ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A Survey on Low Latency towards 5G: RAN, Core Network and Caching Solutions

Udayagiri Mounisha
Department of Electronics & Communication Engineering
JNTUA College of Engineering,
Pulivendula, Andhra Pradesh 516390, India.

Smt.Shaik.Taj Mahaboob
Department of Electronics & Communication Engineering
JNTUA College of Engineering,
Pulivendula, Andhra Pradesh 516390, India.

ABSTRACT:

The demand for rich multimedia services over mobile networks has been soaring at a tremendous pace over recent years. However, due to the centralized architecture of current cellular networks, the wireless link capacity as well as the bandwidth of the radio access networks and the backhaul network cannot practically cope with the explosive growth in mobile traffic. Recently, we have observed the emergence of promising mobile content caching and delivery techniques, by which popular contents are cached in the intermediate servers (or middle boxes, gateways, or routers) so that demands from users for the same content can be accommodated easily without duplicate transmissions from remote servers; hence, redundant traffic can be significantly eliminated. In this article, we first study techniques related to caching in current mobile networks, and discuss potential techniques for caching in 5G mobile networks, including evolved packet core network caching and radio access network caching. A novel edge caching scheme based on the concept of content-centric networking or information-centric networking is proposed. Using trace-driven simulations, we evaluate the performance of the proposed scheme and validate the various advantages of the utilization of caching content in 5G mobile networks. Furthermore, we conclude the article by exploring new relevant opportunities and challenges.

INTRODUCTION

Along with recent advances in mobile communication technologies, an ever growing number of mobile users are continuously attracted to enjoy a wide plethora of multimedia services using smartphones and tablets [1]. While the demand for rich multimedia content has increased tremendously in recent years, the capacity of the wireless link , mobile radio network, mobile backhaul, and mobile core network cannot practically cope with the explosively growing mobile traffic due to the centralized nature of mobile network architectures [1]. Indeed, despite the continuous efforts of mobile network operators (MNOs) and network equipment vendors to enhance the wireless link bandwidth by adopting sophisticated techniques at both the physical (PHY) layer and medium access control (MAC) layers in Long Term Evolution (LTE) and LTE-Advanced systems, such as massive multiple-input multiple-output (MIMO), carrier aggregation, and coordinated multipoint (CoMP) transmission, the utilization efficiency of the radio spectrum is notably reaching its theoretical cap. Stemming from the observation regarding the traffic explosion problem as studied in [2–4], an important portion of mobile multimedia traffic is due to duplicate downloads of a few popular contents (e.g., popular music videos) with large sizes. Therefore, researchers and engineers have been investigating effective ways to reduce the duplicate content transmissions by adopting intelligent caching strategies inside the mobile networks, and enabling mobile users to access popular content from caches of nearby MNO gateways (e.g., using selective IP traffic offload techniques [5, 6]). From the perspective of Internet

Cite this article as: Udayagiri Mounisha & Smt.Shaik.Taj Mahaboob, "A Survey on Low Latency towards 5G: RAN, Core Network and Caching Solutions", International Journal & Magazine of Engineering, Technology, Management and Research, Volume 6 Issue 1, 2019, Page 116-124.

service providers (ISPs), this also helps reduce traffic exchanged inter- and intra-ISPs [3], not to mention significant reduction in the response time required to fetch a content file. Thus, the impact of Internet traffic dynamics on variation of the response latency can be eliminated. Furthermore, reducing traffic load via intelligent caching of popular content would enhance the energy efficiency of 4G networks contributing to the evolution of green 5G networks effectively.

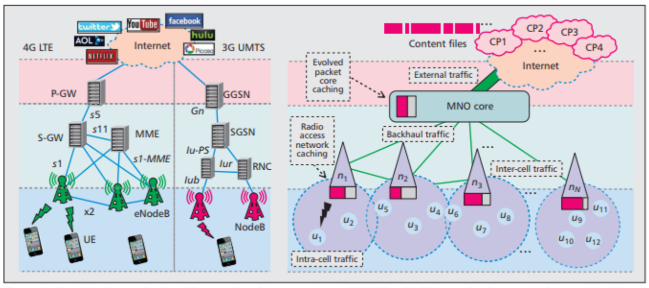


Figure 1. *Typical and generalized mobile network architectures: a) typical UMTS and LTE network infrastructure; b) general mobile network incorporating caching at different layers.*

Caching in 3G mobile networks [2] and caching in 4G LTE [4] networks have both been proven to be able to reduce mobile traffic by one third to two thirds. Furthermore, from related studies, it has become apparent that the technical key issues fall into the following three questions: Where to cache?: As illustrated in Fig. 1, due to the all-IP nature of current cellular networks, two places can be envisioned for deploying caches: •The evolved packet core (EPC), which consists of (among many other nodes) the serving gateway (S-GW), packet data network gateway (P-GW), and mobility management entity (MME) in LTE networks •The radio access network (RAN), which consists of NodeBs in 3G networks and evolved NodeBs (eNBs) in 4G LTE networks, as well as non-Third Generation Partnership Project (3GPP) accesses, such as Wi-Fi access points and worldwide interoperability for microwave access (WiMAX) base stations (BSs) It shall be noted that while in this article we mainly focus on 4G LTE networks as a representative example, and thus consider EPC for the core network, conclusions will be applicable to the core of 3G mobile networks consisting of (among many other nodes) the serving general packet radio service (GPRS) support node (SGSN) and the gateway GPRS support node (GGSN). In addition, throughout this article, we primarily refer to eNBs but without the intention of excluding NodeBs in 3G networks or BSs in WiMAX networks. The deployment of caches within EPC using conventional content delivery networking (CDN) techniques has been studied in [2, 3], while studies relevant to radio access network (RAN) caching have been made in [7, 8]. What to cache?: Caching aims to achieve a trade-off between the transmission bandwidth cost, which is usually expensive, especially for the inter-ISP traffic bandwidth, and the storage cost, which is becoming much cheaper. However, the scale of content acquired by content providers (CPs) is growing significantly and it is thus all but impossible to cache all content. It is hence important to decide what content to cache taking into account content popularity. As practically captured in [3, 4], only a small amount of popular content is accessed by a large portion of mobile user requests, while a long tail of contents remains unpopular [4]. In order to reduce outbound traffic or, say, inter-ISP traffic, it is of vital importance to improve the diversity of the cached content to ensure that most content can be fetched within the MNO’s network. On the other hand, to reduce intra-ISP traffic, mobile users’ requests for content may have to be locally satisfied by eNBs to which they respectively attach; neighboring eNBs do not necessarily have to cache similar contents since they can share and exchange contents (e.g., using interface X2 in an evolved packet system). It shall be noted that different file types also have different “cacheability,” as studied in [4]. For instance, among all the content types, images and videos have the highest revisit rate, and cacheable contents from Facebook have the highest probabilities of revisits. How to cache?: Caching policies, deciding what to cache and when to release caches, are crucial for overall caching performance. It is effectively important to estimate the gain behind a content by evaluating its current popularity, potential popularity, storage size, and locations of existing replicas over the network topology. Rather than adopting traditional caching policies, such as least recently used (LRU), least frequently used (LFU), and first-in first-out (FIFO), it is challenging to design cooperative caching policies for EPC and RAN caching to appropriately improve the cache hit ratio. In Fig. 2, we show four different scenarios wherein there is no caching, or caching is used within the mobile core network (i.e., in theevolved packet core, EPC, and/or the RAN). The figure illustrates how traffic is duplicated and can be reduced thanks to caching. We assume that 9 out of 12 mobile users (u1, u2, ..., and u12) are requesting the same content from CP via several eNBs in the MNO (n1, n2, ..., and nN). Green lines are for cables connecting EPC and RAN, while black lines represent the delivered copies of the requested content. The rectangle at the MNO core and eNBs represent the caching storage. Without any in-network caching, as shown in Fig. 2a, traffic is transmitted in a duplicative manner (with a high redundancy of nine copies). When caching at EPC is used, as illustrated Fig. 2b, the inter-ISP traffic can be significantly reduced to just one single copy, but the intra-ISP traffic from EPC to users via eNBs in the RAN remains the same (i.e., nine copies) with high redundancy. Furthermore, if caching is deployed at the RAN, as shown in Fig. 2c, the intra-ISP traffic can be reduced to four, and each eNB can locally satisfy the mobile users' requests, while the EPC caching storage can be reduced. If eNBs can exchange content with neighbor eNBs, the caching redundancy can be further reduced.

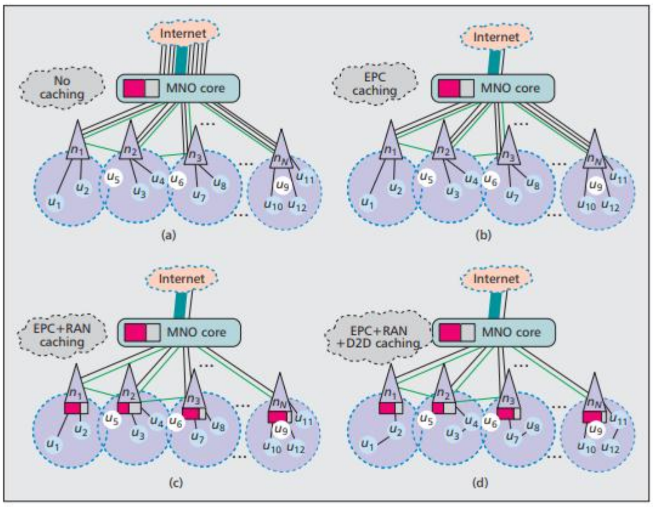


Figure 2. *Comparison of a) no caching; b) EPC caching; c) EPC+RAN caching; d) EPC+RAN+D2D caching.*

In this article, we investigate emerging caching techniques for mobile cellular networks, and explore potential research challenges and opportunities. This article is organized as follows. We discuss EPC and RAN caching techniques, respectively. CCN-based caching is further discussed. Related performance evaluation is shown. We detail new challenges and open issues. The article concludes.

CACHING WITHIN EPC

Both EPC caching and RAN caching can significantly reduce user-perceived latency as well as the transmission of redundant traffic over the network. Furthermore, caching at the edge of the network has the effect of smoothing traffic spikes and balancing the backhaul traffic over a long period of time. As shown in Fig. 1, 3G and 4G cellular networks, as well as the expected 5G networks, generally have a centralized architecture where all packets are concentrated within a small number of EPCs in a country. With the improvement of RAN technologies and wider deployments of cell towers, the backhaul links to those EPCs will suffer from huge traffic loads. Current widely deployed caching functions mostly take place within the EPC, that is, at the P-GW forming the so-called mobile content delivery network (mobile CDN) [3, 4]. Locating cache servers in a centralized fashion with EPC nodes at the premises (e.g., data center) of an MNO admittedly eases the management of both EPC and mobile CDN. This can also easily scale horizontally with increased demand for content as well as the diversity of content, and will improve the visibility of content dynamics and user characteristics so that caching at the EPC will enjoy a higher hit rate. However, downstream of the EPC (i.e., at the RAN), currently contents are transmitted via eNBs by GPRS Tunneling Protocol (GTP) with encapsulation, so it is technically easier to deploy content-aware caching at the EPC than at the RAN. By caching content at the mobile core network of 3G [2] and 4G [4] systems, one to two thirds of mobile traffic can be reduced. Currently, there are generally two kinds of caching techniques: • Web caching, which is object-oriented caching with content awareness,and consists of web caching based on uniform resource locator (URL) and prefix-based web caching • Redundancy elimination

(RE), which is protocol-independent and flow-oriented or packet-oriented, and consists of chunk-level RE, TCP-RE, and packet-level RE. These are described in details hereunder.

REDUNDANCY ELIMINATION

The RE technique is protocol-independent as it does not depend on the delivered service or application, but rather on the monitored chunks or packets. RE can remove duplicate byte strings from arbitrary network flows by deploying two middle-boxes on the link between PDN (e.g., Internet) and P-GW, and on the link between SGW and RAN. The first middle-box divides incoming content into small chunks or packets, and verifies that there are existing cached chunks with the same size and hash; it then sends a stream of chunk hashes to the second middlebox, which reconstructs the original content with those hashes and subsequently delivers it to mobile users. RE needs to be fast, adaptive, and parsimonious in memory usage in order to opportunistically leverage the caching resource. There are three popular types of RE techniques: • Chunk-level RE • TCP-level RE • Packet-level RE, depending on the scale of the penetration of the traffic flows.

Chunk-Level RE —In chunk-based caching, all files are split into chunks, while each chunk is hashed by the caching servers. If the hash and the size of two chunks are the same, they are considered as duplicates, and future requests will go to the same cached chunk. Chunk-level RE solves issues relevant to aliasing, uncacheable content, and content update. However, it lacks awareness of content and does not facilitate cooperative caching strategies for global optimization.

TCP-Level RE — Due to the wide dominance of HTTP-based traffic (above TCP/IP), we hereby also discuss TCP-level RE, which is similar in spirit to chunk-based caching. TCP-RE works over the content of each TCP flow. The caching middle-boxes reassemble the segments in the same TCP flow and divide them into smaller chunks (either fixed-sized or variable-sized) that serve as the unit of caching. While TCP-level RE has additional overhead to manage TCP flows, it allows scalable cache management as well as flow-based caching policies.

Packet-Level RE — One popular caching approach is packet-level RE, which detects candidate fragments for RE within each IP packet on smaller chunk sizes (e.g., 32–64 bytes), typically using content-based fingerprinting functions. By packet caching middle-boxes, the upstream middle-box removes redundant bytes, and the downstream middle-box reconstructs full packets. However, there are a few drawbacks with packet-level RE in high-speed cellular backhaul networks: • Small chunks can easily explode the index sizes in a high-speed network. • A small chunk size would incur a higher hashing and content reconstruction cost, stressing the memory system, which typically becomes the performance bottleneck in memory-based RE. MNOs may want to apply RE to only large file downloads or specific content types that exhibit high redundancy with a small management overhead (e.g., multimedia content like videos, images, and content from famous social network services such as Facebook). Based on [3], bypassing any HTTP responses smaller than 32 kbytes, the cost of managing 90.6 percent of the flows could be eliminated while still ensuring 31.6 percent bandwidth savings.

CACHING AT RAN

Moving application processing resources toward the network edge closer to mobile users will make it possible to simultaneously reduce network traffic and improve quality of experience. Moving application processing resources toward the network edge, closer to mobile users, will make it possible to simultaneously reduce network traffic and improve quality of experience. This can further optimize the most expensive part of the network operating cost of the various fiber leased lines that connect eNBs to EPC. CHEN_LAYOUT_Layout 1/30/14 1:03 PM Page 134 IEEE Communications Magazine • February 2014 135 This can further optimize the most expensive part of the network operating cost of the various fiber leased lines that connect eNBs to EPC. The research work conducted in [7] focuses on caching popular video clips in the RAN. The “FemtoCaching” concept in [8] also discusses and evaluates the video content delivery through distributed “caching helpers” in femtocell networks. However, most RAN caching solutions face the same implementation issues: eNBs establish “tunnels” between users and the EPC, while content files are packetized and then encapsulated by GTP tunneling, making it difficult to carry out object-oriented or contentaware caching. To overcome this shortcoming, the natural evolution of caching is to cache repetitive portions of an object, known as byte caching [9]. Byte caching is a protocol-, port-, and IPaddress-independent bidirectional caching technology that functions at the network layer by looking for common sequences of data in the bytes of packet flows. Byte caching represents an enhanced fine-grained approach similar to the packet-level RE scheme. It does not need to split flows into segments, but continuously penetrates into the byte strings to cache the often used bytes and eliminate any redundant ones. For example, if an enterprise logo is used on all documents, byte caching will identify this common part of bytes on the different files and prevent that data from being transmitted redundantly. Unlike centralized EPC caching, one important issue of RAN caching consists of the fact that the caching space at each eNB is practically small, and the number of users served by individual eNBs is usually small (i.e., unless in highly dense areas such as Times Square in New York City), resulting in low-to-moderate hit ratios. Therefore, intelligent caching resource allocation strategies and cooperative caching policies among (neighboring) eNBs is mandatory for efficient RAN caching.

CCN-BASED CACHING

An important requirement behind 4G has been the all-IP feature of its architecture. Incorporating CCN techniques could become an important feature of 5G mobile networks. As per the need for decentralizing mobile CDN services, CDNs are getting distributed further, incorporating information-centric and content-aware caching techniques, and forming the so-called information-centric (ICN) or content-centric networking (CCN) architecture [10] for the future Internet. The primary goal of CCN is to facilitate in-network data storage for universal caching in every network node. Major CCN designs have the following common attributes: • Receiver-oriented and chunk-based transport • In-network per-chunk caching • Name-based forwarding • Uniquely identifiable content naming For further details on CCN, the interested reader is referred to [10, 11]. In CCN, a user requests a particular content by issuing interest packets to neighbors. If the requested content can be retrieved in the caching store (CS) of any device, the content is directly delivered from the device. Otherwise, routers propagate interest towards appropriate content sources and store information for each forwarded interest in the pending interest table (PIT). Routers and other network nodes will push the cached or fetched content toward their requester(s) based on the information stored in their respective PITs and also on a strategic caching policy. In mobile networks, there are many routers within the EPC and RAN whereby CCN-based caching can be enabled. 5G mobile networks are expected to include CCN-capable gateways, routers, and eNBs. For example, the adaptive mobile video streaming with offloading and sharing in wireless named data networking (AMVS-NDN) framework [11] has realized a CCN architecture in a commercialized WiMAX BS implementing CCN-based caching and achieving significant traffic reduction. CCN-based caching will be universal and pervasive in future mobile networks. Due to the wide distribution of caching resources, cooperative caching policies should carefully consider content popularity, freshness, diversity, and replica locations over the network topology in order to achieve the following objectives:

Minimization of inter-ISP traffic (outbound traffic): This can be guaranteed when the cached content in all storage has the highest diversity so that content can be fetched within the same ISP (supporting the same EPC and RAN) as much as possible. That is, for any content, only one copy is stored, and content is cached according to popularity until the whole storage of caches within the EPC and RAN is full.

Minimization of intra-ISP traffic (traffic within the EPC and RAN): This can be guaranteed when most popular contents are cached at each eNB, so most requests can be locally satisfied without many exchanges among eNBs. This objective may induce situations when many eNBs cache the same popular content, so it is somehow contradictory to the requirement of content diversity in the aforementioned objective.

Minimization of content access delay of all users: Users may fetch content from local eNBs, routers in the RAN and EPC, and even from remote CP servers with different delays. To optimize the quality of service (QoS) of all users, it is important to consider effective caching policy (i.e., placing an object oi in a caching eNB node, RAN node, and/or EPC node) to minimize the total delay (Eq. 1), where oi denotes a content, pi is the popularity (request frequency) of oi , and teNB, tRAN, tEPC, and tCP represent the transmission delays of the corresponding part of the network. There should be a balance among caching choices at the local eNBs, RAN, and EPC with respect to content popularity. Good cooperative caching policies are still challenging.

Open Issues And Challenges Distributed Cache Resource Management And Cooperative Caching Policy

An important limitation of RAN caching conlies in the fact that the caching space at individual eNBs is usually small and the user base is quite limited, which will potentially result in low hit ratios. Generally speaking, a RAN consists of thousands of eNBs interconnected via high-capacity links. A well tuned inter-eNB cooperative caching policy is therefore needed that considers the dynamics of local user demands, caching status of neighboring eNBs, and global optimization of caching resource utilization as well as the maximization of QoS of all users. There is practically a trade-off between the diversity of content stored inside EPC and RAN and the redundancy of the replicas of popular content in eNBs. Figure 4 shows a typical scenario with a cooperative caching policy in the EPC, RAN, and intermediate routers, wherein some popular content are stored at local eNBs, and some are cached in the EPC, while in a small area with several neighboring eNBs, one content does not necessarily need to be cached with multiple copies. In order to achieve efficient cooperation among nodes (devices) in the EPC and RAN, the caching status should be effectively shared among nodes (devices), most importantly incurring only a minimal signaling overhead. Additionally, as there are always numerous new content coming, the underlying caching policy should also consider the dynamic changes of the content popularity and user demands so that adequate caching decisions are made online and in real time. Another important issues of RAN caching pertains to service continuity (i.e., upon a handoff operation, termination of content transmission from source eNB, and prefetching of content in the target attaching eNB). Mobility-aware caching and prefetching in the RAN are highly needed. In this vein, solutions similar to the Follow Me Cloud concept can be envisioned.

REALISTIC RAN CACHING IMPLEMENTATION

In spite of the above mentioned challenges, there are already a few RAN products with caching support. Notable examples are BlueCoat ProxySG appliances, CacheMARA's caching proxy techniques, Altobridge's "Dataat-the-Edge" solution, Saguna Networks' Content Optimization Delivery System with caching in small cells, and Nokia Siemens Networks Radio Application Cloud Server (RACS) for liquid application service for BSs. These existing RAN caching techniques are mainly based on packet-level RE or byte caching, which are not content-aware at all. They thus may suffer from high implementation complexity, as well as a scalability problem. There is already one study for CCN-based content-aware RAN caching in the AMVS-NDN framework [11]. In this study, the CCN architecture is imported into a WiMAXfemtocell station, while CCNbased caching for LTE eNBs is still under realization.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

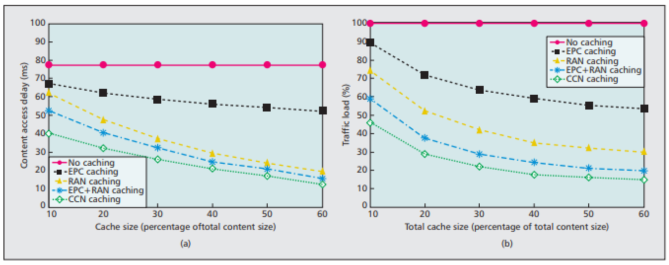
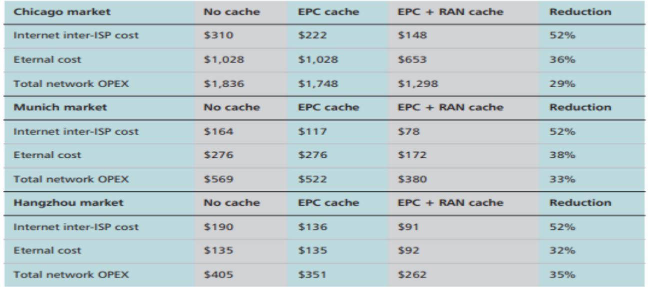


Figure 3. *Performance evaluation of EPC caching, RAN caching, and CCN caching: a) user content access delay; b) number of content transmissions.*

The 3GPP specifications relevant to local IP access (LIPA) and selective IP traffic offload (SIPTO) via a home eNB [15] enable 3G/4G small cells (femtocells) to directly connect to the Internet. Hence, caching at eNBs in the RAN will become like caching at the distributed and collocated EPC (since local PDN GWs will be a part of EPC). An effective caching solution for small cells deploying LIPA and SIPTO [8] may be required, especially for coordinating among neighboring small cells, and proactively and predictively caching content for home users regarding the continuity and optimality of the content delivery may become a promising research topic, which may bring further technological advances toward the realization of next generation content-aware 5G mobile networks.



Chicago market	No cache	EPC cache	EPC + RAN cache	Reduction
Internet inter-ISP cost	$310	$222	$148	52%
Eternal cost	$1,028	$1,028	$653	36%
Total network OPEX	$1,836	$1,748	$1,298	29%
Munich market	**No cache**	**EPC cache**	**EPC + RAN cache**	**Reduction**
Internet inter-ISP cost	$164	$117	$78	52%
Eternal cost	$276	$276	$172	38%
Total network OPEX	$569	$522	$380	33%
Hangzhou market	**No cache**	**EPC cache**	**EPC + RAN cache**	**Reduction**
Internet inter-ISP cost	$190	$136	$91	52%
Eternal cost	$135	$135	$92	32%
Total network OPEX	$405	$351	$262	35%

Table 1. *Ten-year network OPEX estimation [13]; unit: $1 million.*

INTEGRATION WITH VIRTUALIZATION OF CELLULAR NETWORKING DEVICES

The potential use of software-defined networking (SDN) [16] in 4G mobile networks as well as future 5G networks has been gaining lots of momentum recently. In 4G mobile networks, the control and data planes are separated, but with use of the emerging network function virtualization (NFV) technique [17], networking nodes become controllable, programmable, and, most important, scalable and elastic for resource utilization adapting to user and content dynamics. This may handle the capacity needs of each station and ramp up or scale down node capability depending on the mobile traffic dynamics. SDN is mostly focusing on the adaptability and controllability of network functions using virtualized resource management systems. SDN provides a fine-grained packet classifier and flexible routing, which can easily direct a chosen subset of traffic through a set of middle-boxes. Content-aware forwarding and caching may then become easier, particularly with the support of deep-packet inspection (DPI) technologies. New caching and delivery mechanisms for an SDNbased 5G mobile network environment become more critical and important, but are still pending.

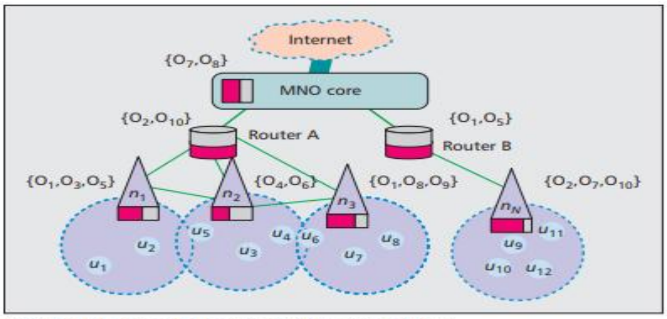


Figure 4. *An example of a cooperative caching policy.*

CACHING WITH MULTICAST AND DEVICE-TO-DEVICE

Although caching techniques can reduce traffic within a mobile network, this does not impact the load of conveying the content via the wireless link. For effective delivery of the same content to a group of mobile users, the evolved multimedia broadcast multicast service (eMBMS) of LTE may be envisioned, integrated along with necessary caching techniques in the EPC and RAN. Furthermore, the wireless link load can be further reduced with the use of local caching and sharing techniques such as device-to-device (D2D), a hot topic currently under investigation within 3GPP 4G LTEAdvanced. Using D2D, mobile users can use operator authorized spectrum for direct communication without the support of infrastructure, so users can cache content locally in their own devices and share them with friends directly . As shown in Fig. 2d, caching at the RAN along with caching and sharing among mobile users via D2D can significantly reduce the traffice load over the cellular links from eNBs to users. However, it is quite challenging to design efficient caching and sharing strategies in the RAN (eNBs) and user devices. A potential research direction for offloading 5G traffic may be how to effectively exploit user social relationships to facilitate caching and sharing activities by D2D communication.

CONCLUSIONS:

It is generally agreed that current RAN deployments and mobile backhaul networks cannot cope with the ever growing demand of mobile users for rich multimedia services. Caching can be a potential solution. In this article, we have provided an overview of some emerging caching techniques for current 3G, 4G, and future 5G networks, comparing between EPC caching, RAN caching, and CCN-based caching. Based on trace-driven evaluation and related analysis, we have demonstrated that the deployment of in-network caching into mobile networks can potentially help reduce mobile traffic. We have also presented a number of promising research opportunities and relevant challenges, particularly related to distributed cache resource management, cooperative caching policy, and content-aware RAN caching. Conclusively, we have highlighted the roles that NFV, eMBMS, and D2D can play in further improving the gains that can be acquired from caching in mobile networks.

REFERENCES

[1] X. Wang et al., "A Survey of Green Mobile Networks: Opportunities and Challenges," ACM/Springer Mobile Networks and Applications (MONET), vol. 17, no. 1, Feb. 2012, pp. 4–20.

[2] J. Erman et al., "To Cache or Not to Cache — The 3G Case," IEEE Internet Computing, vol. 15, no. 2, Mar. 2011, pp. 27–34.

[3] S. Woo et al., "Comparison of Caching Strategies in Modern Cellular Backhaul Networks," ACM MobiSys, June, 2013.

[4] B. A. Ramanan et al., "Cacheability Analysis of HTTP traffic in an Operational LTE Network," Wireless Telecommun. Symp., Apr. 2013.

[5] K. Samdanis, T. Taleb, and S. Schmid, "Traffic Offload Enhancements for eUTRAN," IEEE Commun. Surveys & Tutorials., vol. 14, no. 3, 3rd qtr. 2012, pp. 884–96.

[6] T. Taleb, Y. Hadjadj-Aoul, and S. Schmid, "Geographical Location and Load Based Gateway Selection for Optimal Traffic Offload in Mobile Networks," IFIP Networking, Valencia, Spain, May 2011.

[7] H. Ahlehage and S. Dey, "Video Caching in Radio Access Network: Impact on Delay and Capacity," IEEE WCNC, Shanghai, China, Apr. 2013.

[8] N. Golrezaei et al., "FemtoCaching: Wireless Video Content Delivery through Distributed Caching Helpers," IEEE INFOCOM, Mar. 2012.

[9] F. Le, M. Srivatsa, and A. K. Iyengar, "Byte Caching in Wireless Networks," IEEE ICDCS, June 2012.

[10] B. Ahlgren et al., "A Survey of Information-Centric Networking," IEEE Commun. Mag., vol. 50, no. 7, July 2012, pp. 26–36.

Author Details

Udayagiri Mounisha studied B.Tech from Kandula Obul Reddy Memorial College Of Engineering, Kadapa, Andhra Pradesh. And now pursuing Master of Technology in Digital Electronics and communication Systems from JNTUA College of Engineering, Pulivendula, Andhra Pradesh.

Smt.Shaik.Taj Mahaboob is currently working as Assistant Professor of ECE, JNTUA College of Engineering Pulivendula, Andhra Pradesh. Her main research work focuses on Image Processing and Embedded Systems. She has 14 years of teaching experience and 8 years of research experience.