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

MMTC Communications – Review

IEEE COMMUNICATIONS SOCIETY

Vol. 11, No. 4, August 2020

TABLE OF CONTENTS

Message from the Review Board Directors

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

This issue comprises three reviews that cover multiple facets of multimedia communication research including reconfigurable intelligent surface-based wireless communications, cooperative video caching, and cross-domain manipulation and classification. These reviews are briefly introduced below.

The first paper, published in IEEE Access and edited by Zhe Xing and Rui Wang, designed a prototype and experiments of antennas used for surface based wireless communications to support high data-rate communications.

The second paper is published in IEEE Transactions on Multimedia and edited by Dr. Jinbo Xiong. The paper proposes a cooperative video caching and delivery in a D2D network where devices assist the video dissemination among each other.

The third paper, published in IEEE Transactions on Image processing and edited by Dr. Debashis Sen. This paper aims at jointly tackling multi-domain and multi-modal image manipulation in a unified framework. A novel network architecture called M2RD is proposed for cross-domain image synthesis and classification.

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

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

Prototype and Experiment of the Intelligent Reflecting Surface

A short review for "Reconfigurable Intelligent Surface-Based Wireless Communications: Antenna Design, Prototyping, and Experimental Results"

(Edited by Zhe Xing and Rui Wang)

L. Dai, et al., " Reconfigurable Intelligent Surface-Based Wireless Communications: Antenna Design, Prototyping, and Experimental Results," IEEE Access, vol. 8, pp. 45913-45923, Mar. 2020.

Researches on the 6th generation (6G) mobile communications have already begun since the partial deployment of the 5th generation (5G) commercial networks in the first half of 2020. Compared with the current 5G, the future 6G mobile communication networks will face more challenges, such as providing Tbps magnitude of data rate, near 100% geographical coverage, intelligent and personalized wireless transmission, full automation, et al [1]. Although various key technologies including the millimeter wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra-dense network (UDN), have been proposed to enable huge mobile data traffic and high-speed data transmission [2], they still exhibit drawbacks of high hardware cost (HWC) and energy consumption (EC). To tackle with these problems, 6G should not only be capable of supporting massive machine type communications (mMTC) and ultra-reliable and low latency communications (uRLLC) as 5G does, but also be able to reduce the HWC, improve the energy efficiency and promote the communication intelligence.

Recently, from the perspective of controlling the wireless propagation environment, a state of the art approach, named intelligent reflecting surface (IRS), has been proposed to establish a novel communication paradigm and been envisioned as one of the candidate 6G technologies [3]. Distinguished from the conventional active relay which consumes power and retransmits the signal, or the general physical wall which merely reflects the electromagnetic wave, the IRS is a planar array composed of a large quantity of low-cost passive reflecting units, each of which can generate an adjustable phase shift on the impinging signal wave, and reflect the adjusted signal to the receiving terminal. As a low-cost passive reconfigurable reflecting apparatus, the IRS is regarded as a promising technique to solve the problems of HWC and EC.

The theoretical characteristics and potential applications of the IRS have been extensively investigated [4]. However, the IRS was mostly considered as a concept to be integrated into the traditional wireless communication system models. A few researches that focused on practically fabricating the IRS and measuring its performance in the real world could be found in [4]-[6], where the varactor-based analog IRS and the IRS with 1-bit elements were developed. Nevertheless, the varactor-based IRS had long response time and unsatisfactory phase accuracy, while the 1-bit IRS suffered from significant phase errors which caused antenna gain reductions. Therefore, motivated by the contributions on the IRS with 2-bit elements which struck a tradeoff between the above two cases [7], in this paper, L. Dai, et al. [8] design a 2-bit IRS based prototype system and evaluate its performance at 2.3 GHz and 28.5 GHz.

First, a square IRS with $16 \times 16 = 256$ 2-bit reflecting elements is fabricated. Each reflecting element contains an upper patch, a ground plane and a slot-loaded plane. The upper patch and the ground plane are responsible for receiving or radiating energy and suppressing the back radiation, respectively, while the slot-loaded plane is responsible for inducing the phase shifts, which are generated by positive intrinsic-negative (PIN) diodes and controlled by an FPGA-based phase shift controller, on the coming signal waves. In the slot-loaded plane, 5 PIN diodes are integrated into 4 sets of slots. Each PIN diode can be turned ON or OFF to create 4 different phase states, corresponding to 4 element configurations, with a phase increment of approximately $90°$. Simulation results demonstrate that the phase states remain stable and the magnitude responses degrade by less than 1.2 dB within the frequency band from 2 GHz to 2.6 GHz.

Subsequently, an IRS-assisted wireless communication prototype system is designed. The prototype system consists of a host, a universal software radio peripheral (USRP) equipment and an IRS with 256 2-bit reflecting elements at the transmitter side, and a host and a USRP at the user side. The transmitter host configures the signal parameters and drives the transmitter USRP to perform encoding, modulation, up-conversion and signal emission. The IRS induces phase shifts on the signals and reflects them to the receiver USRP. The receiver USRP is driven by the receiver host to accomplish signal reception, down-conversion, demodulation and decoding.

Finally, the experiments are carried out to testify the feasibility and efficiency of the fabricated IRS. The measurements are first performed in the compact anechoic chamber at 2.3 GHz, and then in the over-the-air (OTA) test environment with high-definition virtual reality (VR) video streams at both 2.3 GHz and 28.5 GHz. Results show that the IRS can achieve 21.7 dBi and 19.1 dBi antenna gains at 2.3 GHz and 28.5 GHz, respectively. Compared to the conventional phased arrays, the IRS can reduce the power consumption by 58.6% in terms of effective isotropic radiated power when maintaining the similar performance.

In conclusion, as an attractive and promising approach, the IRS has prompted the researchers to make theoretical analyses, optimizations and functional designs on its conceptual model. On the contrary, to the best of our knowledge, the practical fabrication and application of the IRS are still in the exploratory stage and may face many problems in terms of phase accuracy, real-time control and response, deployment, et al. As a result, the attempt to build the IRS prototype is of great significance to facilitate its practical application in the future 6G wireless communication networks.

References:

[1] X. You, *et al*., "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, for review.

[2] M. Agiwal, A. Roy and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," *IEEE Commun. Surv. Tut.*, vol. 18, no. 3, pp. 1617–1655, Third Quarter 2016.

[3] N. Rajatheva, *et al*., "White paper on broadband connectivity in 6G," *6G Research Visions*, no. 10, Jun. 2020.

[4] S. Gong, *et al*., "Towards smart wireless communications via intelligent reflecting surfaces: a contemporary survey," *IEEE Commun. Surv. Tut.*, to be published. DOI: 10.1109/COMST.2020.3004197

[5] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing indoor spectrum sharing capacity using smart reflect-array," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1-6.

[6] M. T. Zhang, *et al*., "Design of novel reconfigurable reflectarrays with single-bit phase resolution for ku-band satellite antenna applications," *IEEE Trans. Antennas Propag*., vol. 64, no. 5, pp. 1634-1641, May 2016.

[7] P. Nayeri, F. Yang, and A. Z. Elsherbeni, *Reflectarray Antennas: Theory, Designs, and Applications*. Hoboken, NJ, USA: Wiley, 2018.

[8] L. Dai, *et al*., "Reconfigurable intelligent surface-based wireless communications: antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45913-45923, Mar. 2020.

Zhe Xing received the B.S. degree in communication engineering from the College of Electrical Engineering, Sichuan University, Chengdu, China, in 2016, and the M. S. degree in communication and information systems from the College of Electronics and Information Engineering, Sichuan University in 2019. He is currently pursuing the Ph.D. degree in information and communication engineering with the College of Electronics and Information Engineering, Tongji University, Shanghai, China. His research interests include intelligent reflecting surface in wireless communication systems, joint communication and localization, and automatic modulation classification in non-cooperative communication environment.

Rui Wang received his Ph.D. degree in 2013 from Shanghai Jiao Tong University, China. From Aug. 2012 to Feb. 2013, he was a visiting Ph.D. student at the Department of Electrical Engineering of University of California, Riverside. From Oct. 2013 to Oct. 2014, he was with the Institute of Network Coding, the Chinese University of Hong Kong as a post- doctoral research associate. From Oct. 2014 to Dec. 2016, he was with the College of Electronics and Information Engineering, Tongji University as an assistant professor, where he is currently an associate professor.

A Cooperative Mechanism for Video Caching and Delivery

A short review for "Cache Less for More: Exploiting Cooperative Video Caching and Delivery in D2D Communications"

Edited by Jinbo Xiong

D. Wu, Q. Liu, H. Wang, Q. Yang and R. Wang, "Cache Less for More: Exploiting Cooperative Video Caching and Delivery in D2D Communications," IEEE Transactions on Multimedia, vol. 21, no. 7, July. 2019.

According to Cisco's latest Visual Networking Index (VNI) report, the global mobile data traffic will increase nearly 9-fold from 2015 to 2020, of which mobile video services will account for nearly 75% [1]. This problem is compounded by the fact that mobile users tend to access/watch videos online, which puts high bandwidth requirements and stringent delay constraints of online video streaming. Therefore, the ever-increasing demand for videos poses a serious challenge to existing cellular network infrastructures, in delivering videos with a good quality of experience (QoE) to users.

In order to cope with this ongoing increasing mobile data demand, device-to-device (D2D) communications, as defined in the 3rd generation partner project (3GPP), was proposed to allow mobile devices in close proximity to transmit data between each other, without base station (BS) forwarding [2],[3]. Therefore, it can effectively reduce the load on BS, dramatically improving the radio spectrum utilization and network throughput. However, applying D2D communication in video streaming is still an open problem [4],[5], and it is further complicated by the mobility and sharing willingness of users, unstable channel conditions, and high QoE requirements. Particularly, the movement of users makes D2D links exhibit a discontinuous feature in both temporal and spatial domains, which may cause channel errors and packet losses in video transmissions. Furthermore, interrupted video transmissions or transmissions with large jitters/delays, would significantly degrade a user's QoE.

In this paper, the author borrowed ideas from a user-centric virtual cell technology, and considering that the deployment of more fixed access points (APs) will cause increased communication costs. Therefore, the author proposes a user-centric D2D communication mechanism by using user devices instead of traditional APs, and each user who subscribes to the video service has a dynamic service set that consists of its neighboring users. Specifically, a service set consisting of several service providers, i.e., regular mobile users, is dynamically configured for a service consumer. As such, requested videos are first pushed from BS(s) to service providers and then transmitted to the service consumer, via D2D communications.

The mechanism allows regular mobile users to cache and transmit videos between each other in a cooperative manner. Furthermore, to achieve a QoE-guaranteed video streaming service in a cellular network, the proposed solution jointly considers users' similarity in accessing videos, users' sharing willingness, mobility, storage capacity, and their QoE requirements.

First, mobile users in this proposal are divided into two groups, namely service providers (SP) and service consumers (SC). When a SC successfully subscribes to a video service, its serving BS will check the SC's video accessing history and generates a cooperative cache list (CCL) that contains the videos that are likely to be accessed by the SC in the future. By comparing the similarity between a SP and the SC, regarding to their video accessing history, the CCL of a SP will be identified. From the CCL, a SP determines which videos need to be cached in its local storage. Based on the amount and quality of videos that a SP can provide; a group of SPs are selected to construct the SC' service set.

After the service set is established, videos are pushed to SPs at off-peak time, if possible, to avoid network congestion caused by concurrent video transmissions during peak time. Considering that the storage capacity of the mobile device is limited, the user cannot cache multiple demand videos. Besides, in order not to affect the other operations of consumer, e.g., limited work space, only SPs receive the video

pushed by BS. Thus, the user can prefetch data to reduce latency when network congestion happened. After the so-called base station to device (B2D) communication, videos will be cached on the SPs. Later on, when the SC and SPs are in close proximity, cached videos are transmitted to the SC via D2D communications, in a "many-to-one" manner, that is, a consumer obtains videos from multiple potential providers. The mode can be regarded as the independent operation of several D2D links. When the requested video is completely delivered to the SC, the BS will update the SC's and SPs' CCLs accordingly.

Numerical results indicate that the proposed communication mechanism is able to not only improve users' experience, but also offload the data traffic on the BS(s).

In summary, the authors explore the possibility of caching videos on individual users in a cellular network, and letting similar users share videos to each other, via a hybrid communication mode, i.e., B2D+D2D. The major instrument of this work is the cooperative cache list that considers users' similarly in accessing videos to determine the videos that need to be cached. As a result, the videos are likely to be accessed by service consumers and those needed to be cached on service providers are identified.

References:

[1] M. Abana, M. Peng, Z. Zhao, and L. Olawoyin, "Coverage and rate analysis in heterogeneous cloud radio access networks with device-to-device Communication," *IEEE Access*, vol. 4, pp. 2357-2370, 2016.

[2] X. Cheng, J. Liu, H. Wang, and C. Wang, "Coordinate live streaming and storage sharing for social media content distribution," *IEEE Trans. Multimedia*, vol. 14, no. 6, pp. 1558–1565, Dec. 2012.

[3] D. Wu, L. Zhou, and Y. Cai, "Social-aware rate-based content sharing mode selection for D2D content sharing scenarios," *IEEE Trans. Multimedia*, vol. 19, no. 11, pp. 2571-2582, Nov. 2017.

[4] C. Li, L. Toni, J. Zou, H. Xiong, and P. Frossard, "QoE-driven mobile edge caching placement for adaptive video streaming," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 965-984, 2018.

[5] X. Ge, L. Pan, Q. Li, "Multipath cooperative communications networks for augmented and virtual reality transmission," *IEEE Trans. Multimedia*, vol. 19, no. 10, pp. 2345-2358, 2017.

Jinbo Xiong, Ph.D, is a Professor and Ph.D. supervisor with the Fujian Provincial Key Laboratory of Network Security and Cryptology and the College of Mathematics and Informatics at Fujian Normal University. He received the Ph.D. degree in Computer System Architecture from Xidian University, China, in 2013. His research interests include connected and autonomous vehicle, secure deep learning, cloud data security, mobile media management, privacy protection, and Internet of Things. He has published more than 100 papers in prestigious journals such as IEEE WCM, IEEE TII, IEEE TCC, IEEE IoT J, IEEE TNSE, FGCS and in major International conferences such as IEEE ICCCN, IEEE TrustCom, IEEE HPCC and IEEE ICPADS. He has applied 15 patents and two monographs in these fields. He is a member of IEEE.

Learning Disentangled Representations for Cross-Domain Image Analysis

A short review for "A Multi-Domain and Multi-Modal Representation Disentangler for Cross-Domain Image Manipulation and Classification"

Edited by Debashis Sen

F.-E. Yang, J.-C. Chang, C.-C. Tsai and Y.-C. F. Wang, "A Multi-Domain and Multi-Modal Representation Disentangler for Cross-Domain Image Manipulation and Classification," IEEE Transactions on Image processing, vol. 29, 2795 – 2807, 2020.

Recently, deep learning has achieved significant progresses in the areas of computer vision and machine learning. While exhibiting promising performances, the success of deep learning models typically relies on large-scale labeled datasets for training the associated networks. In practice, collecting and annotating an enormous amount of training data might not be always feasible. With a range of real-world applications such as autonomous driving, person re-identification, scene understanding and style transfer, performing visual synthesis or recognition tasks with representation across image domains is therefore a challenging yet important task.

As solutions to the above issue, a number of methods utilize domain adaptation techniques and learn mapping functions across source and target domains. These are to derive domain-invariant latent representations for encoding cross-domain input data. While these models exhibit satisfactory results for transfer between two domains, they generally suffer from scalability issues with extension of their models for more than two domains. Recently, to perform domain transfer across multiple domains, StarGAN [1] and UFDN [2] present unified model structures with the aid of domain label to translate input images to multiple style of interests.

In the aforementioned visual analysis tasks, one might not only need to perform translation and adaptation across data domains. Take style transfer for example, it would be necessary to exploit intra-domain variations and introduce multi-modal diversity during image synthesis. That is, a single input image may correspond to diverse possible outputs of the same style (e.g., daytime, summer, etc.). Due to the mode collapse problem encountered in Generative Adversarial Nets (GANs) [3], it might not be easy to produce diverse outputs using GAN-based models. To enforce the diversity of the generated images, DRIT [4] and MUNIT [5] apply representation disentanglement techniques to decompose latent representation into disjoint features, which describe content and style information for diverse image-to-image translation. By manipulating different style features, their models would be capable of synthesizing outputs with sufficiently multi-modal diversity.

In the paper being reviewed, the authors aim at jointly tackling multi-domain and multi-modal image manipulation in a unified framework. The authors propose a novel network architecture of *Multi-domain and Multi-modal Representation Disentangler* (M^2RD) for cross-domain image synthesis and classification, with the ability to manipulate image data with particular attribute of interest while exhibiting sufficient diversity. Without the need to collect pairwise image training data, the M^2RD learns to factorize latent image representation into domain-invariant content feature with associated domain-specific one. While the former embeds content information from input data, the latter shows diversity during image translation across multiple domains.

To be more specific, derivation of domain-invariant feature is achieved by the use of domain confusion objectives, which eliminate the domain differences between the features encoded from distinct domains, while the disjoint part of domain-specific feature is ensured to describe style information via reconstruction guarantee. The authors additionally advance adversarial learning with auxiliary domain classifier to guarantee the quality of synthesized images and satisfy the proper domain information. Through fitting domain-specific features to a Gaussian distribution, M^2RD allows sampling different noise vectors as style features during inference to exhibit multi-modal diversity of corresponding output images. Finally, the content consistency constraint further preserves content information during the generation process. In addition to cross-domain image translation, the authors show

that the derived domain-invariant content features can be applied to the visual classification task of *unsupervised domain adaptation* (UDA), which desires a common feature representation shared by different domains for adaptation purposes.

For evaluation, the authors present the synthesized images using different types of datasets (i.e., digits, faces, summer-to-winter, and photo-to-art) in multiple settings, including multi-domain and multi-modal image manipulation. The results of multi-modal image translation show that the model presented by the authors is capable of generating images with sufficient diversity from the given input image, which further proves its excellent ability to perform one-to-many image translation. On the other hand, the model is also successfully applied to translate input images across multiple domains with intra-domain diversity. Apart from the qualitative results, the authors also show t-SNE visualization and classification accuracy results in their experiments, showing that the feature representation learned by the model successfully exhibits domain-invariant characteristics.

In summary, the novel M^{2}RD framework of the paper being reviewed is unique and innovative. The novel network architecture tackles cross-domain image manipulation and classification across multiple domains. It exhibits satisfactory multi-modal diversity by factorizing and modeling the feature representations of interest. As confirmed by the experimental results, M^{2}RD produces promising visual outputs with multi-modal diversity across multiple domains, while the derived feature representations are successfully applied for solving unsupervised domain adaptation tasks.

Acknowledgement:

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

References:

[1] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Chao, "StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation," *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp.8789–8797

[2] A. H. Liu, Y.-C. Liu, Y.-Y. Yeh, and Y.-C. F. Wang, "A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation," in *Advances in neural information processing systems*, 2018, pp. 2590–2599.

[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[4] H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. K. Singh and M.-H. Yang, "Diverse Image-to-Image Translation via Disentangled Representations," *in Proceedings of European conference on computer vision*, 2018, pp. 35–51.

[5] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal Unsupervised Image-to-Image Translation," *in Proceedings of European conference on computer vision*, 2018, pp. 172–189.

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

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

Multimedia Communications Technical Committee Officers

Chair: Honggang Wang, University of Massachusetts Dartmouth, USA
Steering Committee Chair: Sanjeev Mehrotra, Microsoft Research, US
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.