RESEARCHERID
THOMSON REUTERS
[ICEMESM-18]
IC™ Value: 3.00
ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

REVIEW ON MOBILITY-AWARE CACHING IN D2D NETWORKS

Vikas B Wasade[1] Nilesh Bodne[2] , Ruchika Nikhare[3]

[1]Department of Electronics (Comm. Engineering), R.T.M. Nagpur University [2]Department of Electronics (Comm. Engineering), R.T.M. Nagpur University [3]Department of Electronics (Comm. Engineering), R.T.M. Nagpur University

DOI: 10.5281/zenodo.1232295

ABSTRACT

Caching at mobile devices can facilitate device-to-device (D2D) communications, which may significantly improve spectrum efficiency and alleviate the heavy burden on backhaul links. However, most previous works ignored user mobility, thus having limited practical applications. In this paper, we take advantage of the user mobility pattern by the inter-contact times between different users, and propose a mobility-aware caching placement strategy to maximize the data offloading ratio, which is defined as the percentage of the requested data that can be delivered via D2D links rather than through base stations (BSs). Given the NP-hard caching placement problem, we first propose an optimal dynamic programming (DP) algorithm to obtain a performance benchmark, but with much lower complexity than exhaustive search. We will then prove that the problem falls in the category of monotone submodular maximization over a matroid constraint, and propose a time- efficient greedy algorithm, which achieves an approximation ratio as 1:2 . Simulation results with real-life data-sets will validate the effectiveness of our proposed mobility-aware caching placement strategy. We observe that users moving at either a very low or very high speed should cache the most popular files, while users moving at a medium speed should cache less popular files to avoid duplication

Keywords: Caching device-to-device communications, human mobility, matroid constraint, submodular function.

I. INTRODUCTION

With the popularity of smart phones, the data traffic generated by mobile applications, e.g., multimedia file sharing and video streaming, is undergoing an exponential growth, and will soon reach the capacity limit of current cellular networks. To meet the heavy demand, network densification is commonly adopted to achieve higher network capacity, which is expected to increase by 1000 times in future 5G networks. However, the increase of access points and user traffic put a heavy burden on backhaul links, which connect base stations (BSs) with the core network. To reduce the backhaul burden, one promising approach is to cache popular contents at BSs and user devices , so that mobile users can get the required content from local BSs or nearby user devices without utilizing backhaul links. Such local access can also reduce the download delay and improve the energy efficiency .

Most previous investigations on wireless caching networks assumed fixed network topologies. However, user mobility is an intrinsic feature of wireless networks, which changes the network topologies over time. Thus, it is critical to take the user mobility pattern into account. On the other hand, user mobility can also be a useful feature to exploit, as it will increase the communication opportunities of moving users. Mobility-aware design has been proved to be an effective approach to deal with lots of problems in wireless networks. For example, exploiting user mobility helps improve capacity in ad hoc networks [20] and reduce the probability of failed file delivery in femto-caching networks.In this paper, we will propose an effective mobility-aware caching strategy in device-to-device (D2D) caching networks to offload traffic from the cellular network.

II. LITERATURE SURVEY

Rui Wang, Jun Zhang, S.H. Song and Khaled B. Letaief are with the Department of Electronic and Computer Engineering, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Khaled B. Letaief is also with Hamad bin Khalifa University, Doha, Qatar Design and Implement Caching in D2D Networks.

Compared with caching at BSs, caching at mobile devices provides new features and unique advantages. First, the aggregate caching capacity grows with the number of devices in the network, and thus the benefit of device caching improves as the network size increases. Second, by caching popular content at devices, mobile users may acquire required files from close user devices via D2D communications, rather than through the BS. This will significantly reduce the mobile traffic on the backbone network and alleviate the heavy burden on the backhaul links. However, mobile caching also faces some new challenges. Specifically, not only the users, but also the caching helpers are moving over time, which brings additional difficulties in the caching design. There have been lots of efforts on D2D caching networks while assuming fixed network topologies. Considering a single cell scenario, it was shown in that D2D caching outperforms other schemes, including the conventional unicasting scheme, the harmonic broadcasting scheme, and the coded multicast scheme. Assuming that each device can store one file, Golrezaei *et al.* analyzed the scaling behavior of the number of active links in a D2D caching network as the number of mobile devices increases .It was found that the concentration of the file request distribution affects the scaling laws, and three concentration regimes were identified. In the outage-throughput tradeoff in D2D caching networks was investigated and optimal scaling laws of per-user throughput were derived as the numbers of mobile devices and files in the library grow under a simple uncoded protocol. Meanwhile, the case using the coded delivery Scheme was investigated in. There are some preliminary studies considering user mobility. In Poularakis *et al.* studied a femto-caching network with user mobility. A Markov chain model was adopted to represent which helper, i.e., a particular femtocell BS, was accessed by a specific user in different time slots. However, in D2D caching networks, such a model cannot be adopted, since there is no fixed caching helper, and all the mobile users may move over time. The effect of user mobility on D2D caching was investigated via simulations in which showed that user mobility does not have a significant impact on a random caching scheme. However, such a caching scheme failed to take advantage of the user mobility pattern. In it showed that user mobility has positive effect on D2D caching. In [30], Lan *et al.* considered the case where mobile users can update caching content based on the file requirement and user mobility. However, it was assumed that one complete file can be transmitted via any D2D link when two users contact, which is not practical, considering the limited communication time and transmission rate. More recently, several design methodologies for mobility-aware caching were proposed in but more thorough investigations will be needed, especially on practical implementations.

III. EXISITING SYSTEM

In this section, we will firstly introduce the mobility model adopted in our study, together with the caching strategy and file transmission model. The caching placement problem will then be formulated.

A. User Mobility Model:

The inter-contact model can capture the connectivity information in the user mobility pattern, and has been widely investigated in wireless networks,Thus, it is adopted in this paper to model the mobility pattern of mobile users. In this model, mobile users may contact with each other when they are within the transmission range. Correspondingly, the *contact time* for two mobile users is defined as the time that they can contact with each other, i.e., they may exchange files during the contact time. Then, the *inter-contact time* for two mobile devices is 6 defined as the time between two consecutive contact times. Specifically, we consider a network with Nu mobile users, whose index set is denoted as D = {1, 2, ...,Nu}. Same as [32], we model the locations of contact times in the timeline for any two users i and j as a Poisson process with intensity _i,j . For simplification, we assume that the timelines for different device pairs are independent. We call _i,j as the *pairwise contact rate* between users i and j, which represents the average number of contacts per unit time. The pairwise contact rate can be estimated from historical data, and we will test our results on real-life data-sets in the simulation.

B. Caching Strategy and File Delivery Model:

We consider a library of Nfile files, whose index set is denoted as F = {1, 2, ...,Nfile}. Rateless Fountain coding is applied where each file is encoded into a large number of different segments, and it can be recovered

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

by collecting a certain number of encoded segments. Accordingly, it can be guaranteed that there is no repetitive encoded segment in the network. Specifically, we assume that each file is encoded into multiple segments, each with size s bits, and file f can be recovered by collecting Kf encoded segments. Note that the value of Kf depends on the size of file f. It is assumed that each mobile user reserves a certain amount of storage capacity for caching, which can store at most C encoded segments. The number of encoded segments of file f cached in user i is denoted as xi,f .

Mobile users will request files in the file library based on their demands. For simplicity, we assume that the requests of all the users follow the same distribution, and file f is requested by one user with probability pf , where

$$\sum_{f \in \mathcal{F}} p_f = 1$$

P f∈ F pf = 1. When a user i requests a file f, it will start to download encoded segments of file f from the encountered users, and also check its own cache. We assume that the duration of each contact of users i and j is tci ,j seconds, and the transmission rate from user j to user i is ri,j bps Accordingly, we consider that

$$B_{i,j} = \left\lfloor \frac{t_{i,j}^{c} r_{i,j}}{s} \right\rfloor$$

Segments can be transmitted within one contact from user j to user i, which is more practical than most existing works, e.g.,. We also assume that there is a delay constraint, denoted as Td. If user i cannot collect at least Kf different encoded segments of file f within Td, it will request the remaining segments from the BS. For example, as shown in Fig. 1, user 1 requests a file, which is not in its own cache. Thus, it starts to wait for encountering the users storing the requested file. Within time t < Td, user 1 first gets one segment from user 2, and then another segment from user 3. As user 1 collects enough segments via D2D links to recover the requested file, it does not need to download the file from the BS.

C. Problem Formulation:

In this paper, we investigate the caching placement strategy to maximize the data offloading ratio, i.e., the percentage of the requested data that can be delivered via D2D links, rather than via the BS. If the D2D links can offload more data from the BS, it will lead to higher spatial reuse efficiency and also significantly reduce the backhaul burden. Specifically, for user i, the data offloading ratio is defined as

$$E_i = \mathbb{E}_{f \in \mathcal{F}}\left[\mathbb{E}_{u_{i,f}}\left[\frac{\min\left(u_{i,f}, K_f\right)}{K_f}\middle| \text{User } i \text{ requests file } f\right]\right] = \sum_{f \in \mathcal{F}} p_f \left\{\frac{\mathbb{E}_{u_{i,f}}\left[\min\left(u_{i,f}, K_f\right)\right]}{K_f}\right\},$$

D. Divide and Conquer Algorithm:

To Evaluate the Objective To evaluate the objective function of the optimization problem, the distribution of the summation of Nu independent random variables need to be obtained. A direct calculation will incur high computational complexity, which increases exponentially with the number of users Nu. To efficiently solve, we first develop a time-efficient approach to evaluate the objective.

With the definition of the expectation, we can rewrite the objective function as

$$\frac{1}{N_u}\sum_{i \in \mathcal{D}} E_i = \frac{1}{N_u}\sum_{i \in \mathcal{D}}\sum_{f \in \mathcal{F}} \frac{p_f}{K_f}\left\{\sum_{q=0}^{K_f-1} q \Pr\left[\sum_{j \in \mathcal{D}} \min(B_{i,j}M_{i,j}, x_{j,f}) = q\right]\right.$$

$$\left. + K_f - K_f \Pr\left[\sum_{j \in \mathcal{D}} \min(B_{i,j}M_{i,j}, x_{j,f}) \leqslant K_f - 1\right]\right\}.$$

Then, the main difficulty is to get the distribution of the summation of the Nu independent random variables, denoted as

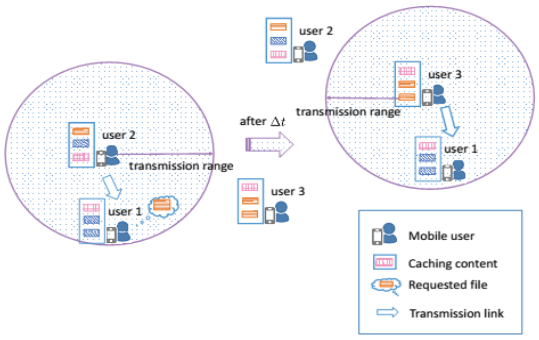
$$S_{i,f}^{u} = \sum_{j \in \mathcal{D}} \min(B_{i,j}M_{i,j}, x_{j,f}).$$

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

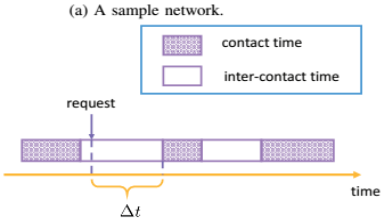
In the following, we will propose a divide and conquer algorithm to reduce the computation complexity of getting the cumulative distribution function (cdf) of Su i,f . The algorithm decomposes the original problem into multiple sub-problems, and solve them one by one. Then, the solution of the original problem is constructed based on those for the sub-problems.

IV. RESEARCH METHODOLOGY

The information of user connectivity in the user mobility pattern is captured with the inter-contact model Specifically, as shown in Fig. 1b, for an arbitrary pair of mobile users, the timeline consists of both contact times, defined as the times when the users are within the transmission range, and inter contact times, defined as the times between contact times.



(a) A sample network.



(b) The timeline of user 1 and user 3.

Each file is encoded into several segments by the rateless Fountain code ,and a mobile user needs to collect enough encoded segments to recover the requested file. Once a user is in contact with some other users who cache part of its requested file, it will get some of the segments via D2D communication. Our objective is to design caching placement to maximize the data offloading ratio. The main contributions of this paper are summarized as follows:

As each mobile user may get the requested file via multiple contacts with other users, the complexity of calculating the objective function, i.e.,the data offloading ratio,increases exponentially with the number of users,which brings a major difficulty for algorithm design.We first propose a divide and conquer algorithm to efficiently evaluate the objective,with quadratic complexity with respect to the number of users.

1- The caching placement problem is shown to be NP-hard and a dynamic programming (DP) algorithm is proposed to obtain the optimal solution with much lower complexity compared to exhaustive search.Although the DP algorithm is still impractical, it can serve as a performance benchmark for systems with small to medium sizes. Moreover, its computation complexity is linear with the number of files, and thus it can easily deal with a large file library.

2- To propose a practical solution, we reformulate the problem and prove that it is a monotone submodular maximization problem over a matroid constraint.The main contribution in this part is to prove that the complicated objective function is a monotone submodular function.With the reformulated form, a greedy algorithm is developed,which achieves at least 1 2 of the optimal value.

V. CONCLUSIONS

In this paper, we exploited user mobility to improve caching placement in D2D networks using a coded cache protocol. We took advantage of the inter- contact pattern of user mobility when formulating the caching placement problem. To assist the evaluation of the complicated objective function, we proposed a divide and

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

conquer algorithm. A DP algorithm was then developed to find the optimal caching placement, which is much more efficient than exhaustive search. By reformulating it as a monotone submodular maximization problem over a matroid constraint, we developed an effective greedy caching placement algorithm, which achieves a near-optimal performance. Simulation results based on both the mathematical model and reallife data-sets showed that the proposed mobility-aware greedy caching strategy outperforms both the random caching strategy and popular caching strategy. It is observed that slow users tend to cache the most popular files to support themselves, since they have few opportunities to communicate with others via D2D links. Fast moving users also tend to cache the most popular files, but the purpose is mainly to fulfill the requests from other users, since they have much more opportunities to establish D2D links with others. Meanwhile, users with medium velocity tend to cache less popular files to better exploit the D2D communications. For future works, it would be interesting to exploit other information of the user mobility pattern to further improve the caching efficiency. It will also be interesting to investigate joint caching at both BSs and devices, as well as distributed algorithms for implementation in large- size networks.

VI. REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[2]] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green Cloud-RAN," IEEE Trans. Wireless Commun., vol. 13, no. 5, pp. 2809– 2823, May 2014.
[3] Y. Shi, J. Zhang, K. B. Letaief, B. Bai, and W. Chen, "Large-scale convex optimization for ultra- dense Cloud-RAN," IEEE Wireless Commun., vol. 22, no. 3, pp. 84–91, Jan. 2015.
[4] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: A survey," IEEE Commun. Mag., vol. 46, no. 9, pp. 59–67, Sep. 2008.
[5] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," IEEE Commun. Mag., vol. 51, no. 4, pp. 142–149, Apr. 2013.
[6] H. A. Omar, W. Zhuang, and L. Li, " VeMAC: A TDMA-based MAC protocol for reliable broadcast in VANETs , " IEEE Trans. Mobile Comput., vol. 12, no. 9, pp. 1724–1736, Jun. 2013.