

**MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE  
IEEE COMMUNICATIONS SOCIETY**

<http://mmc.committees.comsoc.org/>

**MMTC Communications – Review**



IEEE COMMUNICATIONS SOCIETY

---

**Vol. 12, No. 2, April 2021**

**TABLE OF CONTENTS**

<b>Message from the Review Board Directors</b>	2
<b>SPA-GAN: Spatial Attention GAN for Image-to-Image Translation</b>	3
A short review for “SPA-GAN: Spatial Attention GAN for Image-to-Image Translation” (Edited by Guitao Cao)	
<b>An Auction Mechanism for Video Caching in Heterogeneous Networks</b>	5
A short review for “Double Auction Mechanism Design for Video Caching in Heterogeneous Ultra-Dense Networks” (Edited by Shengjie Xu)	
<b>A Creative and Fundamental Study on the Coding Technique for Wireless Caching Systems</b>	7
A short review for “On the Fundamental Limits of Coded Caching Systems with Restricted Demand Types” (Edited by Cong Shen)	
<b>Federated Learning Review in Cooperative Systems</b>	9
A short review for “Opportunities of Federated Learning in Connected, Cooperative, and Automated Industrial Systems” (Edited by Mohammad H Hasan)	
<b>A Proof of Concept of Multimedia Streaming and its Advantages in a Real-world Learning Environment</b>	11
A short review for “Improving Student Learning Satisfaction by Using an Innovative DASH-based Multiple Sensorial Media Delivery Solution” (Edited by Lucile Sassatelli)	

## Message from the Review Board Directors

Welcome to the April 2021 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises five reviews that cover multiple facets of multimedia communication research including Image-to-image translation, video caching, wireless caching system, connected, cooperative and automated industrial system, and mulsemmedia streaming. These reviews are briefly introduced below.

The first paper, published in IEEE Transactions on Multimedia and edited by Dr. Guitao Cao, proposes a spatial attention GAN model, which computes the spatial attention in discriminator and use it to help the generative network attend to the discriminative regions and produce more realistic images.

The second paper is published in IEEE Transactions on Wireless Communications and edited by Dr. Shengjie Xu. It designs a double auction-based caching mechanism to ensure the efficient operation of the market by maximizing the social welfare.

The third paper, published in IEEE Transactions on Communications and edited by Dr. Cong Shen, proposes a novel caching and delivery coding scheme for caching systems where there's a restriction on the demand of users.

The fourth paper, published in IEEE Communications Magazine and edited by Dr. Mohammad H Hasan, explores emerging opportunities of federated learning for the next-generation networked industrial systems.

The fifth paper, published in IEEE Transactions on Multimedia and edited by Dr. Lucile Sassatelli, builds a mulsemmedia testbed made of a programmable fan and sent diffuser, a haptic mouse and a computer where the students watched the educational STEM video.

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

IEEE ComSoc MMTC Communications – Review  
Directors

Zhisheng Yan  
Georgia State University, USA  
Email: zyan@gsu.edu

Yao Liu  
Binghamton University, USA  
Email: yaoliu@binghamton.edu

Wenming Cao  
Shenzhen University, China  
Email: wmcao@szu.edu.cn

Phoenix Fang  
California Polytechnic State University, USA  
Email: dofang@calpoly.edu

## SPA-GAN: Spatial Attention GAN for Image-to-Image Translation

*A short review for “SPA-GAN: Spatial Attention GAN for Image-to-Image Translation”*

Edited by *Guitao Cao*

*H. Emami, M. M. Aliabadi, M. Dong and R. B. Chinnam, "SPA-GAN: Spatial Attention GAN for Image-to-Image Translation," in IEEE Transactions on Multimedia, vol. 23, pp. 391-401, 2021.*

Image-to-image translation is to learn a mapping between images from a source domain and images from a target domain. As we can imagine, an excellent image-to-image translation system could have an almost unlimited number of applications, including image colorization, generating semantic labels from images [1], image super resolution [2], [3], and domain adaptation [4].

Image-to-image translation aims to learn the mappings between different image domains. How to represent these mappings to generate the desirable results is explicitly related to the generative models, which can model the distribution of the target domain by producing convincing “fake” data, namely, the translated images. In general, an image-to-image translation method needs intrinsic source content preserved and the extrinsic target style transferred. Therefore, it is important to identify the regions of interest. Once the discriminative regions are confused with the noisy regions, it will lead to image distortion or a substantial visual gap. In an unsupervised setting with a large but unpaired number of images between two domains, one must leverage additional information to force the model to focus on the regions of the image that are subject to transfer.

This paper introduces a novel attention mechanism for image-to-image transformations of GANs, called spatial attention GAN (SPA-GAN), which computes attention weights using spatial information of channel dimensions in the discriminator layer. An additional feature map loss contains generator and decoder feature map loss, calculated by the  $L_1$  norm. In this way, the weighted feature maps are fed into the generator network, allowing the generator network to perform better at transferring extrinsic target styles and maintaining domain-specific features.

And SPA-GAN, different from previous methods which require additional attention networks or supervision (e.g., segmentation labels) during training, is a lightweight model with a low computational burden on the GPU.

In this paper, the authors’ major contribution is to propose a spatial attention GAN model, which computes the spatial attention in discriminator and use it to help the generative network attend to the discriminative regions and produce more realistic images. Inspired by the CycleGAN [5] network structure, SPA-GAN outputs the attention graph from the discriminator and inputs it into the generator. Based on the proposed attention mechanism, SPA-GAN modifies the cycle consistency loss to help the generator focus on the most discriminative regions under a stronger cycle-consistency constraint, and adds the feature map loss to penalize the differences between feature maps to preserve more low-level features.

SPA-GAN introduces the attention mechanism directly to the GAN [6] architecture and propose a novel spatial feature map loss for image-to-image translation. In SPA-GAN, the discriminator network classifies the input images into real or fake, while outputs the attention maps defined as the spatial maps [7]. In unsupervised setting, SPA-GAN proposes an additional feature map loss to preserve domain specific features during translation. In SPA-GAN’s generative network, authors constrain the feature maps obtained in the first layer of the decoder [8] to be matched with the identified regions of interest from both real and generated images so that the generated images are more realistic.

Extensive experiments demonstrate the improved performance of the proposed spatial attention GAN (SPA-GAN). The results show that this

model significantly outperforms other state-of-the-art image-to-image translation methods on various benchmark datasets, while requiring a small overhead in terms of additional parameters. In summary, the proposed spatial attention GAN (SPA-GAN) for image-to-image translation is in unsupervised settings. In this work, SPA-GAN, which is based on the CycleGAN network structure, outputs the attention map from the discriminator and inputs it into the generator to assist the generator in focusing on more discriminated regions in the image. SPA-GAN modifies the cyclic consistency loss with additional feature map loss (with the first layer of decoder output operation), which has the lowest KID and the highest classification accuracy. SPA-GAN is a lightweight model and achieves superior performance over current state-of-the-arts both qualitatively and quantitatively.

### References:

- [1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 1125–1134.
- [2] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 4681–4690.
- [3] L. Chen, L. Wu, Z. Hu, and M. Wang, "Quality-aware unpaired image-to-image translation," IEEE Trans. Multimedia, vol. 21, no. 10, pp. 2664–2674, Oct. 2019.
- [4] Z. Murez, S. Kolouri, D. Kriegman, R. Ramamoorthi, and K. Kim, "Image to image translation for domain adaptation," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2018, pp. 4500–4509.
- [5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proc. IEEE Int. Conf. Comput. Vision, 2017, pp. 2223–2232.

- [6] Goodfellow, I. J. , et al. "Generative Adversarial Networks." Advances in Neural Information Processing Systems 3(2014):2672-2680.
- [7] S. Zagoruyko and N. Komodakis "Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer," in Proc. Int. Conf. Learn. Representations, 2017.
- [8] M.-Y. Liu and O. Tuzel, "Coupled generative adversarial networks," in Proc. Adv. Neural Inf. Process. Syst., 2016, pp. 469–477.



**Guitao Cao** obtained her Ph.D. in 2006 from Shanghai Jiao Tong University with a focus on pattern recognition. She is currently a professor of Software Engineering Institute, East China Normal University (ECNU), Shanghai, China. ECNU is the top tier university in China with a high rank (Level A) in Software Engineering in China. She was also a visiting researcher with University of Missouri at Columbia. She has published decades of peer reviewed papers in top venues including IEEE Transactions on Cybernetics, IEEE Transactions on Multimedia, and IEEE Transactions on Biomedical Engineering. Prof. Cao is also the Principal Investigator for many research funding with major sponsors including the National Science Foundation of China, Ministry of Industry and Information Technology of the People's Republic of China, and Science Foundation of Shanghai. Her research interests include pattern recognition, image processing and machine learning.

## An Auction Mechanism for Video Caching in Heterogeneous Networks

*A short review for “Double Auction Mechanism Design for Video Caching in Heterogeneous Ultra-Dense Networks”*

Edited by Shengjie Xu

*J. Du, C. Jiang, E. Gelenbe, H. Zhang, Y. Ren and T. Q. S. Quek, "Double Auction Mechanism Design for Video Caching in Heterogeneous Ultra-Dense Networks," in IEEE Transactions on Wireless Communications, vol. 18, no. 3, pp. 1669-1683, March 2019.*

Mobile data traffic is experiencing a dramatic increasing over cellular networks. There is evidence that content distribution services, such as video on demand, catch-up TV, internet video streaming, etc., have become premier drivers of the exponential traffic growth [1]. A relatively small number of popular video files provided by a certain part of video service providers account for the most of data traffic, and are requested frequently by mobile users (MUs). To deal with this fact of highly redundant video demands from MUs, caching techniques, e.g., storing video files in MUs' devices or potential helper nodes disseminated in the network, have been developed to avoid the high-throughput backhaul to the core network, which is too costly and constitutes a major bottleneck [2].

Recently, the small-cell based architecture has dominated in ultra-dense heterogeneous networks (HetNets). In HetNets, mobile network operates (MNOs) deploy multiple small base stations (SBSs) which work in conjunction with micro base stations (MBSs). For the MNO, the cost for long distant transmission can be saved by this architecture. On the other hand, MUs receive their requested data through low power consumption, low-latency and better-quality communications [3]. Based on this architecture, video caching relying on SBSs constitutes a feasible and low-cost solution to further cope with the increasing video data traffic over backhaul channels with assistance of the small-cell based architecture.

Small-cell based caching mechanisms can help offload data traffic from the MBSs and bring contents closer to the MUs, which will reduce the power consumption, shorten the transmission latency and offloading delay. Due to its significant performance on releasing the increasing mobile

data traffic, video caching has received considerable attention in the wireless communications, and many researches have focused on effect and efficient video caching mechanism design.

In this work, a small-cell based video caching system in ultra-dense HetNets, in which the MNOs operate a set of SBSs and leases SBS storage to multiple video service providers (VSPs) to placing their video files. Based on different VSP preferences, VSP utility functions, the MNO cost function and a social welfare maximization problem are formulated. Then, in order to elicit the hidden information among VSPs and the MNO, i.e., VSP utility functions and the MNO cost function, a double auction model is introduced to solve the caching problem.

Moreover, an alternative optimization problem is formulated where it has the same optimal solution as the social welfare optimization problem. By solving this problem, and applying the designed allocation schemes and pricing rules, the maximum social welfare can be achieved although there exists the hidden information. The detailed proof of convergence and economic properties of the designed double auction-based caching mechanism are validated by simulation results.

Firstly, this work introduces how to implement the video caching in the small-cell based network. Then it analyzes the economic utility that can be obtained by VSPs from receiving the caching service, and the cost of the MNO when providing the caching service to VSPs. The caching procedure in the small-cell based caching system is introduced first, by following three steps. The first step focuses on SBS assignment, where each VSP rents a certain fraction of the SBSs operated

by the MNO for placing its video files. The second step focuses on video file placing, where each VSP accesses and places the most popular video files at its assigned SBSs to achieve an efficient caching. The third step focuses on MU video requests, where an MU sends a request of video. This request is first sent to this MU's nearest SBS which caches this video file. If there exists an SBS that has cached this requested video and its coverage can cover this MU, then this MU will download the video directly from this SBS.

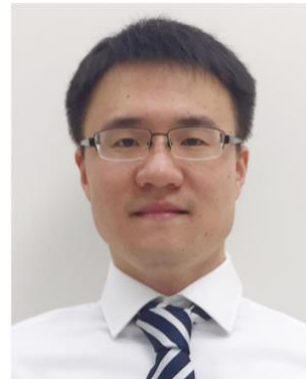
Secondly, a double auction mechanism design for small-cell based caching system. Based on the model and its analysis, the objectives of the MNO and VSPs are opposite to each other. Particularly, VSPs are willing to get as many as fraction of SBSs for caching to achieve a maximum utility. On the other hand, the MNO tends to lease less fraction of SBSs to save its maintaining cost. Therefore, it is difficult for them to reach an agreement. Concerning this problem, a market broker (system controller) is necessary to operate the caching service market effectively and efficiently. In the market, the broker is paid according to the volume of transactions it facilitates. Therefore, in this work, the broker is considered to be honest and has no incentive to distort the transaction efficiency [4]. To solve this problem, an iterative algorithm, Iterative Double Auction (I-DA), is introduced to make sure that the maximum social welfare can be achieved, with assistance of properly designed resource allocation scheme and pricing rule.

In the simulation study, numerical results to demonstrate and test the validity and effectiveness of the proposed I-DA algorithm for the small-cell based video caching system. In addition, the convergence and economic properties of the I-DA based caching mechanism are also verified through the simulation.

Finally, the simulation results revealed the effect of some parameters, such as the step size of iterative algorithm, the number of VSPs, on the system performance and algorithm performance. Moreover, the convergence, economic efficiency, individually rationality, incentive compatibility, and budget balance properties [5] of the proposed I-DA based caching mechanism are also validated by the simulation results.

### References:

- [1] P. Juluri, V. Tamarapalli, and D. Medhi, "Measurement of quality of experience of video-on-demand services: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 401–418, 1st Quart., 2016.
- [2] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 849–869, Feb. 2016.
- [3] K. Poularakis, G. Iosifidis, and L. Tassiulas, "Approximation algorithms for mobile data caching in small cell networks," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3665–3677, Oct. 2014.
- [4] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "A double-auction mechanism for mobile data-offloading markets," *IEEE/ACM Trans. Netw.*, vol. 23, no. 5, pp. 1634–1647, Oct. 2015.
- [5] S. Paris, F. Martignon, I. Filippini, and L. Chen, "An efficient auctionbased mechanism for mobile data offloading," *IEEE Trans. Mobile Comput.*, vol. 14, no. 8, pp. 1573–1586, Aug. 2015.



**Shengjie Xu** [SM'14-M'19] received a Ph.D. degree in Computer Engineering from the University of Nebraska-Lincoln in 2019, and an M.S. degree in Telecommunications from the University of Pittsburgh in 2014. Before that, he held a B.E. degree in Information Assurance in 2012. He has a strong expertise in cyber security, certified by the NSA/DHS National Center of Academic Excellence in Information Assurance and Cyber Defense (CNSS 4011-4015). Presently, Dr. Xu is an assistant professor of The Beacom College of Computer and Cyber Sciences at Dakota State University. He serves as a Technical Editor of *IEEE Wireless Communications* and an Editor of *International Journal of Sensor Networks*. His research interests include cyber security, machine learning, critical infrastructure protection, and intelligent networking system. He is a member of IEEE, ACM, and AAI. He is also involved in IEEE Technical Committees, including CISTC, TCGCC, and TC

## A Creative and Fundamental Study on the Coding Technique for Wireless Caching Systems

*A short review for "On the Fundamental Limits of Coded Caching Systems With Restricted Demand Types"*

Edited by Cong Shen

*S. Shao, J. Gómez-Vilardebó, K. Zhang and C. Tian, "On the Fundamental Limits of Coded Caching Systems With Restricted Demand Types," in IEEE Transactions on Communications, vol. 69, no. 2, pp. 863-873, Feb. 2021.*

Over the last decade, content delivery has been the main reason of cellular traffic [1]. According to [2], video applications account for 70.5% of the total data traffic in 2016. Since content is usually generated completely long before transmission, caching becomes a quite practical and efficient solution. Popular content, such as movies or TV episodes, is cached in the period of low network utilization, so that traffic of network in high network utilization is reduced. Recently, an information theoretic framework is proposed for wireless caching systems [3]. In this framework, a caching system works in two phases. One phase is called the caching phase, and another is called the delivery phase. In the caching phase, each user caches data in advance without knowing its exact demand. Then in the delivery phase, the central server will broadcast data to all users depending on their requests. Also, in [3], it shows that in such wireless caching systems, a carefully designed coded caching and delivery scheme will have a transmission rate gain over uncoded ones. However, the optimal coding scheme in general still remains unknown.

This paper proposes a novel caching and delivery coding scheme for caching systems where there's a restriction on the demand of users. By the result of [4], the transmission rate in the delivery phase varies with user's demand. Meanwhile, the demand that all files are requested is not necessary the case which has the largest communication rate in the delivery phase. Therefore, as a natural idea, coding schemes designed for different class of demands specifically can have a better rate performance than the schemes designed for mixed demands, if the prior information of demand is known to the server. Motivated by this idea, this paper focus on the caching systems where every

file is requested by at least one user, which is a pretty practical setup when the cached files are the really popular ones.

In the prefetching phase, this paper uses a coded prefetching scheme instead of an uncoded one. Previous result [5] shows that when the prefetching scheme is required to be uncoded, the coding scheme in [3] is optimal; however [6] shows that in many cases a scheme with coded prefetching can have a lower transmission rate than ones with uncoded prefetching. In this paper, a mixture of uncoded and coded prefetching is applied. Each file is partitioned into multiple sub-files, and each sub-file has a three dimensioned index. The sub-files with some certain indices are prefetched uncodedly according to the prefetching scheme in [3]. The rest of sub-files are coded by following a product code structure.

Besides the prefetching scheme, the proposed coding scheme has also subtly designed the transmission scheme in the delivery phase. The authors' previous work [7] designed a coding scheme for the case where the number of users equals to the number files, and each file is requested by exactly one user. However, the scheme in [7] cannot easily extend to the cases with general number of users, since the transmitted message will have a linear dependence when some of the files which are requested by even number of users and hence a decoding failure occurs. Therefore, in this paper, the proposed scheme brings out an idea called I/Q copy. That is to say, they further partition a sub-file into an I copy and Q copy. In the delivery phase, if one file is requested by even number of users, an invertible linear transform will be applied on all sub-files of that very file and the linear dependency no longer

exists. Meanwhile since the transform is invertible, successfully decoding the transformed sub-files is equivalent to successfully decoding the original sub-files.

Simulation results show that the proposed scheme can indeed achieve new operational corner points that are not covered by any previous works. Meanwhile this scheme still works if the restriction on demand combination does not hold anymore, however the transmission rate will increase.

In summary, this paper studies the fundamental coding technique for caching systems. Instead of designing a scheme which has a same rate for all different demands, the proposed scheme can have a better rate performance when every file is requested by at least one user, at the cost of rate when such restriction does not hold. Meanwhile, some interesting coding techniques are applied in this paper. All in all, this paper brings us novel solution and inspiring techniques for the coding problems in caching systems.

### References:

- [1] M. A. Maddah-Ali and U. Niesen, "Coding for caching: fundamental limits and practical challenges," in *IEEE Communications Magazine*, vol. 54, no. 8, pp. 23-29, August 2016.
- [2] T. Han, N. Ansari, M. Wu and H. Yu, "On Accelerating Content Delivery in Mobile Networks," in *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1314-1333, Third Quarter 2013.
- [3] M. A. Maddah-Ali and U. Niesen, "Fundamental Limits of Caching," in *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2856-2867, May 2014.
- [4] C. Tian, "Symmetry, outer bounds, and code constructions: A computer aided investigation on the fundamental limits of caching," *Entropy*, vol. 20, no. 8, p. 603, Aug. 2018.
- [5] Q. Yu, M. A. Maddah-Ali and A. S. Avestimehr, "The Exact Rate-Memory Tradeoff for Caching with Uncoded Prefetching," in *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1281-1296, Feb. 2018.
- [6] C. Tian and J. Chen, "Caching and delivery via interference elimination," *Transactions on Information Theory*, vol. 64, no. 3, pp. 1548-1560, Mar. 2018.
- [7] S. Shao, J. Gómez-Vilardebó, K. Zhang and C. Tian, "On the Fundamental Limit of Coded Caching Systems with a Single Demand Type," 2019 *IEEE Information Theory Workshop (ITW)*, Visby, Sweden, 2019, pp. 1-5.



**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 received his Ph.D. degree from the Electrical Engineering Department of UCLA in 2009. He is currently an Assistant Professor in the Electrical and Computer Engineering Department at University of Virginia. He is a Senior Member of IEEE and serves as editor for the *IEEE Transactions on Wireless Communications* and *IEEE Wireless Communications Letters*.



## Federated Learning Review in Cooperative Systems

*A short review for “Opportunities of Federated Learning in Connected, Cooperative, and Automated Industrial Systems”*

Edited by Mohammad H Hasan

*S. Savazzi, M. Nicoli, M. Bennis, S. Kianoush and L. Barbieri, "Opportunities of Federated Learning in Connected, Cooperative, and Automated Industrial Systems," in IEEE Communications Magazine, vol. 59, no. 2, pp. 16-21, February 2021, doi: 10.1109/MCOM.001.2000200.*

Improving the intelligence of devices is an everlasting pursuit in the age of intelligence and the internet of things. Technologies such as self-driving vehicles and collaborative robots (Cobots) are developed. In these technologies, there is an ever-present need for additional data from the environment for dynamical model improvement and for improving the response of the output of the deep neural network (DNN) model in these systems. Device-to-device communication serves to improve the performance of individual devices, with the increased availability of data for model training, and communications, enabling devices to access information about the environment that are otherwise inaccessible locally.

However, free communication of data cooperative systems through a conventional mobile edge cloud (MEC) has multiple drawbacks. One of which is the security concern arising from the transfer of local device data, in applications such as self-driving cars, into an external server [1]. Another concern arises from the high computational overhead due to the large size of data collected locally in each individual device in the cooperative system.

Federated learning (FL) is a distributed learning method which does not require sending local training data to a server [2-3]. Federated averaging is one of the most popular FL implementations. The decentralized learning method can have various applications in mission critical control scenarios, such as collaborative industrial robotics and cooperative automated vehicles.

This paper addresses the opportunities of emerging distributed FL tools for system with autonomous industrial components, such as

vehicles, robotics. FL is proposed to integrate with the sensing-decision-action loop.

The author’s major contribution of this work is the summarization of the state-of-the-art in the field of FL. The paper reviews different means of exchanging local parameters between devices (federated nodes) to facilitate the update of a superior local model in each device. Present two avenues for incorporating FL: cooperative automated driving and cobots. Finally, the paper concludes with a demonstration of the use of FL in cobots, showing an improved performance of robots trained using FL, versus others that utilize opportunistic models with no access to training data externally.

This paper present two main approaches for the exchange of local parameters between federated nodes: consensus-based approaches, and diffusion-based approaches. In the former approach, federated nodes exchange their own local ML parameters, such as connection weights and biases in a DNN with each other. The updated local ML parameters are consequently set to the distributed average of the local parameters exchanged. In this approach, communication overhead is relatively low, especially implementing sparse local ML parameters in the D2D communication. This approach appears to be feasible in the 15-cobot group network example presented in this paper when the dataset is independent & identically distributed (IID).

The second approach for parameter exchange presented in this paper, Diffusion FL, trades off additional communication overhead for a better performance in non-IID datasets. In this case, the federated nodes within the network exchange both their local ML parameters as well as their gradients

to their neighbors. This approach significantly increases the convergence time of the FL as the network size increases.

Finally, this paper presents the potential of using FL to facilitate cooperation in automated driving systems. FL may use distributed ML to combine their sensory information together to improve environment monitoring. Moreover, through FL, vehicles with inaccurate sensors may improve their vision systems through cooperation with vehicles with more accurate sensors.

In summary, this paper highlights some prevalent approaches for FL. It discusses the tradeoffs of each approach and the overall potential of FL in cooperative and industrial systems, concluding with remarks about the potential of FL in cooperative heterogeneous device cooperation tasks.

### References:

- [1] Liu, D., & Simeone, O. (2020). Privacy for free: Wireless federated learning via uncoded transmission with adaptive power control. *IEEE Journal on Selected Areas in Communications*, 39(1), 170-185.
- [2] Konečný J., et al. Konečný, J., McMahan, H.B., Ramage, D. and Richtárik, P., 2016. Federated optimization: Distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527.
- [3] Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z.,

Cormode, G., Cummings, R. and d'Oliveira, R.G., 2019. Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977.



**Mohammad H Hasan** Received his Ph. D in 2020 from University of Nebraska – Lincoln with a focus on neuromorphic (neuro-inspired) computing in non-digital hardware components, artificial intelligence-augmented micro-sensors, novel robotic systems (e.g., soft and micro-robots). He is currently an assistant professor of Robotics Engineering in Columbus State University since 2021. His Research interests include: artificial intelligence, dynamical systems and chaos, and robotics.

## A Proof of Concept of Mulsemmedia Streaming and its Advantages in a Real-world Learning Environment

*A short review for "Improving Student Learning Satisfaction by Using an Innovative DASH-based Multiple Sensorial Media Delivery Solution"*

Edited by Lucile Sassatelli

*T. Bi, R. Lyons, G. Fox and G. -M. Muntean, "Improving Student Learning Satisfaction by Using an Innovative DASH-based Multiple Sensorial Media Delivery Solution," in IEEE Transactions on Multimedia, 2020. doi: 10.1109/TMM.2020.3025669.*

The quest for immersion cannot stop at sound and sight only, but should rather incorporate other human senses which are key in connecting a human with an environment, present or remote. Multiple sensorial media, also named mulsemmedia, aim at extending traditional media with tactile, gustatory and olfaction components. These components can enhance presence in virtual environments [1,2]. In this article, Bi et al. investigate whether such additional sensory components can benefit objectives beyond the user's enjoyment, in particular whether it can benefit the learning process in the classroom. The authors focus on the issue of STEM learning, within the framework of NEWTON, a European project aimed at bringing innovative and immersive teaching and learning solutions to the classroom, focusing specifically on STEM.

The authors build a mulsemmedia testbed made of a programmable fan and sent diffuser, a haptic mouse and a computer where the students watched the educational STEM video.

Delivering this kind of augmented experience to the remote classroom requires to design transmission systems able to consider such heterogeneous content, while still enabling adaptation of the delivery to the conditions of the user. To do so, the authors build on a multi-sensorial extension of the MPEG DASH video streaming standard, presented in [3] and named DASHMS. The principle is to enable adaptation of the delivery of mulsemmedia content (video and other sensory components), so as to maximize the user's quality of experience. The quality of experience objective function is defined as a

weighted sum of video quality and energy consumption of the mulsemmedia devices. Indeed, the video content remains the most demanding component in terms of transmission throughput from the server to the client, the other components requiring only metadata to control their actuation. This objective function therefore allows to consider both network conditions and user profile or preference.

The user experiments involved 44 business school students who had to watch STEM videos and were then questioned about their level of enjoyment and annoyance, and were assessed with questionnaires on the acquired knowledge (on the remembering, applying and evaluating Bloom's learning dimensions). Three different conditions were evaluated: (1) no multi-sensory effect with video and sound, (2) multi-sensory effect (airflow, haptic, olfaction) with olfaction scent adapted to the video content, (3) multi-sensory effect with constant olfaction scent.

The majority of the students reported to have enjoyed the experience, and that the haptic mouse was their least favored mulsemmedia effect. While they enjoyed the experience, a significant fraction of users reported being disturbed in their learning task by the multi-sensory effect. In terms of learning improvement thanks to mulsemmedia effects, results were less conclusive, as some smells on certain videos improved the learning outcomes while others impaired it. It seems that the learning improvement depends on the type of smell, and that matching the smell with the video content may not always be straightforward.

## IEEE COMSOC MMTC Communications – Review

The interaction of learning and olfactory environment therefore warrants more investigation, as other works bring supporting evidence of a beneficial interaction [4]. Of particular importance is the matching of the video content with the type of smell depending on the learning objective.

The authors have therefore demonstrated a real-world deployment of a server-to-client multimedia streaming systems based on DASHMS and incorporated in a full-fledged learning management system [5]. This opens up the way for more diverse experiments of immersive media accessed from remote classrooms for innovative and more efficient learning solutions.

### References:

- [1] N. Murray, B. Lee, Y. Qiao, and G.-M. Muntean, "Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges," *ACM Computing Surveys (CSUR)*, vol. 48, no. 4, p. 56, 2016.
- [2] N. Ranasinghe, T. N. T. Nguyen, Y. Liangkun, L.-Y. Lin, D. Tolley, and E. Y. Do, "Vocktail: A virtual cocktail for pairing digital taste, smell, and color sensations," in *Proceedings of ACM Multimedia Conference (MM'17)*, ACM, New York, 1139–1147, 2017.
- [3] L. Zou, T. Bi, and G.-M. Muntean, "A dash-based adaptive multiple sensorial content delivery solution for improved user quality of experience," *IEEE Access*, vol. 7, pp. 89 172–89 187, 2019.

- [4] J. Degel and E. P. Köster, "Odors: implicit memory and performance effects," *Chemical Senses*, vol. 24, no. 3, pp. 317–325, 1999.
- [5] NEWTELP platform webpage, <http://newtelp.eu>.



**Lucile Sassatelli**, Ph.D, is an associate professor with Université Côte d'Azur, France, since 2009. She has been nominated a junior fellow of Institut Universitaire de France (IUF) in 2019. She obtained a PhD from Université of Cergy Pontoise, France, in 2008 and her professorial habilitation in 2019. She was a postdoctoral fellow at MIT, Cambridge, USA. She focuses on the problems of multimedia transmission of virtual and augmented reality, specifically on machine learning-based streaming approaches.

She has published papers in prestigious journals such as *IEEE Transactions on Information Theory*, *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and conferences such as *IEEE INFOCOM*.

MMTC Communications – Review Editorial Board

DIRECTORS

**Zhisheng Yan**

Georgia State University, USA  
Email: zyan@gsu.edu

**Wenming Cao**

Shenzhen University, China  
Email: wmcao@szu.edu.cn

**Yao Liu**

Binghamton University, USA  
Email: yaoliu@binghamton.edu

**Phoenix Fang**

California Polytechnic State University, USA  
Email: dofang@calpoly.edu

EDITORS

**Carsten Griwodz**

University of Oslo, Norway

**Mengbai Xiao**

Shandong University, China

**Ing. Carl James Debono**

University of Malta, Malta

**Marek Domański**

Poznań University of Technology, Poland

**Xiaohu Ge**

Huazhong University of Science and Technology,  
China

**Roberto Gerson De Albuquerque Azevedo**

EPFL, Switzerland

**Frank Hartung**

FH Aachen University of Applied Sciences,  
Germany

**Pavel Korshunov**

EPFL, Switzerland

**Ye Liu**

Nanjing Agricultural University, China

**Luca De Cicco**

Politecnico di Bari, Italy

**Bruno Macchiavello**

University of Brasilia (UnB), Brazil

**Yong Luo**

Nanyang Technological University, Singapore

**Debashis Sen**

Indian Institute of Technology - Kharagpur, India

**Guitao Cao**

East China Normal University, China

**Mukesh Saini**

Indian Institute of Technology, Ropar, India

**Roberto Gerson De Albuquerque Azevedo**

EPFL, Switzerland

**Cong Shen**

University of Virginia, USA

**Qin Wang**

Nanjing University of Posts &  
Telecommunications, China

**Stefano Petrangeli**

Adobe, USA

**Rui Wang**

Tongji University, China

**Jinbo Xiong**

Fujian Normal University, China

**Qichao Xu**

Shanghai University, China

**Lucile Sassatelli**

Université de Nice, France

**Shengjie Xu**

Dakota State University, USA

**Tiesong Zhao**

Fuzhou University, China

**Takuya Fujihashi**

Osaka University, Japan

## IEEE COMSOC MMTC Communications – Review

### Multimedia Communications Technical Committee Officers

**Chair:** Jun Wu, Fudan University, China

**Steering Committee Chair:** Joel J. P. C. Rodrigues, Federal University of Piauí (UFPI), Brazil

**Vice Chair – America:** Shaoen Wu, Illinois State University, USA

**Vice Chair – Asia:** Liang Zhou, Nanjing University of Post and Telecommunications, China

**Vice Chair – Europe:** Abderrahim Benslimane, University of Avignon, France

**Letters & Member Communications:** Qing Yang, University of North Texas, USA

**Secretary:** Han Hu, Beijing Institute of Technology, China

**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.