-embedding scheme of the algorithm will be triggered and driven. The authors are not only focusing on accepting more VNs, but also concentrating on guaranteeing the QoS performance of each accepted VN.

Besides of CPU, node location and communication bandwidth constraints, node memory, storage and QoS demand are implemented as virtual constraints in this paper. While in previous VNE publications [4][5], the number of virtual constraints is not more than three (CPU, node location and communication bandwidth). This is another highlight of this paper, comparing with previous publications.

Another novel performance metric, average virtual element delay, is proposed to evaluate VN QoS performance for the first time in VNE research. Previous VNE publications [3][4][5] have not formulated related metrics for representing the VN QoS performance yet. In ref. [3], multiple preliminary performance metrics are categorized into the QoS metric type: path length, stress level, utilization, throughput, delay and jitter. With respect to the first two performance metrics (path length and stress level), they are usually adopted to measure the amount of consumed physical resources for quantifying VN embedding strictly [4]. Hence, path length and stress level are not suitable to be the metrics for VN QoS performance. With respect to the utilization, it strictly belongs to the resource spending category [3]. With respect to the throughput and jitter, they are usually measured based on the VN prototype implementation that is out of the scope of this paper. The authors aim at proposing dynamic efficient algorithm in this paper. With respect to the delay, it just measures the time a packet needs to travel across a virtual link. However, the authors cannot simply focus on one isolated virtual link delay. Therefore, the authors further define the average virtual element delay (Equation 1 in the paper), representing the VN QoS performance in this paper. The average virtual element delay originates from the preliminary metric delay, enabling to reveal the delay level of certain one VN service. Hence, optimizing average virtual element delay does benefit to fulfilling different time-sensitive services in the next generation network [4].

Extensive simulation results are plotted in order to validate the strength of the dynamic heuristic algorithm. For instance, VN acceptance ratio of the dynamic heuristic algorithm improves at least 13% higher than its dynamic version without adjustment. Other performance metrics (e.g. average virtual element delay) are illustrated to further highlight the dynamic algorithm.

In summary, the proposed dynamic and QoS driven mapping algorithm is able to allocate virtualized resources efficiently, comparing with existing dynamic algorithms. In addition, the proposed algorithm guarantees QoS performance per accepted VN.

References:

[1] R. Jain, S. Paul, "Network virtualization and software defined networking for cloud computing: A survey," IEEE Commun. Mag., vol. 51, no. 11, pp. 24-31, 2013.

[2] A. Fischer, J. Botero, M. Till Beck, H. De Meer and X. Hesselbach, "Virtual network embedding: A survey," IEEE Commun. Surveys Tuts., vol. 15, no. 4, pp. 1888-1906, 4th Quart. 2013.

[3] H. Cao, H. Hu, Z. Qu and L. Yang, "Heuristic solution of virtual network embedding: A survey," China Commun., vol. 15, no. 3, pp. 186-219, Mar. 2018.

[4] E. J. Davalos, B. Baran, "A survey on algorithmic aspects of virtual optical network embedding for cloud networks," IEEE Access, vol. 6, no. 1, pp. 20893-20906, 2018.

[5] H. Yao, X. Chen, M. Li, P. Zhang and L. Wang, "A novel reinforcement learning algorithm for virtual network embedding," Neurocomputing, vol. 284, pp. 1-9, 2018.

Qin Wang, Ph.D, is an assistant professor with the College of Engineering and Computing Sciences at New York Institute of Technology. She received B.S. and Ph.D degrees in Communication Engineering from Nanjing University of Posts and Telecommunications, China, in 2011 and 2016, where she is currently a Postdoctoral Research Fellow. She was a visiting Ph.D. student in the Department of Computer Science, San Diego State University. Her research interests include multimedia pricing, network resource allocation, and Internet of Things. She has published papers in prestigious journals such as IEEE Transactions on Vehicular Technology and IEEE Communications Magazine, in prestigious conferences such as IEEE INFOCOM SDP Workshop.

Adversarial Learning for Video-based Person Re-Identification

A short review of "Few-Shot Deep Adversarial Learning for Video-Based Person Re-Identification"

(Edited by Debashis Sen)

L. Wu, Y. Wang, H. Yin, M. Wang and L. Shao. "Few-Shot Deep Adversarial Learning for Video-based Person Re-identification," IEEE Transactions on Image Processing, vol. 29, pp. 1233-1245, 2020.

An essential task in visual surveillance system is to automatically associate individuals across disjoint camera views, which is known as person re-identification (re-ID). It has gained considerable popularity in video surveillance, multimedia, and security system by its prospect of searching persons of interest from a large amount of video sequences. Videos or image sequences are often largely available from surveillance cameras, which inherently contain more information than independent images. Videos are abundant and rich source of human motion and appearance information. For example, given sequences of images, temporal information related to a person's motion can be captured, which may disambiguate difficult cases that arise in the case of recognising the person in a different camera. However, working on videos creates new challenges such as dealing with video sequences of arbitrary length, and the difficulty of learning effective representations while disentangling nuisance factors caused by visual variations.

Before the paper being reviewed, most approaches for video-based person re-ID were based on supervised learning to optimise a discriminant distance metric under which minimised intra-video and maximised inter-video distances can be achieved [1]–[3]. They typically extracted spatio-temporal features (e.g., HOG3D [4]) on each fragment from which videos were represented as a set of extracted features. To learn data-dependent high-level video features, [5], [6] used Long-Short Term Memory (LSTM) networks to aggregate frame-wise CNN features into video-level representations while mapping into hidden states with temporal dependency. However, videos are very high dimensional entities, and it becomes increasingly difficult to do credit assignment on each frame selection to learn long-range relationship among the frames, unless substantial amount of labeled data is collected or feature engineering (e.g., computing the right kinds of flow features) is performed to keep the dimensionality low. As a matter of fact, realization of such video analysis for person re-ID by collecting massive amount of labeled pairs of video sequences becomes difficult.

One major contribution of the paper being reviewed is to develop a few-shot deep adversarial learning framework to produce effective video representations for video-based person re-ID when only a few labeled training paired videos are available. The proposed model based on variational recurrent neural networks (VRNNs) [7] is able to map video sequences into latent variables which serve as video representations by capturing the temporal dynamics. The approach is to learn the latent variables as representations for cross-view video-based person re-ID when only a few labeled training video pairs are available. The latent variables of the VRNNs that model the representations are designed to capture the temporal dependencies across time.

Another contribution of the paper is to addresses the distribution disparity of learned deep features by performing adversarial training [8] in a way to ensure the view-invariance. Further, the algorithm is based on few-shot learning, and thus it has advantage in its generalization capability to widen its application to large-scale networked cameras. To achieve view-invariance, the authors perform adversarial training in a way to produce the latent representations invariant across camera views. Specifically, the VRNNs contain variational auto-encoders that provide a class of latent variables to capture the input dynamics, all of which conditioned on the previous encoders through the hidden states of RNNs. To promote the learned features' view-invariance, the network is augmented with cross-view verification to update adversarially into the view changes, and it encourages view-invariant features to emerge in the course of the optimization. The proposed approach is generic, as it can be created atop any existing feed-forward architecture that is trainable by back-propagation. Meanwhile, the network can be easily optimized by adding a gradient reversal layer [8] that leaves the input unchanged during forward passing and reverse the gradient by multiplying it with a negative scalar during back-propagation.

The authors finally validate the proposed solution and evaluate its performance on real-world benchmark datasets to show that notable improvement is achieved by the proposed method. In order to evaluate the learned representations, the authors qualitatively and quantitatively analyse the predictions and matching rates in people recognition made by the model.

Acknowledgement

The editor would like to thank the authors for providing a preliminary draft of this review.

References:

[1] T. Wang, S. Gong, X. Zhu, and S. Wang, “Person re-identification by video ranking,” in Proc. ECCV, 2014, pp. 688–703.
[2] J. You, A. Wu, X. Li, and W.-S. Zheng, “Top-push video-based person re-identification,” in Proc. CVPR, Jun. 2016, pp. 1345–1353.
[3] X. Zhu, X.-Y. Jing, F. Wu, and H. Feng, “Video-based person re-identification by simultaneously learning intra-video and inter-video distance metrics,” in Proc. IJCAI, 2016.
[4] A. Klaser, M. Marszalek, and C. Schmid, “A spatio-temporal descriptor based on 3D-gradients,” in Proc. BMVC, 2008, pp. 1–11.
[5] N. McLaughlin, J. M. del Rincon, and P. Miller, “Recurrent convolutional network for video-based person re-identification,” in Proc. CVPR, Jun. 2016, pp. 1325–1334.
[6] Y. Yan, B. Ni, Z. Song, C. Ma, Y. Yan, and X. Yang, “Person re-identification via recurrent feature aggregation,” in Proc. ECCV, 2016, pp. 701–716.
[7] J. Chung, K. Kastner, L. Dinh, K. Goel, A. C. Courville, and Y. Bengio, “A recurrent latent variable model for sequential data,” in Proc. NIPS, 2015, pp. 1–9.
[8] Y. Ganin et al., “Domain-adversarial training of neural networks,” J. Mach. Learn. Res., vol. 17, no. 1, pp. 2030–2096, May 2015.

Debashis Sen is an Assistant Professor in the Department of Electronics and Electrical Communication Engineering and a faculty in the Centre of Excellence in Advanced Manufacturing Technology of Indian Institute of Technology - Kharagpur. He received his Ph.D. in Image Processing from Jadavpur University, Kolkata, India and his M.A.Sc. in Electrical Engineering from Concordia University, Montreal, Canada. He was a postdoctoral researcher at the Multimedia Analysis and Synthesis Laboratory, National University of Singapore and at the Center for Soft Computing Research, Indian Statistical Institute. He currently heads the Vision, Image and Perception research group and the ArtEye Lab in his department, which are funded by multiple agencies of Government of India and prominent industries in India. His current research interests are in Vision, Image and Video Processing, Uncertainty Handling, Eye Movement Analysis, Machine Vision and Deep Learning. He has authored/co-authored more than 50 research articles in high impact journals and conferences. Dr. Sen is on the editorial board of IET Image Processing, Springer's Circuits, Systems and Signal Processing, and IEEE MMTC Communications - Review. He has received a young scientist award from The Institution of Engineers (India), a Qualcomm Innovation Fellowship, an ERCIM Alain Bensoussan Fellowship, a Ministry of Manpower (Singapore) Research Fellowship and a couple of best paper awards from IET.

Deep Learning Interpretability in Medical Diagnosis and Applications

A short review for: "VINet: A Visually Interpretable Image Diagnosis Network"
(Edited by Ying Zhang)

Donghao Gu, Yaowei Li, Feng Jiang, Zhaojing Wen, Shaohui Liu, Wuzhen Shi, Guangming Lu, Changsheng Zhou "VINet: A Visually Interpretable Image Diagnosis Network", IEEE Transactions on Multimedia, Early Access Article, 2020.

Recent advances in deep learning, are significantly changing the computer vision industry. We have seen a wide range of deep learning applications in the medical field as well, which aim to identify, classify, and quantify patterns in medical images. The applications include, but are not limited to the detection of, skin cancer [1], breast cancer [2], lung cancer [3], and cervical cancers [4]. Most of these works formulate the medical diagnosis as a classification problem where a single binary prediction label is outputted. But very few of these works explain how the output is reached. Thus, the trustiness of these decisions is sometimes questioned and this makes many solutions to be hard to adopt in the clinical practice. This is also named as the *black-box* problem for deep learning and is exceedingly complicated to be mathematically formulated [5]. Thus, we have witnessed a lot of recent research attempts to make conclusions obtained from deep learning methods more explainable, so that the doctors can review such deep learning conclusions easily and make diagnosis more reliably.

Among these efforts, some popular methods include semantic segmentations and attention models [7-9]. Although these studies made great achievements in explainable deep learning, the results, when applied to the medical field, are sometimes still ambiguous and unclear, which leave great space for further improvements.

To address such an interpretability concern and make the deep learning results closer to human understanding, this paper introduces a visually interpretable image diagnosis network. The network is named as VINet. As an overview, VINet consists of two major components: an importance-estimation network and a deep-classification network. The importance estimation network predicts and quantifies each pixel's importance for the later classification. VINet is trained in an end to end manner and is evaluated on a medical challenge dataset with promising results.

The key component of VINet is the "importance-estimation network". It generates kind of pixel-wise probability map, indicating how important the visual features at those locations may contribute to the final classification decision. This importance-estimation network can be built on any existing network models as long as the output shares the same size with the input. The only difference is to replace the last layer's activation to a proposed normalized sofctmax. This activation function is found to be more powerful than Sigmoid for a fast converge in the experiments. The generated probability map will be used as a visual explanation.

Once the importance-map is obtained, VINet goes to the classification phase to output a classification label. The decision is made from the image regions that belong to the most valued region in the importance-map. Instead of using a multiplication between the feature map and the importance map, which has been popularly used in many existing work, VINet adds uniform noise to the unwanted regions. Such a processing is named as "irrelevant feature destruction". Furthermore, such processing is operated on multiple scales of the image feature maps. Experiments show that such multi-sale noise destruction is much more effective than a single-scale operation.

From the above two key networks, the final loss of VINet is consisted of two parts, a cross-entropy based classification loss and an importance loss. The classification loss is to obtain more accurate classification results and the importance loss aims to eliminate as many features as possible.

Finally, the proposed work is evaluated on the LUNA16 dataset [6]. This is a challenge dataset used for lung nodules detection. The evaluation is conducted both objectively and visually. For the objective evaluation, three popular visualization methods for CNNs are selected as baseline for comparison. These three methods are CAM [7], VBP [8], and LRP [9], all these three solutions are able to output visual explanation maps for the underlying CNNs. The evaluation shows that VINet can produce satisfactory classification results by very few pixels.

To test the classifier robustness, the authors conduct an experiment by occluding part of the image and seeing how the performance is affected. The test results show that the visual interpretation of VINet has the highest occlusion sensitivity. With masking the least pixels (2.37%), the effect on classification accuracy is the most significant (from 82.57% to 8.11%).

As for the visual evaluation, VINet's explanation map is significantly clearer and more accurate than the other three baseline methods. Concretely, the results look to be much sharper. The probability values tend to distribute more densely on value 0 and 1s. This is very likely to be a benefit from the "irrelevant feature destruction".

The paper also reports ablation study to discuss some possible influence factors. It includes a discussion of using different network structures and various feature destruction strategies.

In summary, this work proposes an end-to-end trainable system for medical image diagnosis. VINet can generate diagnostic visual interpretations and classification decisions at the same time. The method is indicated to be verified in clinical practice. The feedback from the doctors might be of great interest to the field.

References:

[1] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115.

[2] Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K. and Li, S., 2017. Breast cancer multi-classification from histopathological images with structured deep learning model. Scientific reports, 7(1), pp.1-10.

[3] Coudray, N., Ocampo, P.S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A.L., Razavian, N. and Tsirigos, A., 2018. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature medicine, 24(10), pp.1559-1567.

[4] Alyafeai, Z., Ghouti, L., 2020. A fully-automated deep learning pipeline for cervical cancer classification. Expert Systems with Applications 141, 112951.

[5] Razzak, M.I., Naz, S. and Zaib, A., 2018. Deep learning for medical image processing: Overview, challenges and the future. In Classification in BioApps (pp. 323-350). Springer, Cham.

[6] Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B. and van der Gugten, R., 2017. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Medical image analysis, 42, pp.1-13.

[7] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. and Torralba, A., 2016. Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2921-2929).

[8] Bojarski, M., Choromanska, A., Choromanski, K., Firner, B., Jackel, L., Muller, U. and Zieba, K., 2016. Visualbackprop: efficient visualization of cnns. arXiv preprint arXiv:1611.05418.

[9] Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.R. and Samek, W., 2015. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PloS one, 10(7).

Ying Zhang received the B.E. degree in computer science from Northwestern Polytechnical University, Xi'an, China, in 2009, and the Ph.D. degree in computer science from the National University of Singapore, in 2014. She is currently a Senior Research Fellow with NUS. Before that, she was the Head of Data Security Unit and Research Scientist with the department of Cyber Security Intelligence, Institute for Infocomm Research (I2R), A*STAR, Singapore. Her research interests include machine learning, image and video processing, multimedia. Various parts of her work have been published in first-tier venues including ACM MM, TMM, TOMM, TIFS, TKDE, etc.

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: Qing Yang (qing.yang@unt.edu), Roger Zimmermann (rogerz@comp.nus.edu.sg), Wei Wang (wwang@mail.sdsu.edu), and Zhou Su (zhousu@ieee.org). 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 Communications – Review Editorial Board

DIRECTORS

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

Roger Zimmermann
National University of Singapore, Singapore
Email: rogerz@comp.nus.edu.sg

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

Zhou Su
Shanghai University, China
Email: zhousu@ieee.org

EDITORS

Koichi Adachi
Institute of Infocom Research, Singapore

Joonki Paik
Chung-Ang University, Seoul, Korea

Xiaoli Chu
University of Sheffield, UK

Mukesh Saini
Indian Institute of Technology, Ropar, India

Ing. Carl James Debono
University of Malta, Malta

Gwendal Simon
Telecom Bretagne (Institut Mines Telecom), France

Marek Domański
Poznań University of Technology, Poland

Cong Shen
University of Science and Technology, China

Xiaohu Ge
Huazhong University of Science and Technology, China

Alexis Michael Tourapis
Apple Inc., USA

Carsten Griwodz
Simula and University of Oslo, Norway

Qing Wang
New York Institute of Technology, USA

Frank Hartung
FH Aachen University of Applied Sciences, Germany

Rui Wang
Tongji University, China

Pavel Korshunov
EPFL, Switzerland

Jinbo Xiong
Fujian Normal University, China

Ye Liu
Nanjing Agricultural University, China

Michael Zink
University of Massachusetts Amherst, USA

Bruno Macchiavello
University of Brasilia (UnB), Brazil

Zhiyong Zhang
Henan University of Science & Technology, China

Debashis Sen
Indian Institute of Technology, Kharagpur, India

Jun Zhou
Griffith University, Australia

Multimedia Communications Technical Committee Officers

Chair: Honggang Wang, University of Massachusetts Dartmouth, USA
Steering Committee Chair: Sanjeev Mehrotra, Microsoft Research, USA
Vice Chair – America: Pradeep K Atrey, University at Albany, State University of New York, USA
Vice Chair – Asia: Wanqing Li, University of Wollongong, Australia
Vice Chair – Europe: Lingfen Sun, University of Plymouth, UK
Letters & Member Communications: Jun Wu, Tongji University, China
Secretary: Shaoen Wu, Ball State University, USA
Standard Liaison: Guosen Yue, Huawei, USA

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.