2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS): PERFECTO 2017: Performance and Fault Management in Cellular Networks

Mobile Core Network Redimensioning for Efficient Resource Utilization

Eyuphan Bulut
Department of Computer Science
Virginia Commonwealth University
401 West Main St. Richmond, VA 23284
ebulut@vcu.edu

Abhimanyu Gosain
Raytheon BBN Technologies
10 Moulton Street
Cambridge MA 02138
agosain@bbn.com

***Abstract*—In response to the mobile data tsunami, the research community has studied WiFi offloading based efforts to provide an immediate solution to mitigate the load on both the cellular space and mobile core network. However, this has caused mobile network operators (MNO) to lose control and visibility on the user data since the traffic is usually directed towards the Internet bypassing the operator's core network. Therefore, to offload the traffic from the cellular space but still have the control and visibility and continue offering operator services (i.e., VoLTE) even over WiFi, operators have extended their core network by adding new gateways (e.g. ePDG) that integrate the trusted and untrusted WiFi offloaded traffic. On the other hand, such an extension have caused the core network data load to remain the same while requiring a different solution space. In this paper, we approach the challenges triggered by mobile data explosion from core network gateways' perspective whose load is not alleviated even with WiFi offloading. We first propose the formation of Social-Offloading-Groups (SOGs) to offload the core network in terms of the cellular connection count (i.e., session) while maintaining all regular activities of SOG members over the SOG lead. Secondly, we provide a collaborative core network architecture between geographically distributed gateways to balance their operations and reduce the utilization of core network resources.**

***Index Terms*—Mobile core network, data offloading, resource utilization efficiency, cooperative networks.**

I. Introduction

The wide-spread proliferation of mobile devices has triggered a tremendous increase in demand for the mobile data usage. The number of mobile users has already reached over 50% of world's population [1], [2]. Mobile phone users check their devices 110-150 times a day and spend 60-90 minutes depending on device type. A recent report [3] shows mobile Internet traffic will increase eight fold between 2015 and 2020 with a compound annual growth rate (CAGR) of 53 percent.

Service providers (i.e., operators) look for affordable ways of handling this huge growth without losing their subscribers and even converting it to a monetizing opportunity. These solutions include but are not limited to [4]: (i) upgrading the radio access network (RAN) to Long Term Evaluation (LTE) or LTE-A; (ii) improving the infrastructure by adding more radio towers (i.e. macrocells) or (iii) small-scale base stations such as femtocells, (iv) applying tiered data plans for taming data usage, and (v) utilizing WiFi access points (AP) for offloading cellular network traffic [5].

Among the aforementioned solutions, WiFi offloading is considered a more viable, economic and immediate remedy, and thus, has attracted a great attention from both academic and industrial research community. Direct offloading to the fixed WiFi access point (AP) or indirect offload after hop by hop transfer of data between multiple devices are commonly used techniques. There is also some work done on the efficient deployment of new APs [7] or drop zones [23] and recruitment of third-party APs via auctions [11].

Even though WiFi offloading techniques have provided some relief on cellular load, they are mainly designed for unmanaged offloading [4], which allows mobile data go to Internet bypassing the core network. However, this caused operators to lose visibility and control over user data and prevented users from accessing services provided by operators (i.e., VoLTE). Thus, operators are adding new gateways (e.g., ePDG) to their core networks that provide seamless integration of their core network with the new access modules (e.g., WiFi) while still providing visibility on the user data. Operators not only can authorize their subscribers connected from WLAN, let them use their services and bill them for their usage but they can also continue analyzing user behavior and updating their network and services depending on the user needs. This *managed offloading* process, however, does not change the amount of traffic passing through the core network gateways. Thus, the benefit of WiFi offloading stays limited to cellular space.

In this paper, we present solutions to increase resource usage efficiency in mobile core network. Specifically, we target low utilization of user sessions and waste in the number of sessions and propose to merge sessions of close users. Moreover, we propose a collaborative core network architecture to reduce the variations in gateway resource utilization.

The rest of the paper is organized as follows. We provide the motivation on the ideas in Section II. In Section III, we discuss the proposed solutions for efficient resource utilization. We discuss the related work in Section IV and propose future evaluation work in Section V. The conclusion is provided in Section VI.

978-1-5386-2784-6/17/$31.00 ©2017 IEEE

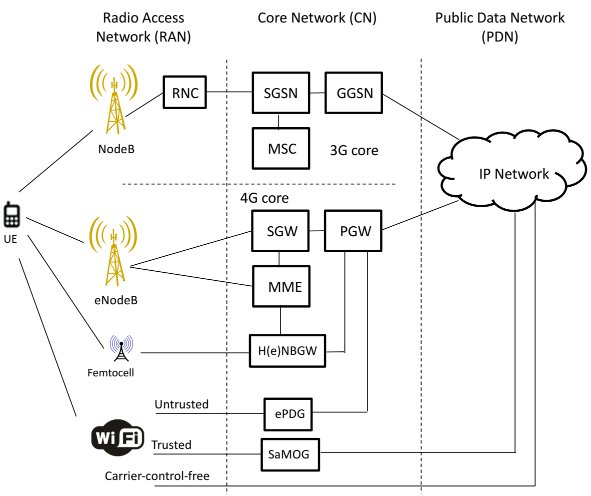

Fig. 1: Mobile core network and offloading architecture.

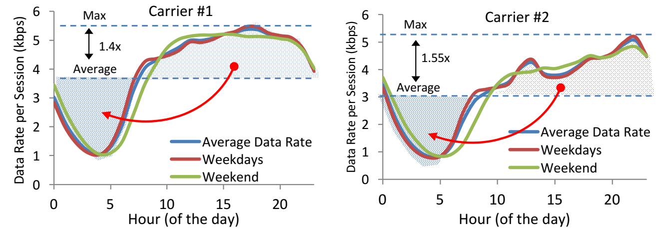

Fig. 2: Daily data usage trending is predictable and diurnal [6]. Degrading the utilization variation help increase capacity.

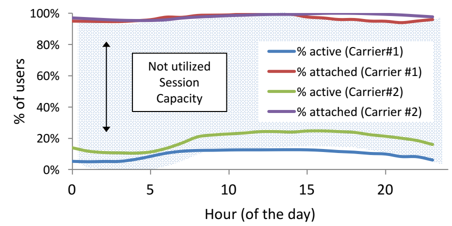

Fig. 3: Session capacity of core network gateways are under-utilized. Only 10-20% of sessions send/transmit data actively.

II. Motivation

Figure 1 illustrates a simple view of the mobile core network architecture with offloading gateways. If the UE connects to a femtocell, its traffic is forwarded to a H(e)NBGW (Home (e)NodeB Gateway) which integrates traffic from many femtocells to the existing core. If the UE connects to WLAN, different core gateways can play a role depending on access type and carrier's involvement. If the user does not use any service (voice, apps etc.) provided by carrier, the traffic can go to IP network without any interruption by the carrier. Otherwise, the carrier integrates the traffic to its existing core architecture through ePDG (Evolved Packet Data Gateway) and SaMOG (S2a Mobility based on GTP) gateways. If the WiFi AP is untrusted (i.e. not controlled by carrier), ePDG secures the data transmission from the UE to core network via IPSec tunnels. If the WiFi AP is trusted, SaMOG handles the interworking between the UE and core network.

Previous work [6] on the analysis of aggregate mobile data usage (traffic on core network gateways (i.e., PGW)) shows that (Figure 2) general data usage trending is predictable (i.e., subscribers of different carriers have similar usage graphs) and diurnal [13]–[15]. There are some remarkable differences due to user behavior changes (i.e. late nights during weekend, lunch time during weekday) at weekends and weekdays. Moreover, special events and holidays can cause deviations from general usage. The operators should make the capacity of their gateways ready for maximum expected usage. However, due to the uneven data usage distribution during a day, resources (i.e., CPU) on gateways are not utilized efficiently. For example, around 40% saving could be obtained in Figure 2 if data usage could be balanced and trending is flattened.

Moreover, even at peak usage times, users are not actively using the sessions they are allocated. One of the significant property that defines the volume of traffic on a gateway is the percentage of sessions that are actively transmitting/receiving data. Even though a session is allocated to a user device on the gateway in the core network, it is not actively used most of the time. Moreover, when there is lack of activity in a User Equipment (UE), it switches to idle state after some time to save energy. This results in release of the channel and deletion of bearers between SGW and MME. However, PGW and MME still keep the information about the bearers and consume corresponding memory resources at these gateways. UE returns back to connected or active mode when the UE responds to a paging request or it has some signaling data to send such as Packet Data Network (PDN) connectivity request. During this process, all of the required S1 interface connections for all of the established EPS bearers are reestablished. The key point here is PGW and MME resources do not benefit from active-idle transition. The only way the sessions on these gateways will decrease is by turning off these phones.

Figure 3 shows the percentage of sessions connected on a PGW (of two tier-1 MNOs in US) during a day and the percentage of these sessions that send at least a byte of traffic in the last minute. User sessions stay connected to the main gateway in the core network even at nights (since most are still turned on), thus, they still allocate *memory resources* at PGW which usually can support a maximum fixed session count. Moreover, the user activity (10-20%) is very limited on each session even at busy hours. For example, as an intuitive approach, five of these sessions could be merged and a higher activity (50-100%) could be achieved on each session while the memory consumption at core network could be reduced to 20% of current usage.

III. Proposed Solutions

We present two different schemes to achieve efficient resource utilization and increase capacity in mobile core network gateways. In the first scheme, we aim to decrease cellular connection (i.e., session) count at a core network gateway (i.e., PGW) without decreasing the subscriber count served and without interrupting any user operation. This will then result in decreased memory usage and open room for new subscribers, yielding larger capacity at the gateway. In the second scheme, we want to achieve effective and balanced utilization of processing resources (i.e., CPU) by changing diurnal aggregate user traffic on a gateway to a more balanced load with a reduced maximum usage during a day.

A. Session Aggregation through Social Offloading Groups (SOG)

In the first scheme, we look at the feasibility and implementation challenges of Social Offloading Groups (SOG) which will group several mobile users based on their social relations and introduce them to the rest of the network as a single user and session as a representative of all. As Figure 4 illustrates, there will be a SOG lead which will provide the cellular connection (i.e. session) towards the backhaul. The SOG lead will be rotating among the members of SOG considering the battery and efficiency of the UE. The intra-communication within SOG members will be achieved via D2D communication. Each SOG member will send/receive its data towards/from the SOG lead via a non-delayed (which could be over multiple nodes) or minimally delayed (opportunistic) links.

The scenario we consider for the formation of SOGs is a trusted network like a home or work environment where members know and trust each other. Considering that more than 80% of mobile data is produced indoor [3] today (via potentially mobile but non-moving devices), this assumption covers most of the realistic scenarios in practice. This trusted environment also abolishes a possible competition among the users for accessing the shared resources (spectrum). However, there are challenges including: (i) the analysis of mobile data usage trends for each user and understanding its correlation between the user's interactions with other users; (ii) on-the-fly and efficient formation of SOGs, its lead selection and rotation; and (iii) handling the relations of the UEs in a SOG with core network to continue their complete similar functioning in the SOG as in regular case.

The proposed scheme is different than tethering and its variants [25] in which there is still a cellular connection for users served by a main user with a data plan and they consume resources in core network gateways. Moreover, tethering approaches are on demand and controlled by users. This scheme proposes a self managed system where both the data calls and voice calls are accomplished over the SOG lead while there is no cellular connection maintained by the other members of the SOG. Moreover, the focus of the proposal is different than the clusterhead-based topology formation defined in 3GPP Proximity Services (ProSe) [26], which looks at control and scheduling of D2D transmission by a master UE which serves like an eNodeB for other nodes in its range. This proposed scheme aims to benefit from co-localization of subscribers so that they can connect to mobile backhaul over one of the users in the vicinity. This may not only help MNOs to save resources in their core network but also MNOs can offer reduced prices to subscribers for using their service as they will be sharing their sessions with others.

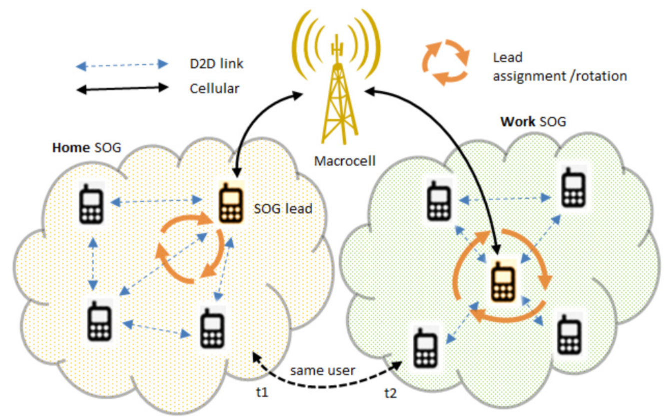

Fig. 4: Social Offloading Group (SOG) architecture. Socially close users (e.g., family members) form SOGs and communicate to the core network via single cellular connection over SOG lead which is selected and rotates based on energy consumption and data activity concerns. SOG members exchange data via D2D links benefiting from caching. The same user can be part of multiple SOGs at different times.

In order to build the prototype of the proposed system, with minimal or no change to call flows of cellular system which is already complicated, multiple ways can be considered for the treatment of SOG in the core network. All phone calls for any member of the SOG need to be received by the SOG lead, then it will be directed to the right member in the SOG. From core network perspective, registration and paging of all SOG member UEs should be managed over the SOG lead. Possible prototypes could be built using dynamic SIM (Subscriber Identity Module) allocation [17], multi-SIM card [18] and PhoneLets [31], which is a system that allows a single mobile subscription to be shared between multiple nearby devices. Moreover, to allow subscribers the options of individual connection and connection with SOGs, hybrid solutions could be utilized (e.g., by turning off/on the individual connection and starting/leaving to use PhoneLet).

Since the connection to core network will be achieved by the lead of SOG, the selection of lead and rotation of it among members of SOG is significant for energy consumption. This can be determined based on the user relations [8], [9], [37] and user activity. To determine the SOG lead, several parameters can be considered. There will be a cellular energy consumption at the lead node for the transmission of data of all SOG members. Moreover, there will be energy consumption for the transport of the data from other members of SOG to the lead. For example, 3G cellular power consumption of a node is modeled by three states: idle, transmission and tail. Highest power is consumed at transmission (DCH) state while producing the lowest delay and highest throughput. A smartphone switches to idle state to save energy but it also waits in tail (FACH) state for some time in case new data transmission requests come in short time so that an immediate switch to DCH state can be achieved and promotion cost is prevented. Some recent work [10], [34] look at the minimization of tail energy by scheduling the tasks and transferring them in bulks. Our goal is to find out the lead of a SOG which can minimize the total energy consumption in the SOG.

Consider a set of N users. Each user $j \in N$ has I_j tasks (i.e. data to be transmitted). Task i of node j is denoted as T_j^i and D_j^i corresponds to the size of the task. The energy that node j consumes to complete the task i ($\mathcal{E}_j(T_j^i)$) via direct cellular connection can then be calculated as:

$$\mathcal{E}_j(T_j^i) = \frac{D_j^i}{\beta(j)} P(j) + \mathcal{E}_{pro} + \mathcal{E}_{tail}$$

If the node sends its data to a neighbor node k first via a local (i.e., D2D) connection then it is sent to core network from the neighbor node, the energy consumption becomes:

$$\mathcal{E}_j^k(T_j^i) = \frac{D_j^i}{\beta(j,k)} P(j,k) + \frac{D_j^i(1 - p(T_j^i, k))}{\beta(k)} P(k) + \mathcal{E}_{pro} + \mathcal{E}_{tail}$$

Here, $\beta(j)$ denotes the cellular bandwidth of node j while $\beta(j,k)$ denotes the local bandwidth between node j and k. $P(j)$ is the power consumption rate for cellular communication of node j and $P(j,k)$ is the power consumption rate for D2D communication from node j to node k. We also denote the probability of the similarity of the data in the task with any cached data at node k with $p(T_j^i, k)$. If there is a match found, the data will not be transmitted over cellular radio, thus, extra saving will occur.

The optimization function, which gives the SOG lead that minimize the total cost in the SOG, is:

$$k \leftarrow \arg\min \left\{ \sum_{j \in N - \{k\}} \left(\sum_{i \in I_j} \mathcal{E}_j^k(T_i) \right) + \sum_{i \in I_k} \mathcal{E}_k(T_i) \right\}$$

The key parameters here are $p(T_j^i, k)$, $\mathcal{E}_{pro}$ and $\mathcal{E}_{tail}$. $p(T_j^i, k)$ can be obtained by analyzing the data similarity of users who are socially close (or within the same SOG). $\mathcal{E}_{pro}$ is the energy consumed in promotion from idle to DCH state and $\mathcal{E}_{tail}$ is the energy consumption during tail state. To decrease the overall energy cost spent in a SOG for $\mathcal{E}_{pro}$ and $\mathcal{E}_{tail}$, overlap of user activity of SOG members that will give better scheduling of tasks can be considered in SOG formation.

B. *Collaboration of Core Networks at Different Time-zones*

The data rates of users change drastically throughout a day. But aggregate data usage graphs on gateways show similar diurnal behavior. The capacity of core network gateways should be configured based on peak time traffic, thus, this can cause under-utilization of resources (i.e., CPU) at non-peak traffic duration. While the user behavior can be altered to some extent with ideas such as tiered pricing [35], it cannot be changed drastically. For this problem, we propose the collaboration of core network gateways (of different or same MNOs) at different time zones so that they can achieve a balanced load at different times of the day with a reduced maximum load. In other words, we propose a user data exchange system between the gateways of different time zones. The gateway having low usage on its system will offload some of the user traffic load of other gateways having peak time.

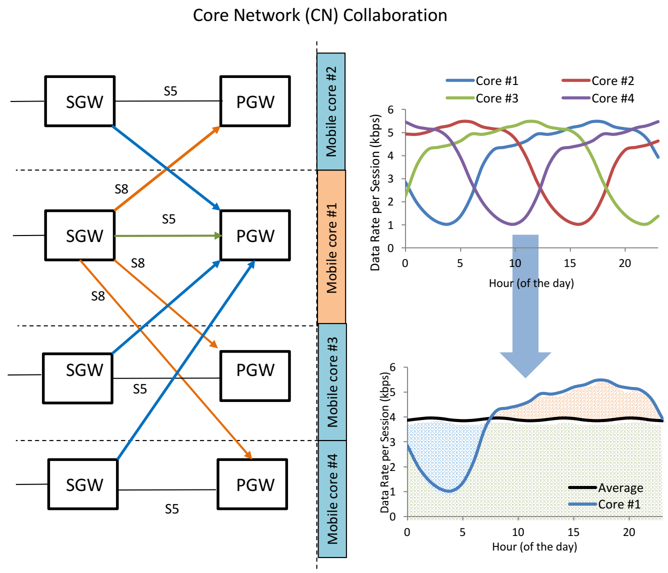

Fig. 5: Core network collaboration example with four operators with 6 hours apart from each other (GMT+0, GMT+6, GMT+12, GMT+18). At non-peak durations (until 7 a.m.), carrier #1's PGW can help processing other carriers traffic, then it offloads some of its traffic to other carriers' PGW by forwarding from its own SGW (over S8 interface).

The idea is illustrated in Figure 5 with an example of four core networks. Each of these core networks can obtain very balanced average usage and reduced maximum usage (i.e., leaving room for extra capacity) on their PGWs with proper direction of traffic between each other (this could be considered like a managed roaming activity). Figure 6 shows the amount of capacity increase with different collaborating cores and hour differences in their time zones.

Such a system can easily be conducted between the gateways of the same MNO as they are managed by the same team. Between the gateways of different operators, this can still be achieved through agreements on a perfect and trusted collaboration and by exchanging heavy user traffic. As a result both MNOs mutually benefit, and a balanced and reduced maximum utilization of each of the core network gateways can be achieved.

It is also possible that multiple MNOs can participate in the proposed system. Moreover, each MNO can show different

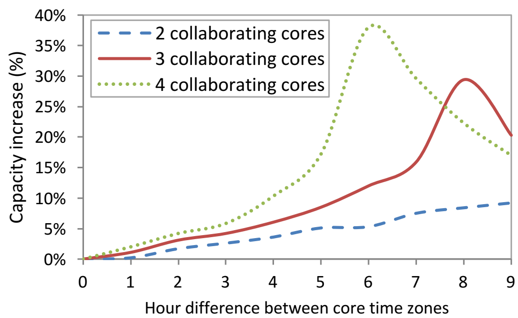

Fig. 6: The potential capacity increase with different collaborating core networks and hour differences in the their time zones.

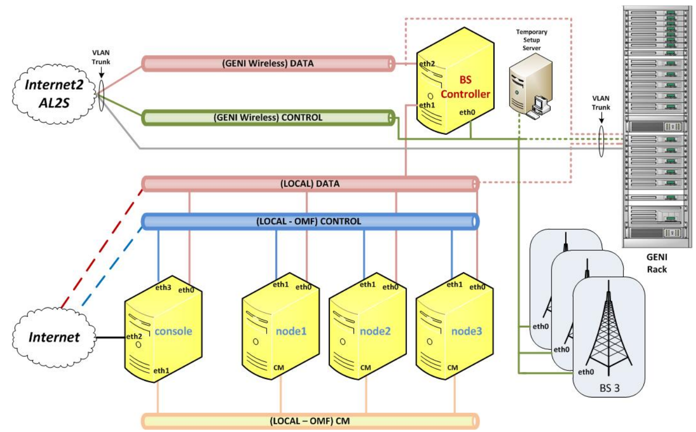

Fig. 7: 4G-LTE GENI Campus Deployment

level of trust and willingness to collaborate. In such a scenario with many participating operators from different time zones, traffic loads and levels of trust and willingness to collaborate, a game theoretical approach can be used to achieve highest benefit for each operator.

Assume U_{ij} is user $i \in N_j$ of operator $j \in O$, d^t_{ij} shows the data rate associated with U_{ij} at time t. We define a boolean parameter, β^t_{ijk} which is set to 1 if d^t_{ij} is offloaded to operator k. C^t_{ijk} is the virtual credit paid by operator j to k for its offloading d^t_{ij}. The load on operator j at time t, L^t_j, can be calculated as:

$$L^t_j = \sum_{i \in N_j} d^t_{ij} + \sum_{k \in O} \sum_{i \in N_k} \beta^t_{ikj} d^t_{ik} - \sum_{i \in N_j} \left(\sum_{k \in O} \beta^t_{ijk} \right) d^t_{ij}$$

$$\text{s.t.} \sum_{k \in O} \beta^t_{ijk} = \{0, 1\}$$

Here, the first term refers to the total data created by operator j's users. Second term is the data of other operators that it offloads and the last term shows its users data that is offloaded to other users. The objective function is to minimize the deviation of total load on each operator's core network during a day. Assuming a discrete model based on T time frames in a day (i.e., L_t represents data load in time frame t), for operator j, this can be calculated with $\sigma = \sqrt{\sum_{t=1}^{T} (L^t_j - \mu)^2 / T}$ where $\mu = \sum_{t=1}^{T} L^t_j / T$. Moreover, in order to promote collaboration and prevent selfishness, the total virtual credit balance of each operator j, V_j, should be set to zero (after trades between operators): By this way, different cores or operators could be encouraged to serve for others while being served by others.

IV. Related Work

As smartphones have built-in WiFi capability and there is ubiquitous access opportunity to connect to WLAN, offloading the mobile data through WiFi access points (AP) has been considered as an immediate solution for unloading the burden on cellular network. Several studies have been conducted on the deployment and recruitment of WiFi APs [7], [11], [12], [27]–[29] with the extensions of delayed [20], [22] and on-the-spot offloading [21]. There are also studies that look at the hybrid usage of WiFi and cellular access [24], and quality aware offloading [10].

All these studies offer a promising immediate solution for decreasing cellular burden only when the WiFi traffic is not managed by MNOs. Moreover, MNOs have been adding new gateways to their backhaul to integrate trusted or untrusted WiFi traffic to their cellular architecture. As a result, the load on mobile core network gateways remain unchanged, and thus, alternative approaches are still needed.

The core network offloading is also studied by introducing a new gateway between UE and PGW. The goal of this gateway is to offload the traffic that is destined to Internet. This can be achieved by using different identifiers like QCI and APN. If such a traffic is detected, it is forwarded towards IP network bypassing the core network gateways.

Media optimization solutions also indirectly help offload some traffic on core network gateways [4]. Using various techniques like compressing, the data traffic volume is reduced and a better network resource consumption with a better user experience is accomplished.

These media optimization servers are located at the Internet side of core network gateways. Other techniques could be listed as video caching, bandwidth adjustment, video pacing and web optimization. In video caching, popular videos are stored in the media optimization servers and users are provided content from these local servers rather than bringing them from original servers. In bandwidth adjustment, the bitrate of video is adjusted based on available bandwidth for user. Video pacing aims to pace the downloading and buffering of video as it is watched so that users who open a video but close it without completely watching it will not cause waste. Finally, for web optimization, more efficient mobile web browsers [36] and optimized website downloading and loading are targeted.

V. Future Work

The evaluation of the two proposed schemes requires visibility and control of the mobile core network and UE. To evaluate them in a deployed scenario, we plan to use the GENI (Global Environment for Network Innovations) wireless testbed [33]. GENI provides access to a fully programmable end to end 4G-LTE deployment at twelve campuses across the USA. Each deployment consists of a local cloud compute and storage cluster called a GENI Rack, which runs the EPC and uses high speed fiber connectivity to the LTE eNodeB. GENI provides UE's in the form of USB dongles and Android handsets. These deployments are all interconnected over a research network backbone, Internet2 (see Figure 7). GENI also provides WiFi AP's that will allow us to evaluate the SOG and Core Network collaboration schemes.

VI. Conclusion

In this paper, we present two different solutions for efficient resource utilization in mobile core network. While the first targets low utilization of user sessions and waste in the number of sessions, the second one aims to balance the daily usage and reduce maximum utilization. The solutions proposed here do not change overall user traffic load on the core network but they target efficient utilization of resources on the core network. Thus, the capacity of the core network gateway is increased and more user traffic can be handled by the same hardware. Moreover, all the optimization techniques presented in related work are orthogonal to our solutions and they can be applied over our proposed models for further benefit.

References

[1] Pew Research Center, *Mobile Technology Fact Sheet*, last accessed Sept 2015.

[2] Ericson, *Ericsson Mobility Report*, last accessed Sept 2015.

[3] Cisco Visual Networking Index, *http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html*, Feb, 2016.

[4] A. Aijaz, H. Aghvami, M. Amani, *A survey on mobile data offloading: technical and business perspectives*, IEEE Wireless Communications, 20 (2) (2013) 104112.

[5] K. Lee, I. Rhee, J. Lee, S. Chong, and Y. Yi, *Mobile data offloading: how much can wifi deliver?* in Proc. of Co-NEXT, pp. 26:1-26:12, 2010.

[6] E. Bulut, and B. K. Szymanski, *Understanding User Behavior via Mobile Data Analysis*, in Proc. of IEEE ICC Workshops, London, UK, June 8, 2015, pp. 1548-1553.

[7] E. Bulut and B. K. Szymanski, *WiFi Access Point Deployment for Efficient Mobile Data Offloading*, ACM Mobile Computing and Communications Review 17(1):71-78, January, 2013.

[8] E. Bulut and B. K. Szymanski, *Exploiting Friendship Relations for Efficient Routing in Mobile Social Networks*, IEEE Transactions on Parallel and Distributed Systems, 23 (12), pp.22542265, 2012.

[9] E. Bulut, S. Geyik, and B. Szymanski, *Utilizing Correlated Node Mobility for Efficient DTN Routing*, Pervasive and Mobile Computing 13 August, 2014, pp. 150-163.

[10] W. Hu and G. Cao, *Quality-Aware Traffic Offloading in Wireless Networks*, in Proc. of ACM MobiHoc, 2014.

[11] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, *A Double Auction Mechanism for Mobile Data Offloading Markets*, in IEEE/ACM Transactions on Networking, 99, 2014.

[12] X. Zhuo, W. Gao, G. Cao, and S. Hua, *An incentive framework for cellular traffic offloading*, in IEEE Trans. Mobile Comput., vol. 13, no. 3, pp. 541555, Mar. 2014.

[13] Y. Zhang, *User Mobility from the View of Cellular Data Networks*, in Proc. of the 33rd Annual IEEE Conference on Computer Communications (INFOCOM 2014), Toronto, Canada, April 2014.

[14] Y. Jin, N. Duffield, A. Gerber, P. Haffner, W. Hsu, G. Jacobson, S. Sen, S. Venkataraman and Z. Zhang, *Characterizing Data Usage Patterns in a Large Cellular Network*, In Proc. of the ACM SIGCOMM workshop on CellNet, 2012, pp. 7-12.

[15] L. Meng, S. Liu, A. Striegel, *Analyzing the longitudinal impact of proximity, location, and personality on smartphone usage*, in Computational Social Networks, December 1:6, 2014.

[16] J. Yang, Y. Qiao, X. Zhang, H. He, *Characterizing User Behavior in Mobile Internet*, IEEE Transactions on Emerging Topics in Computing. PP (99), 2014.

[17] Evolving Systems, *Dynamic SIM Allocation Solution Overview White Paper*, online, last accessed on Sept 2015.

[18] , Multi-SIM Card, *https://en.wikipedia.org/wiki/Multi-SIM_card*, online, last accessed on Sept 2015.

[19] C. Joe-Wong, S. Sen, S. Ha, and M. Chiang, *Do Mobile Data Plans Affect Usage? Results from a Pricing Trial with ISP Customers*, in IEEE PAM 2015.

[20] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, *Cellular traffic offloading onto network-assisted device-to-device connectiozns*, IEEE Commun. Mag., vol. 52, no. 4, pp. 2031, Apr. 2014.

[21] F. Rebecchi, M. D. Amorim, V. Conan, A. Passarella, R. Bruno and M. Conti, *Data Offloading Techniques in Cellular Networks*, IEEE Communications Surveys & Tutorials, 2014.

[22] M. Cheung, and J. Huang, *DAWN: Delay-Aware Wi-Fi Offloading and Network Selection*, in IEEE Journal on Selected Areas in Communications, 33 (6), pp. 1214-1223, 2015.

[23] I. Trestian, S. Ranjan, A. Kuzmanovic, and A. Nucci, *Taming the Mobile Data Deluge with Drop Zones*, In IEEE/ACM Transactions on Networking, to appear, 2012.
vol. 46, no. 9, pp. 59-67, September 2008.

[24] A. Balasubramanian, R. Mahajan, and A. Venkataramani, *Augmenting Mobile 3G Using WiFi*, in ACM MobiSys'10, Jun. 2010, pp. 209-222, 2010.

[25] A. Sharma, V. Navda, R. Ramjee, V. N. Padmanabhan, and E. M. Belding, *Cool-tether: energy efficient on-the-fly wifi hot-spots using mobile phones*, in Proc. of the ACM International Conference on Emerging Networking Experiments and Technologies (CoNEXT09), ACM, 2009.

[26] 3GPP, *3rd Generation Partnership Project; Technical Specification Group SA; Feasibility Study for Proximity Services (ProSe) (Release 12)*, TR 22.803 V1.0.0, Aug. 2012.

[27] S. Dimatteo, P. Hui, B. Han, and V. O. Li, *Cellular traffic offloading through WiFi networks*, in Proc. IEEE MASS, 2011, pp. 192201.

[28] N. Ristanovic, J.-Y. L. Boudec, A. Chaintreau, V. Erramilli, *Energy Efficient Offloading of 3G Networks*, in Proc. of IEEE MASS, 2011.

[29] X. Zhuo, W. Gao, G. Cao, and Y. Dai, *Win-Coupon: An Incentive Framework for 3G Traffic Offloading*, in Proc. of IEEE ICNP, 2011.

[30] E. J. Vergara, and S. N. Tehrani, *Watts2Share: Energy-Aware Traffic Consolidation*, in Proc. of IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber Physica and Social Computing, 2013.

[31] A. Aucinas, and J. Crowcroft, *Demo: PhoneLets - Offloading the Phone off your Phone for Energy, Cost and Network Load Optimization*, in ACM Mobicom, 2014.

[32] A. Disessa, M. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, and E. Stage, *Toward a Scientific Practice of Science Education*, Lawrence Erlbaum Associates, July 1990.

[33] A. Gosain, M. Berman, M. Brinn, T. Mitchell, C. Li, Y. Wang, H. Jin, J. Hua, and H. Zhang, *Enabling Campus Edge Computing Using GENI Racks and Mobile Resources*, 2016 IEEE/ACM Symposium on Edge Computing, 28 Oct 2016.

[34] Z. Tang, S. Guo, P. Li, T. Miyazaki, H. Jin, and X. Liao, *Energy-Efficient Transmission Scheduling in Mobile Phones Using Machine Learning and Participatory Sensing*, IEEE Transactions on Vehicular Technology, vol. 54, no.7, July 2015.

[35] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, *TUBE: Time-dependent pricing for mobile data*, in Proc. of ACM SIGCOMM, pp. 247258, ACM, 2012.

[36] B. Zhao, Q. Zheng, G. Cao, and S. Addepalli, *Energy-Aware Web Browsing in 3G Based Smartphones*, in Proc. of IEEE ICDCS, 2013.

[37] S. C. Geyik, E. Bulut and B. K. Szymanski, *Grammatical inference for modeling mobility patterns in networks*, IEEE Transactions on Mobile Computing, 12 (11), pp.2119–2131, 2013.