Wireless Cooperative Caching in Device-to-Device Networks: Simulation and Modeling

Wei Li[1], Jiayi Lu[1], Yi Xu[1], Zeru Wei[1], Yuanwen Tian[1], Yiming Miao[1,*]
*Corresponding author: Yiming Miao

School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
`{weili_epic,yimingmiao}@hust.edu.cn,{jiayilu.epic,xuyi.epic,zeruwei.epic}@qq.com,yuanwentian1997@gmail.com`

Abstract. The ever increasing demand for content is straining operators networks, which encourages the development of content delivery mechanisms. Mobile edge caching and device-to-device technologies constitute a promising solution for reducing the effects of demand growth by placing popular content files in proximity to uses. Mobile Edge caching enables users obtain the requested contents from small cellular or other user equipments, rather than the content server through the mobile core networks, so as to alleviate the bottleneck of backhual in 5G network and reduce the delay. In this paper, we simulate the caching scenario in 5G networks by fusing with content centric network using OPNET Modeler, the obtained simulation results show that edge cache and local D2D cooperative cache can achieve higher hit rate, effectively reduce the delay of users' access to content, and reduce the core network traffic of 5G.

Key words: caching, device-to-device networks, simulation.

1 Introduction

With the popularity of terminal equipments and growth of data service demands [1], the network service providers suffer the huge pressure of data transmission. The effect of introducing the cache technology to reduce the network load has been increasingly prominent, its main performance is placing popular contents in the cache of terminal users or in the nearby servers, which can effectively reduce the network load in the peak of data requests and improve the bandwidth utilization [2], reduce the the transmission energy consumption for obtaining data services directly from the base station to multiple end users.

D2D is one of the key technologies of the fifth generation communication system (5G), mobile users can not only communicate with the local cellular base station, but also communicate with other mobile users with the support of device to device (D2D) communication [3]. Some popular videos are often repeatedly requested by mobile users, which has triggered the rise of D2D communication-

s [4]. Through the use of D2D links, adjacent mobile devices can share cached contents without communication via base stations to reduce the communication cost and delay [5, 6, 7]. With the increasing of performance of user terminal hardware, mobile phones also have potential capabilities of storage and computing [8, 9, 10, 11]. Therefore, content caching and sharing of mobile users has great potential for improving network performance [12, 13, 14, 15, 16]. The cooperative communication has already been widely researched in wireless sensor network [17, 18] and artificial intelligence [19], while not mature in wireless cellular network.

There are four ways for mobile user obtaining requested contents in cellular network that supports D2D [20]:

– **Local caching:** when user requests content, he or she first checks whether the content is cached locally. If cached, he or she gets content directly from the local without any delay.
– **D2D caching:** if not satisfied at local, he or she will request from neighboring users via D2D. If neighboring user caches the content, it would be sent to the requesting user.
– **SBS caching:** if not satisfied within D2D range, the interest request would be forwarded to the SBS (Small cell Base Station) to obtain the content.
– **MBS caching:** if not satisfied with above methods, the interest request would be forwarded to MBS (Macro cell Base Station).

The costs consumed with above ways are different, in which the cost of the cellular network is usually the communication cost, and the cost of obtaining content through the D2D network includes communication delay cost and collaboration cost (considering the selfishness of mobile users), which is much less than the communication cost of cellular networks [21, 22].

The contributions in this paper are as follows:

– A cooperative caching method is proposed that maximizes the probability of obtaining content from local or cooperative devices through the broadcast of adjacent nodes when content is not obtained within D2D range, thus reducing the total cost and the backhaul load.
– We realize the caching model in D2D networks by importing CCN strategy to all network elements on top of IP layer in the OPNET simulator [23, 24], and validate the effect of our proposed method more realistically by way of network simulation.

The remainder of this paper is organized as follows. Section 2 provides some related research work. Section 3 introduces the envisioned system architecture, Section 4 shows the simulation setup and discusses simulation results. Finally, Section 5 concludes this paper.

2 Related Work

Caching is a technique to improve performance well known in many wired network domains, e.g., content-centric networks (CCN) [25, 26, 27]. In traditional

wireless networks, caching popular contents in the edge could effectively reduce the backhaul cost [28]energy consumption and transmission delay [29, 30, 31, 32] as well as improving user experience, caching is also an effective way to improve the throughput of wireless communication system as well as the reliability of information routing in relay-based ad hoc systems [33, 34]. In reference [35], the authors cached popular contents both in core networks and access networks; Reference [36, 37] considered energy consumption, they formulate the optimal caching problem to obtain the optimal solution for minimizing the energy consumption of the base station and relay points. Considering that backhaul becomes a bottleneck in ultra dense networks (UDNs) of 5G while disk size increases quickly at a relatively low cost, the authors in [38] suggested equipping caches at the BSs to mitigate the load of backhaul links.

Recent studies have found that large amounts of multimedia contents, such as video, file, and audio occupy the majority of network traffic, however, only a small part (such as popular content) is frequently accessed by the majority of users [39, 40, 41, 42]. On the other hand, recent devices (such as smartphones, tablets) have larger and larger storage capacity and increase without expensive consumption, and they can cache more and more content in the future. So, the key issue of designing content placement policy is where to place and which contents should be placed.

The existing CCN caching policies have been widely studied, but their common approach is to use tools to track a period of content first, and then replay data in the cache algorithm simulation. Reference [43] fomulates the content placement problem and devises an optimal content placement policy to reduce the server load by serving as many requests as possible by the caches. Their methodology is based on expressing the content placement problem as the maximization of a submodular function subject to uniform matroid constraints with numerical experiments. The problem of above approach is its difficulty to statistic large data, and to perform simulation, while the CCN model we built is compatible with the existing Internet, and has a clear dynamic evaluation algorithm.

3 System Model

In this section, we first introduce the basic scenario of cooperative caching between mobile devices in UDNs of 5G utilizing D2D technology, then introduce the proposed caching policy and the basic data workflow of network model realized in this paper.

Fig. 1 illustrates the general scenario of user requesting interest content from adjacent devices in a way of D2D, or SBS and MBS when not satisfied at D2D ranges in ultra dense networks of 5G. When user A requests content that is not cached in local, first he will broadcast his request within D2D range, then nearby user B receives the request and checks in the cache, if found user B will deliver the content to user A directly, else broadcast the request again, then the content

would be hit at user C and delivered to user A through user B. If still not hit after broadcast twice, the request would be forwarded to MBS.
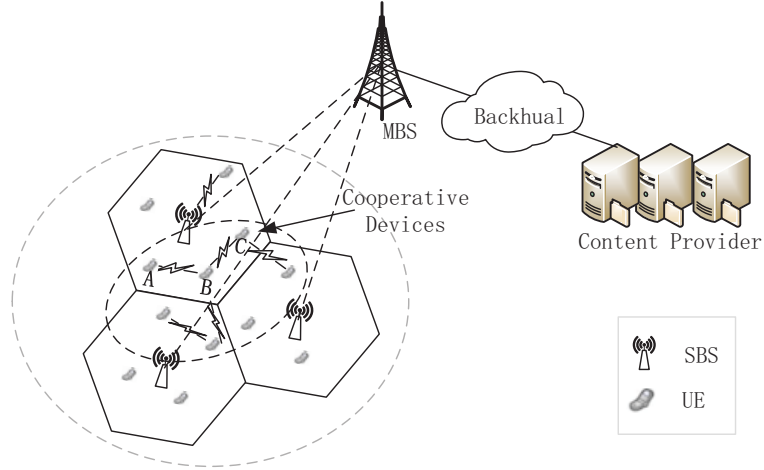

Fig. 1. Illustration of cooperative caching in D2D networks.

3.1 Caching policy

If each device stores the popular content, then each device stores almost the same content, then collaboration from neighboring nodes would have no effect. In order to ensure content diversity and avoid the similarity between cooperative devices, the cache is divided into two parts: unique part (cache the local request) and shared part (cache the common request in D2D network).

Thus we proposed a cooperative caching policy (CC, in short), when a network device received a content, it would do the following process:

– If the content came from cellular network, it would be compared with contents in unique part. If it has higher popularity, then replace with the lower one; Else cached the content if not full in unique part.
– If the content came from D2D network, meaning there are at least 2 copies of the content, it would be compared with contents in shared part. If it has higher popularity, then replace with the lower one; Else cached the content if not full in shared part.

3.2 Data workflow

In the network model, the mobile terminal continuously sends interest requests which follows the Pareto principle, that is, the common requested contents of 80% users only accounts for 20% of all contents. At the beginning, all the devices

do not cache any contents, with the continuous requesting and forwarding of interest packets and the response of data packets, each node device updates the local PIT (Pending Interest Table) table and determines whether a data packet should be cached or not according to the above cache policy and updates the local CS (Content Store) continuously.

The basic data workflow of each user equipment (UE, in short) requesting interest is depicted as Fig. 2. In order to hit the content at UEs as much as possible through D2D communication, UEs check the count that interest packet has been forwarded when not hit at local cache, if the count is smaller than a certain value, The UE would broadcast the interest packet, or forwards it according to FIB, thus increasing the hit rate of terminals.
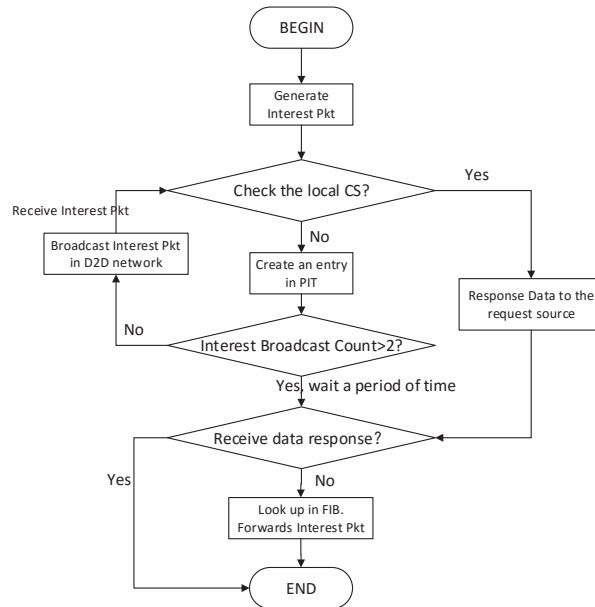
Fig. 2. Basic workflow of UE requesting content.

4 Simulation and Results

4.1 Network architecture

We realize the caching model in D2D networks by importing CCN strategy to all network elements on top of IP layer using OPNET 16.1. Though an open source NDN model named ndnSIM has been built up by NS3 [44], it is limited to mobile network simulation (e.g. 2.5G/3G/4G), while OPNET doesn't. The CCN model we built is compatible with the existing Internet, CCN is overlayed over the IP

layer. We integrated CCN processing modules into all netwrok elements, such as WLAN mobile station, Wireless Access Point (AP, function as SBS mentioned above), WiMAX BS, content server and IP Cloud.

Fig. 3 shows the hybrid simulation framework of Wlan and WiMax, the simulation network has 3 cells, each has an WiFi Access Point (AP) and 5 wireless mobile workstations (MSs). The MSs and AP should be configured in IEEE 802.11g standard, and access network applies WiMax broadband wireless access technology, connected to the backbone network through backhaul link, and then connected to the core network and content server through the wired link.
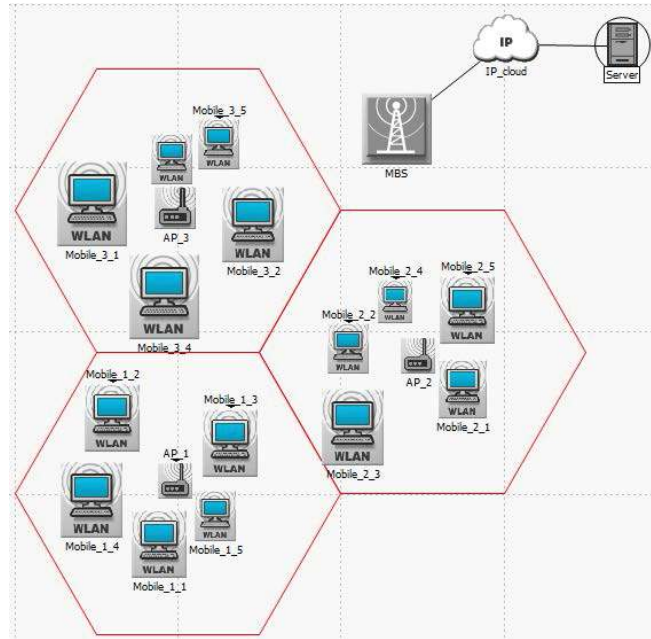

Fig. 3. Envisioned network architecture.

In the simulation scenario, all mobile terminal UEs need to obtain video contents from the server, which follows Pareto principle, the total number of files is 500 thousand. Each UE has a certain capacity of local cache, which can be set to 0.05/0.1/0.15/0.2/0.25/0.3% of the total number of files. UEs directly communicate with each other through WiFi Direct, adopting AODV protocol and forming ad-hoc network, thus realizing the communication mode of D2D. In this way, UE can obtain content from caches of neighboring users.

To realize the caching policy in wireless network, we build NDN layer over IP layer based on the architecture of content centric network, Fig. 4 shows the processor model of a CCN-enabled wlan workstation, from which we see that

NDN_device module is added overlaying on ip_encap to process NDN protocol. Other nodes has the similar module added to realize the caching.
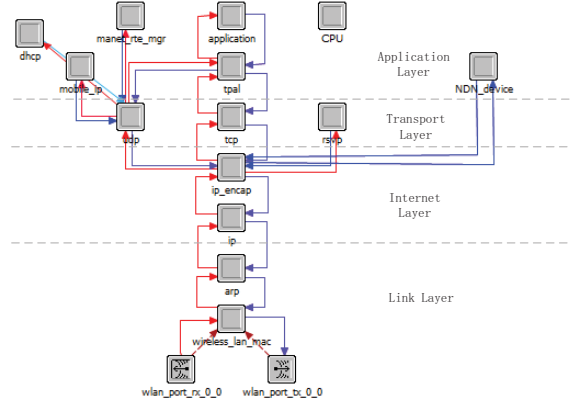

Fig. 4. Processor model of a CCN-enabled wlan workstation.

Table 1 lists the values of important parameters considered in the simulations. These values were selected to reflect real-world implementations of cellular network and that is considering prior research work. The simulations were run multiple times and the presented results are an average of these runs.

4.2 Simulation results

The simulation duration is set to 21000 seconds, this size of total files we used is 500000. The size of all the file is set to 1Mb, thus the total file size will be 500000Mb. Besides, the cache size is set to 100M, 300M, 500M, 800M and 1000M, 0.02%, 0.06%, 0.1%, 0.16% and 0.2% of total size respectively. All the mobile workstations send interest packet every 5 seconds, then the total number of interest packets CCN nodes received will be 5*3*21000/5=63000. The simulation results are obtained and analysed as follows.

(1) Effects of cache size on hitting rate.

In Fig. 5, we evaluate the performance of different caches on hitting rate. The relative cache size is set to 0.02%, 0.06%, 0.1%, 0.16% and 0.2% respectively, the final corresponding hitting rates are %, %, % and % respectively. The obtained simulation results show that increasing cache size can improve the hitting rate, but it is also obvious that the increment of the hitting rate is not linear to the cache volume. Besides, from the figure we can see that when the cache size increases to 0.16%, trying to enlarge the cache volume would not have big help on enhancing the hitting rate.

(2) Effects of caching scheme on hitting rate.

As shown is Fig. 6, we compared 3 different caching schemes, the well-known random and LRU schemes, and the scheme CC we proposed. In the simulation,
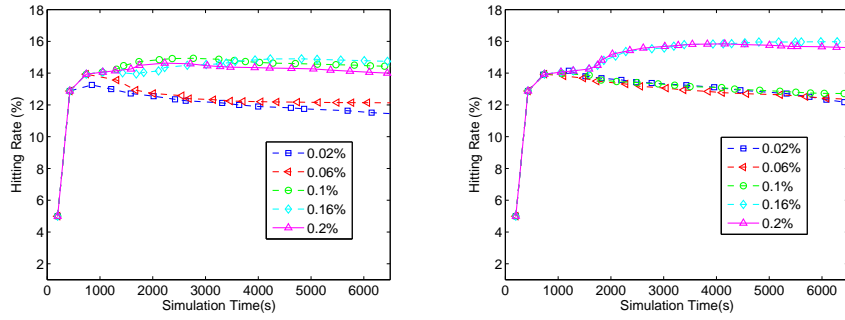
(a) Performance of random for cache sizes.

(b) Performance of CC for cache sizes.
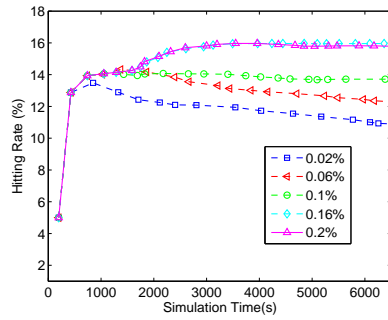
(c) Performance of LRU for cache sizes.

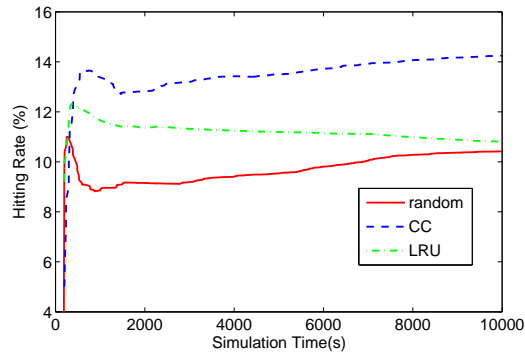**Fig. 5.** Performance of random, CC, LRU for varying relative cache sizes

Fig. 6. Multiple replacement policies with same cache size.

Table 1. Simulation parameters

Element	Attribute	Value
WAN	Link for server	1,000 BaseX
Server	CCN root	ccnx:// epic/video/
	Publish root Interval	100 seconds
	Number of files	500000 files
	File size	1 MB
	Packet size	1024 bits
UEs	CCN dir	ccnx:// epic/video/
	File popularity	Pareto principle
	Start time	100 + random(10) seconds
	Stop time	20000 seconds
	Interest packet Interval	5 + random(2) seconds
	Data packet Time Out	2 seconds
	Buffer size	0.05/0.1/0.15/0.2/0.25/0.3%
	Wireless interface	802.11g @ 54Mbps
	Routing protocol	Ad-hoc (AODV)
AP, MBS	Buffer size	0.05/0.1/0.15/0.2/0.25/0.3%
	Replacement policy	random/LRU/CC
	MBS PHY Profile	WirelessOFDMA 20MHz
	AP PHY Profile	802.11g @ 54Mbps

we set the caching scheme of mobile workstations in cell 1 as random caching scheme, and scheme of that in cell 2 as CC scheme, and LRU scheme for cell 3. During the simulation, we track the hitting rate of local cache in Mobile_1_5, Mobile_2_5, and Mobile_3_5 respectively. Fig. 6 shows the hitting rate varies with time by different scheme when caching size is set to 0.06% of total files, we can see from the figure that the hitting rate of Mobile_1_5 with random replacement scheme is the lowest, and Mobile_2_5 with CC replacement scheme has the highest hitting rate, while Mobile_3_5's hitting rate is in the middle, that is the rand of hitting rate is $CC > LRU > random$. The results demonstrates that the our cooperative caching scheme has better performance on hitting rate.

Fig. 7 shows the hit rate of various policies under different case size, it can be seen that the cache size has little impact on the hit rate, which enlightens that designing good replacement policy is more important than improving the cache capacity.

(3) Effects of caching scheme on server load.

If the cache policy has a higher hit rate, the less interest packets are forwarded to the server, resulting that the server's load will decrease. Fig. 8 shows the server load contrast under different caching policies, the obtained results well prove that the server load of CC scheme is lower than the others, the value is close to 15.4Mbps, while the server load of random scheme is 16.4Mbps.
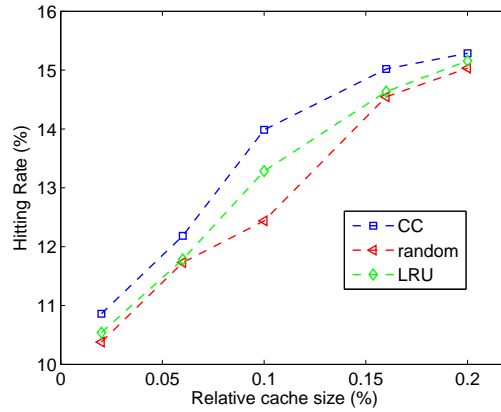
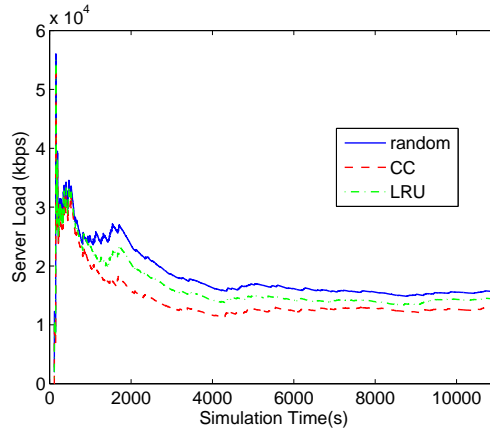**Fig. 7.** Final state of multiple replacement policies.

Fig. 8. Server load with same cache size.

5 Conclusion

Mobile edge caching and device-to-device technologies provide a promising solution for reducing the effects of demand growth by placing popular content files in proximity to uses. In this paper, to maximize the hit rate on local cache and cooperative devices' cache, we divide the cache into unique part and shared part to enriches the diversity of cached files in D2D networks so that enlarging the hit rate. The simulation results show that compared with two common replacement schemes, the cooperative caching CC in this paper achieves better performance. Besides, it reduces the content demands from MBS, which not only decreases the resource cost of MBS, but also mitigates the backhaul load and energy consumption. Moreover, we conduct the simulation with OPNET, reflecting the real features of small cellular network.

References

[1] C. H. Liu, A. Kind, and A. V. Vasilakos, "Sketching the data center network traffic," *IEEE Network*, vol. 27, no. 4, pp. 33–39, 2013.

[2] Y. Li and M. Chen, "Software-defined network function virtualization: A survey," *IEEE Access*, vol. 3, pp. 2542–2553, 2015.

[3] C. H. Liu, J. Fan, P. Hui, J. Wu, and K. K. Leung, "Toward qoi and energy efficiency in participatory crowdsourcing," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4684–4700, 2015.

[4] M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 176–189, 2016.

[5] K. Lin, J. Luo, L. Hu, M. S. Hossain, and A. Ghoneim, "Localization based on social big data analysis in the vehicular networks," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2016.

[6] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin, and S. Chen, "Vehicular fog computing: A viewpoint of vehicles as the infrastructures," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3860–3873, 2016.

[7] H. Wang, F. Xu, Y. Li, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," in *Internet Measurement Conference*, 2015, pp. 225–238.

[8] C. H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, 2011.

[9] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *Journal of Guilin University of Electronic Technology*, vol. 16, no. 4, pp. 1801–1819, 2013.

[10] C. H. Liu, B. Zhang, X. Su, J. Ma, W. Wang, and K. K. Leung, "Energy-aware participant selection for smartphone-enabled mobile crowd sensing," *IEEE Systems Journal*, vol. PP, no. 99, pp. 1–12, 2015.

[11] J. Wang, A. Liu, T. Yan, and Z. Zeng, "A resource allocation model based on double-sided combinational auctions for transparent computing," *Peer-to-Peer Networking and Applications*, no. 10, pp. 1–18, 2017.

[12] S. Liu and A. D. Striegel, "Exploring the potential in practice for opportunistic networks amongst smart mobile devices," in *International Conference on Mobile Computing & NETWORKING*, 2013, pp. 315–326.

[13] K. Lin, W. Wang, X. Wang, and W. Ji, "Qoe-driven spectrum assignment for 5g wireless networks using sdr," *Wireless Communications IEEE*, vol. 22, no. 6, pp. 48–55, 2015.

[14] F. Xu, Y. Li, M. Chen, and S. Chen, "Mobile cellular big data: linking cyberspace and the physical world with social ecology," *IEEE Network*, vol. 30, no. 3, pp. 6–12, 2016.

[15] X. Wang, Z. Sheng, S. Yang, and V. C. M. Leung, "Tag-assisted social-aware opportunistic device-to-device sharing for traffic offloading in mobile

social networks," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 60–67, 2016.

[16] H. Wang, X. Wang, K. Li, J. Ren, X. Zhang, and T. Jiang, "A measurement study of device-to-device sharing in mobile social networks based on spark," *Concurrency & Computation Practice & Experience*, no. 10, 2017.

[17] M. Chen, V. C. M. Leung, S. Mao, and T. Kwon, "Receiver-oriented load-balancing and reliable routing in wireless sensor networks," *Wireless Communications & Mobile Computing*, vol. 9, no. 3, pp. 405–416, 2009.

[18] M. Chen, M. Guizani, and M. Jo, "Mobile multimedia sensor networks: Architecture and routing," *Eurasip Journal on Wireless Communications & Networking*, vol. 2011, no. 1, p. 159, 2011.

[19] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, "Brain intelligence: Go beyond artificial intelligence," *Mobile Networks & Applications*, 2017.

[20] M. Chen, Y. Hao, M. Qiu, J. Song, W. Di, and H. Iztok, "Mobility-aware caching and computation offloading in 5g ultra-dense cellular networks:," *Sensors*, vol. 16, no. 7, p. 974, 2016.

[21] Y. Li, F. Zheng, M. Chen, and D. Jin, "A unified control and optimization framework for dynamical service chaining in software-defined nfv system," *Wireless Communications IEEE*, vol. 22, no. 6, pp. 15–23, 2015.

[22] F. Xu, Z. Tu, Y. Li, P. Zhang, X. Fu, and D. Jin, "Trajectory recovery from ash: User privacy is not preserved in aggregated mobility data," in *International Conference on World Wide Web*, 2017, pp. 1241–1250.

[23] M. Chen, "Opnet network simulation," *Press of Tsinghua University*, vol. 1, 2004.

[24] ——, "Ndnc-ban: supporting rich media healthcare services via named data networking in cloud-assisted wireless body area networks," *Information Sciences*, vol. 284, pp. 142–156, 2014.

[25] N. Choi, K. Guan, D. C. Kilper, and G. Atkinson, "In-network caching effect on optimal energy consumption in content-centric networking," in *IEEE International Conference on Communications*, 2012, pp. 2889–2894.

[26] J. Llorca, A. M. Tulino, K. Guan, J. Esteban, M. Varvello, N. Choi, and D. C. Kilper, "Dynamic in-network caching for energy efficient content delivery," in *IEEE INFOCOM*, 2013, pp. 245–249.

[27] J. Li, B. Liu, and H. Wu, "Energy-efficient in-network caching for content-centric networking," *IEEE Communications Letters*, vol. 17, no. 4, pp. 797–800, 2013.

[28] M. Gregori, J. Gómez-Vilardebó, J. Matamoros, and D. Gündüz, "Wireless content caching for small cell and d2d networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1222–1234, 2016.

[29] A. Liu and V. K. N. Lau, "Asymptotic scaling laws of wireless ad hoc network with physical layer caching," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1657–1664, 2015.

[30] M. Chen, L. T. Yang, T. Kwon, L. Zhou, and M. Jo, "Itinerary planning for energy-efficient agent communications in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3290–3299, 2011.

[31] D. Tian, J. Zhou, Z. Sheng, and V. C. M. Leung, "Robust energy-efficient mimo transmission for cognitive vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3845–3859, 2016.

[32] Y. X. Liu, A. Liu, S. Guo, Z. Li, Y. J. Choi, and H. Sekiya, "Context-aware collect data with energy efficient in cyber-physical cloud systems," *Future Generation Computer Systems*, 2017.

[33] D. Tian, J. Zhou, Z. Sheng, M. Chen, Q. Ni, and V. C. M. Leung, "Self-organized relay selection for cooperative transmission in vehicular ad-hoc networks," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2017.

[34] D. Tian, J. Zhou, Y. Wang, G. Zhang, and H. Xia, "An adaptive vehicular epidemic routing method based on attractor selection model," *Ad Hoc Networks*, vol. 36, no. P2, pp. 465–481, 2016.

[35] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leung, "Cache in the air: exploiting content caching and delivery techniques for 5g systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 131–139, 2014.

[36] C. Yang, Z. Chen, Y. Yao, B. Xia, and H. Liu, "Energy efficiency in wireless cooperative caching networks," in *IEEE International Conference on Communications*, 2014, pp. 4975–4980.

[37] Y. Hao, L. Peng, H. Lu, M. M. Hassan, and A. Alamri, "Energy harvesting based body area networks for smart health." *Sensors*, vol. 2017, no. 7, 2017.

[38] N. Golrezaei, K. Shanmugam, A. G. Dimakis, and A. F. Molisch, "Femto-caching: Wireless video content delivery through distributed caching helpers," in *IEEE INFOCOM*, 2013, pp. 1107–1115.

[39] K. Wang, Z. Chen, and H. Liu, "Push-based wireless converged networks for massive multimedia content delivery," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2894–2905, 2014.

[40] M. Cha, H. Kwak, P. Rodriguez, Y. Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *ACM SIGCOMM Conference on Internet Measurement*, 2007, pp. 1–14.

[41] K. Lin, J. Song, J. Luo, W. Ji, M. S. Hossain, and A. Ghoneim, "Gvt: Green video transmission in the mobile cloud networks," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 27, no. 1, pp. 159–169, 2017.

[42] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for iot localization in smart buildings," *IEEE Transactions on Automation Science & Engineering*, vol. 13, no. 3, pp. 1294–1307, 2016.

[43] K. Poularakis and L. Tassiulas, "On the complexity of optimal content placement in hierarchical caching networks," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2092–2103, 2016.

[44] "Ns-3 based named data networking (ndn) simulator." http://ndnsim.net/index.html.