Guest Editorial Caching for Communication Systems and Networks—Part II

Georgios S. Paschos, *Senior Member, IEEE*, George Iosifidis, Meixia Tao, *Senior Member, IEEE*, Don Towsley, *Fellow, IEEE*, and Giuseppe Caire, *Fellow, IEEE*

WELCOME to the second part of the IEEE JSAC special issue on Caching for Communication Systems and Networks. The goal of this special issue is to present the multiple facets of caching, from information theory to networking and services, and explore the role of memory in communications. This is a very timely topic due to recent technological and theoretical advances summarized in the tutorial paper that appears in the first part of the issue [1].

The concept of caching was introduced more than 40 years ago for enhancing the performance of computer systems [2], and since then has been used to support content delivery in the Internet, and recently over wireless networks [3]. As the user demand for (mobile) content grows fast, caching is expected to play a very important role in future communication systems and networks. Indeed, there is consensus today that caching can increase network performance, reduce expenditures for operators and infrastructure owners, and improve the quality of services offered to users.

Therefore, it is not surprising that a number of different industries focus today on content caching, or leverage this technique in order to improve their services. Prominent examples include content distribution networks such as the Akamai CDN; cloud storage services offered by Amazon; service-specific caching networks as those deployed by Google and Facebook; and caching techniques tailored to wireless networks as those proposed by Cadami. Furthermore, 3GPP working groups have proposed standards for enabling caching at the Radio Access Network or even at base stations, hence opening new opportunities for caching in wireless networks. These developments call for further research on caching theory and caching architectures.

This special issue has called for papers from the entire spectrum of caching research, aiming to present the latest developments in the various caching topics, and to offer a unified view on this important subject. We were delighted to receive 118 papers spanning all research topics in caching (see Table I), and using different tools such as information theory, optimization theory, large-scale simulations, data analysis and experiments. Following a thorough and very competitive review process 33 papers were accepted and are presented in this double special issue. We hope that this selection of papers will highlight recent findings in caching and contribute towards the further development of this important tool. To this end, the first part of the issue includes also a tutorial [1].

TABLE I
CACHING TOPICS STATISTICS

Topic and Keywords	Papers
Information-theoretic Caching Analysis	16
Fundamental Limits of Caching	16
Coded Caching Design with Practical Constraints	12
Scaling Laws of Cache Networks	7
HetNet and Device-to-device Caching	18
Edge Caching, Cooperation and Femtocaching	32
Joint Caching, Scheduling and Routing	19
Secure Caching and Privacy Preservation	6
Content Caching and Delivery	56
Algorithms for Storage Placement	17
Video caching and Streaming	14
Caching Economics	13
Caching models for ICN	12
Popularity Models and Machine Learning	9

The second part of this special double issue includes 16 papers. It begins with the paper "On coding for cache-aided delivery of dynamic correlated content" [4] which generalizes the idea of coded caching by considering correlated content items. The authors show that this correlation can be leveraged to obtain an improved lower bound for the rate-memory trade-off. Next, the paper "Cache aided communications with multiple antennas at finite SNR" [5] introduces a linear program for devising joint cache allocation and (pre-) fetching schemes that minimize content delivery delay. The proposed solution employs zero-forcing and cached interference subtraction, and hence allows each user to be served at the rate of its own channel rather than of the poorest channel in the group. Similarly, the paper "Utility optimal scheduling for coded caching in general topologies" [6] is motivated by the observation that the performance of coded caching over wireless fading channels is limited by the user with the worst fading condition. The paper suggests a new opportunistic content delivery scheduling policy in order to address this problem.

G. S. Paschos is with the France Research Center, Huawei Technologies, 92100 Paris, France (e-mail: georgios.paschos@huawei.com).

G. Iosifidis is with the School of Computer Science and Statistics, Trinity College Dublin, Dublin 2, Ireland (e-mail: george.iosifidis@tcd.ie).

M. Tao is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: mxtao@sjtu.edu.cn).

D. Towsley is with the School of Computer Science, University of Massachusetts Amherst, Amherst, MA 01002 USA (e-mail: towsley@cs.umass.edu).

G. Caire is with the Department of Electrical Engineering and Computer Science, Technical University of Berlin, 10623 Berlin, Germany (e-mail: caire@tu-berlin.de).

Digital Object Identifier 10.1109/JSAC.2018.2844578

The fourth paper in this issue, "Caching and coded delivery over Gaussian broadcast channels for energy efficiency" [7], extends coded caching to wireless broadcast channels with Gaussian noise, and studies methods that improve energy efficiency. The paper "On the effects of subpacketization in content-centric mobile networks" [8] analyzes the throughput-delay trade-offs in an ad hoc network where each node moves according to a reshuffling mobility model, and extends prior work to this scenario where subpacketization is used.

Next, the paper "Optimized base-station cache allocation for cloud radio access network with multicast backhaul" [9] uses a mixed time-scale problem to optimize cache dimensioning at the base stations and beamforming decisions, aiming to improve content delivery in C-RAN wireless networks. C-RAN is also the focus of the paper "Online resource allocation, content placement and request routing for cost-efficient edge caching in cloud radio access networks" [10]. The authors consider a cooperative system where cloud resources at the RAN can be shared, and design joint content placement and routing policies for minimizing system costs.

The paper "Hierarchical edge caching in device-to-device aided mobile networks: modeling, optimization, and design" [11] proposes the idea of improving caching policies in small cell networks and mobile devices by considering social-layer characteristics, such as mutual user interests and mobility patterns. The paper "Mobility prediction-assisted over-the-top edge prefetching for hierarchical VANETs" [12] presents traces from a vehicular ad-hoc network. It is demonstrated that prediction-enhanced content prefetching can indeed increase the network performance. In this case the predictions refer both to content popularity and the location of the vehicles.

Zhang *et al.* [13] "Energy-efficient caching for scalable videos in heterogeneous networks" consider video caching over heterogeneous networks modeled with Poison Point Processes (PPP), and study the impact of different viewing quality requirements on the network's energy efficiency. The work "QoE-assured 4K HTTP live streaming via transient segment holding at mobile edge" [14] studies HTTP-based live streaming services where mobile devices request ultra-high video quality. The authors propose a novel scheme which performs context-aware transient holding of video segments at the cache-enabled mobile edge, aiming to reduce initial startup delay and live stream latency.

The next paper in this issue "Caching with partial adaptive matching" [15], studies user request matching to caches. It compares a scheme focusing on coded server transmissions while ignoring matching capabilities, with a scheme that focuses on adaptive matching while ignoring potential coding opportunities. Next, the paper "Coded joint pushing and caching with asynchronous user requests" [16] extends the coded caching framework to an online approach that designs placement and delivery while considering user demand asynchronism.

In a different setting, the paper "ReD/LeD: An asymptotically optimal and scalable online algorithm for service caching at the edge" [17], proposes an online algorithm to address an important question in hierarchical networks: having a set of edge servers with limited storage capacity, which *services* to cache in order to reduce the network costs? The authors propose an algorithm for making such caching decisions in an online fashion, and without requiring prior knowledge of service popularity parameters.

The paper "Parallel simulation of very large-scale general cache networks" [18] studies how to efficiently simulate very large caching systems, a task known to be very compute-intensive. The authors propose model-driven techniques, and leverage very accurate approximations of caching systems that enable the simulation of caching algorithms in networks with hundreds of caches and trillions of content items. The issue is completed with the paper "Understanding scoped-flooding for content discovery and caching in content networks" [19] which designs a new technique for propagating control signals for content discovery, by building multicast trees routed at the source node. Since replica frequency is expected to relate to popularity, the authors suggest tuning the discovery radius according to content popularity.

ACKNOWLEDGMENTS

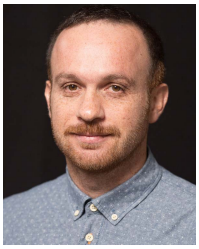
The guest editors would like to thank all the authors who submitted their work in this JSAC issue, and acknowledge the excellent work of all reviewers who participated in the review process. Also, the mentoring and help provided by Prof. Max Henrique Machado Costa, JSAC Senior Editor, and by Laurel Greenidge, Executive Editor, are greatly acknowledged.

REFERENCES

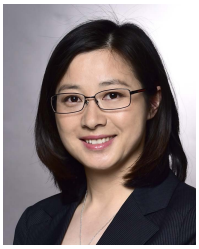
[1] G. Paschos, G. Iosifidis, M. Tao, D. Towsley, and G. Caire, "The role of caching in future communication systems and networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, 2018.
[2] L. A. Belady, "A study of replacement algorithms for a virtual-storage computer," *IBM Syst. J.*, vol. 5, no. 2, pp. 78–101, 1966.
[3] G. Paschos, E. Bastug, I. Land, G. Caire, and M. Debbah, "Wireless caching: Technical misconceptions and business barriers," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 16–22, Aug. 2016.
[4] P. Hassanzadeh, A. Tulino, J. Llorca, and E. Erkip, "On coding for cache-aided delivery of dynamic correlated content," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[5] I. Bergel and S. Mohajer, "Cache aided communications with multiple antennas at finite SNR," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[6] R. Combes, A. Ghorbel, M. Kobayashi, and S. Yang, "Utility optimal scheduling for coded caching in general topologies," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[7] M. M. Amiri and D. Gündüz, "Caching and coded delivery over Gaussian broadcast channels for energy efficiency," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[8] A. Malik, S. H. Lim, and W.-Y. Shin, "On the effects of subpacketization in content-centric mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[9] B. Dai, Y.-F. Liu, and W. Yu, "Optimized base-station cache allocation for cloud radio access network with multicast backhaul," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[10] L. Pu, L. Jiao, X. Chen, L. Wang, Q. Xie, and J. Xu, "Online resource allocation, content placement and request routing for cost-efficient edge caching in cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[11] X. Li, X. Wang, P.-J. Wan, Z. Han, and V. C. M. Leung, "Hierarchical edge caching in device-to-device aided mobile networks: Modeling, optimization, and design," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[12] Z. Zhao, L. Guardalben, M. Karimzadeh, J. Silva, T. I. Braun, and S. Sargento, "Mobility prediction-assisted over-the-top edge prefetching for hierarchical VANETs," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[13] X. Zhang, T. Lv, W. Ni, J. Cioffi, N. C. Beaulieu, and Y. J. Guo, "Energy-efficient caching for scalable videos in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.

[14] C. Ge, N. Wang, W. K. Chai, R. Bradbury, and H. Hellwagner, "QoE-assured 4K HTTP live streaming via transient segment holding at mobile edge," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[15] J. Hachem, N. Karamchandani, S. Moharir, and S. Diggavi, "Caching with partial adaptive matching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[16] Y. Lu, W. Chen, and H. V. Poor, "Coded joint pushing and caching with asynchronous user requests," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[17] T. Zhao, I.-H. Hou, S. Wang, and K. S. Chan, "ReD/LeD: An asymptotically optimal and scalable online algorithm for service caching at the edge," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[18] M. Tortelli, D. Rossi, and E. Leonardi, "Parallel simulation of very large-scale general cache networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.
[19] L. Wang, S. Bayhan, J. Ott, J. Kangasharju, and J. Crowcroft, "Understanding scoped-flooding for content discovery and caching in content networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, 2018.

Georgios S. Paschos (S'97–M'02–SM'15) received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 2002, and the Ph.D. degree in wireless networks from the Electrical and Computer Engineering Department, University of Patras, Greece, in 2006. He held research positions at VTT, Finland, from 2007 to 2008, CERTH-ITI, Greece, from 2008 to 2012, and LIDS, MIT, USA, from 2012 to 2014. From 2009 to 2012, he was with the Electrical and Computer Engineering Department, University of Thessaly. Since 2014, he has been leading the Network Control and Resource Allocation Team, Huawei Technologies, Paris, France, where he is currently a Principal Researcher. Two of his papers received the Best Paper Award, in GLOBECOM 07 and IFIP Wireless Days 09, respectively. He was the Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS special issue for content caching and delivery, and actively serves as an Associate Editor for the IEEE/ACM TRANACTIONS ON NETWORKING and the IEEE NETWORKING LETTERS, and as a Technical Program Committee Member of INFOCOM, WiOPT, and Netsoft.

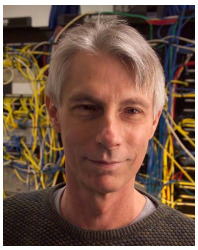


George Iosifidis received the Diploma degree in electronics and communications from the Greek Air Force Academy, Athens, in 2000, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Thessaly, in 2012. He was a Post-Doctoral Researcher with CERTH-ITI, Greece, from 2012 to 2014, and a Post-Doctoral/Associate Research Scientist with Yale University from 2014 to 2017. He is currently the Ussher Assistant Professor in future networks with the Trinity College Dublin, Ireland. He was a co-recipient of the Best Paper Awards in WiOPT 2013 and the IEEE INFOCOM 2017 conferences, and received the SFI Career Development Award in 2018.



Meixia Tao (S'00–M'04–SM'10) received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology in 2003. She was a member of the Professional Staff with the Hong Kong Applied Science and Technology Research Institute from 2003 to 2004, and a Teaching Fellow and then as an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, from 2004 to 2007. She is currently a Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Her current research interests include wireless caching, physical-layer multicasting, resource allocation, and interference management.

Dr. Tao currently serves as a member of the Executive Editorial Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2009. She was a recipient of the WCSP 2012 Best Paper Award in 2012, the IEEE Heinrich Hertz Award for the Best IEEE COMMUNICATIONS LETTERS in 2013, and the IEEE/CIC ICCC 2015 Best Paper Award in 2015. She was on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011, the IEEE COMMUNICATIONS LETTERS from 2009 to 2012, and the IEEE WIRELESS COMMUNICATIONS LETTERS from 2011 to 2015.



Don Towsley received the B.A. degree in physics and the Ph.D. degree in computer science from The University of Texas, in 1971 and 1975, respectively. He has held visiting positions at numerous universities and research labs, including the University of Paris VI, IBM Research, AT&T Research, Microsoft Research, and INRIA. He is currently a Distinguished Professor with the College of Information and Computer Sciences, University of Massachusetts. His research interests include security, quantum communications, and networks and performance evaluation. He is a Co-Founder of *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* and served as one of its first Co-Editor-in-Chiefs. He served as the Editor-in-Chief of the IEEE/ACM TRANSACTIONS ON NETWORKING and on numerous other editorial boards. He has served as the Program Co-Chair for numerous conferences and on the program committees of many other. He is a Corresponding Member of the Brazilian Academy of Sciences and has received numerous IEEE and ACM awards, including the 2007 IEEE Koji Kobayashi Award, the 2007 ACM SIGMETRICS Achievement Award, and the 2008 ACM SIGCOMM Achievement Award. He has also received numerous best paper awards, including the IEEE Communications Society 1998 William Bennett Paper Award, the 2008 ACM SIGCOMM Test of Time Award, the 10+ Year 2010 DASFAA Best Paper Award, the 2012 ACM SIGMETRICS Test of Time Award, and five ACM SIGMETRICS Best Paper Awards. He is a fellow of the ACM.

Giuseppe Caire (S'92–M'94–SM'03–F'05) was born in Turin, Italy, in 1965. He received the B.Sc. degree in electrical engineering from the Politecnico di Torino, Italy, in 1990, the M.Sc. degree in electrical engineering from Princeton University in 1992, and the Ph.D. degree from the Politecnico di Torino in 1994. He was a Post-Doctoral Research Fellow with the European Space Agency, ESTEC, Noordwijk, The Netherlands, from 1994 to 1995, an Assistant Professor in telecommunications with the Politecnico di Torino, an Associate Professor with the University of Parma, Italy, and a Professor with the Department of Mobile Communications, Eurecom Institute, Sophia-Antipolis, France. He is currently a Professor of electrical engineering with the Viterbi School of Engineering, University of Southern California, Los Angeles, and an Alexander von Humboldt Professor with the Electrical Engineering and Computer Science Department, Technical University of Berlin, Germany. His main research interests are in the field of communications theory, information theory, channel and source coding with particular focus on wireless communications. He served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1998 to 2001 and as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2001 to 2003. He received the Jack Neubauer Best System Paper Award from the IEEE Vehicular Technology Society in 2003, the IEEE Communications Society and Information Theory Society Joint Paper Award in 2004 and 2011, the Okawa Research Award in 2006, the Alexander von Humboldt Professorship in 2014, and the Vodafone Innovation Prize in 2015. He has served on the Board of Governors of the IEEE Information Theory Society from 2004 to 2007, and as an Officer from 2008 to 2013. He was the President of the IEEE Information Theory Society in 2011.