

Q-Learning Based Social Community-Aware Energy Efficient Cooperative Caching in 5G Networks

Han Yeo Reum Im, Kyi Thar, and Choong Seon Hong

Department of Computer Science and Engineering, Kyung Hee University, 17104, Republic of Korea

Email: {prosummer, kyithar, cshong}@khu.ac.kr

Abstract— To satisfy the vast demand of data traffic caused by various mobile devices, caching contents at Base Stations (BSs) and User Equipment (UE) has become a promising solution. A lot of recent works have proved that an approach of caching at edge devices is efficient to reduce latency and alleviate backhaul load so that the probability to meet users' Quality of Service (QoS) is much higher. However, due to the selfishness of mobile users, it is hard to convince users to store contents for other users. In this regard, we consider the social connection between users in order to induce users to store contents for the social community. So, to implement the socially aware Device to Device (D2D) caching, we construct two types of graphs: i) physical D2D graph which is constructed based on users communication over D2D links and ii) logical social graph which is consist of a group of people who share the common interest. Then, we formulate socially aware D2D caching problem into Markov Decision Process (MDP). Additionally, we consider the distance and the social ties between users, and mobile devices' capacity to optimally store contents under the constraint that energy consumption of cellular link is much higher than D2D link. Finally, we solve the D2D caching problem by using Q-learning where the goal is to minimize the total energy consumption over Macrocell Base Station (MBS) and mobile devices.

Keywords—Proactive caching, socially aware networking, device-to-device communication, Q learning, 5G Networks

I. INTRODUCTION

The demand for mobile data traffic caused by various types of devices has increased over the past few years. According to the report entitled Cisco visual networking index, the IP traffic over the world will reach 4.8ZB per year by 2022 and Internet traffic in rush hour is growing much faster than that in a non-rush hour [1]. When it comes to video traffic, it requires not only much more traffic but also lots of higher bandwidth. As a lot of mobile User Equipment (UE) requests tremendous contents to the content server every second, backhaul load between the content server and the base station has become heavier more than ever.

To alleviate the heavy burden on backhaul load, an approach of content caching at edge devices has emerged as a promising solution. This paradigm is called *Proactive Caching* which means contents are cached at the storage before users request the contents. To deal with this problem, it is important to predict which contents will be requested at a specific time. Moreover, mobile users move everywhere

every time so that forecasting users' mobility has been recognized as a good approach for proactive caching.

One of the recent works has focused on prediction of content popularity and caching proactively [12]. They predict the future class label for each content and the number of requests for each content respectively using Deep Learning. Then, they cache the contents having a high probability to be requested at the base station. Also, in the cloud radio access networks (CRANs), the machine learning-based caching scheme that predicts content request distribution and mobility pattern at the baseband units (BBUs) level is proposed to maximize capacity gain [13].

On the other hand, several literatures have considered the information of social connection between users [2]-[3]. If the social connection between two users is strong enough, there is a higher chance for users to transmit its content to one another. By taking social ties between users and common interests into consideration, a new hypergraph framework for caching in D2D communication is introduced [2]. Also, for improvement of the content offloading over wireless D2D communications, a caching strategy utilizing the multi-community social context of users is proposed [3]. With a cooperative game method, they model collaborative effects between the set of UEs

However, there are few studies on leveraging user's social community over both BS level and D2D level. Moreover, most of the related works have not considered mobile users' cache capacity. To summarize, our contributions are summarized as follow:

- We consider both of the energy consumption of BS and D2D level. We assume that if requested contents are not stored in the nearby devices, the users will be transmitted the desired content from the MBS.
- The users in the same social community prefer similar short video contents such as YouTube music videos, so the content catalog for a community is defined already. We organize two different graphs: i) physical D2D graph having user nodes as vertices and D2D links as edges, and ii) logical social graph which is consist of a group of people who share the common interest.
- We formulate the content distribution scheme to all users in the same social community with the consideration of the distance and the social ties between users and cache capacity of mobile devices. We solve this problem by using Q-learning method aiming to minimize total energy consumption.

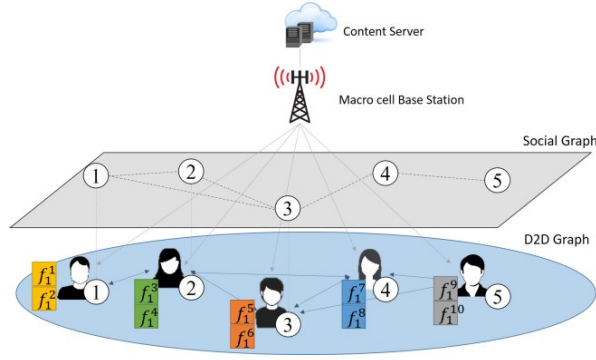


Figure 1 An example of Social Community Aware Caching

II. SYSTEM MODEL

In this paper, we consider 5G networks with MBS and mobile devices. A detailed example is illustrated as Fig. 1. The content server is connected to MBS over backhaul link so that content can be transmitted from the content server to MBS. In the scope of this MBS, there are 5 users named by numbers. The MBS can transmit requested content to each user over a cellular link. If channel gain between two users is high enough, we assume that there is a D2D link between them. As described in the figure, let's say that the 5 users share the common interest. Then, we can say that they are in the same social community. Moreover, we can construct a social graph with edges of the social tie which can be frequency, intimacy/closeness, longevity, reciprocity, recency, multiple social contexts, and mutual confiding (trust) [10]. Also, the users in the graph cache several contents of the content catalog under the constraint of their cache capacity.

A. Graph Model

In this scenario, there are N number of users in the range of MBS θ . A set of mobile users is denoted by U_N . Their location is represented by $P_N = (x_N, y_N)$. Then, the D2D Graph $G_D(U_N, E_D)$ can be constructed which is connected-graph where E_D is the edge of graph ($E_D = \{E_{m,n} | U_m, U_n \in U_k \text{ and } m \neq n\}$). It is weighted by the distance of two user devices over D2D communication $r_D = \{r_{m,n}^D | U_m, U_n \in U_k \text{ and } m \neq n\}$. We consider K number of social community and each of community constructs social graph $G_S^k(U_k, E_S^k)$ which is connected-graph where the edge of a graph is weighted by a social tie between two different users in social community k . We can denote a set of users in social community k as U_k . Then, the edge of this graph, $E_S^k = \{E_{m,n}^k | U_m, U_n \in U_k \text{ and } m \neq n\}$ can be weighted by the social tie between user m and n .

We assume that popular contents differ from each social community. Then we can define content catalog for social community k which can be denoted by N_f^k . Also, we assume that all the contents have the same size C_s . We assume that the MBS distribute several popular contents of

each social community to all users who belong to the social community.

There are several assumptions on the proposed system model. We assume a static content catalog which is consist of N_f contents ordered by content popularity. For modeling content popularity, Zipf distribution is commonly used. When each mobile user requests a content, then the probability of requesting f^{th} popular content is denoted by [5]:

$$p_f = \frac{f^{-\alpha}}{\sum_{j=1}^{N_f} j^{-\alpha}} \quad (1)$$

Also, we assume that each mobile user equipment has the same limited cache capacity C . When it comes to distance, the distance between one-hop neighbors should less than d .

B. Energy Consumption Model

In the proposed environment, we assume overlay inband D2D communication that D2D users are allocated to dedicated spectrum resources [6]. As a result of dedicated resources, the problem of interference is much alleviated in overlay inband D2D communication. So we do not consider the interference on D2D communication. Also, we assume that BS manages the beaconing channel so that BS can support D2D users which makes more energy efficient. Similar with [7], we define the energy consumption of D2D communication. According to [8], the average data rate of D2D communication is denoted by:

$$R_D = \mathbb{E} \left\{ W_D \log_2 \left(1 + \frac{T_D h_D^2 r_D^{-\delta}}{\sigma_D^2} \right) \right\} \approx W_D \log_2 \left(1 + \frac{T_D r_D^{-\delta}}{\sigma_D^2} \right) \quad (2)$$

T_D is the mobile device transmission power. W_D is the channel bandwidth from D_i to D_j . h_D denotes channel gain which follows zero mean complex Gaussian distribution with unit variance. r_D represents the distance between D_i and D_j . σ_D^2 is white Gaussian noise and δ is path loss factor. In [9], they modeled the power consumption of D_j as $\rho_D T_D + C_D + H_D$, where C_D denotes the circuit power consumption of the mobile device, and ρ_D represents the inverse of power amplifier efficiency factor and H_D is the energy consumption of caching hardware devices. We do not take energy for delivery of the cached content into consideration. In case that D_j transmits L bits file to D_i , the energy consumption of D_j is followed by:

$$E_D = \frac{C_s}{R_D} (\rho_D T_D + C_D) \quad (3)$$

Regarding energy cost for MBS, we can model the downlink speed of the mobile device is formulated as:

$$R_B = \mathbb{E} \left\{ W_B \log_2 \left(1 + \frac{T_B h_B^2 r_B^{-\delta}}{\sigma_B^2} \right) \right\} \approx W_B \log_2 \left(1 + \frac{T_B r_B^{-\delta}}{\sigma_B^2} \right) \quad (4)$$

In this formulation, T_B is the MBS transmission power. W_B denotes the downlink transmission bandwidth from

MBS to the mobile device, h_B represents the channel power gain which follows zero mean complex Gaussian distribution with unit variance. r_B is the distance between D_i and MBS. σ_B^2 is the variance of additive white Gaussian noise and β is the path loss factor. According to [9], when we do not consider the energy consumption for file storage in MBSs, in case that MBS send L bits file to D_i , the energy consumption at the MBS is denoted by:

$$E_B = \frac{C_s}{r_B} (\rho_B T_B + C_B) \quad (5)$$

C_B represents an offset of site power including the baseband processor, the cooling system and etc, and ρ_B denotes the inverse of power amplifier efficiency factor.

Finally, the optimization problem of minimization of total energy consumption is represented by:

$$\min(\Sigma E_D + \Sigma E_B) \quad (6a)$$

$$\text{s. t.} \quad r_{m,n}^D \leq d, \forall U_m, U_n \in U_N \quad (6b)$$

$$r_{m,n}^B \leq \theta, \forall U_m, U_n \in U_N \quad (6c)$$

$$0 < C_s |S_{m,n}| \leq C, \forall U_m, U_n \in U_N \quad (6d)$$

III. Q-LEARNING BASED CONTENT DISTRIBUTION

To solve the content distribution to mobile user devices, we utilize Q-learning method [10] which yield great performance for getting the optimal policy corresponding to the given Markov Decision Processes (MDPs).

- *State*: In our problem, we consider a social community k to distribute several contents in content catalog N_f^k to all users in U_k . Then, we can define a matrix $S_{U_k \times N_f^k}$ which indicates the state of which content is stored at which user device. When user U_i stored content f_l , the indicator can be denoted by $\mathbb{I}_{U_i, f_l} = 1$. Otherwise, the indicator has zero value $\mathbb{I}_{U_i, f_l} = 0$. The matrix $S_{U_k \times N_f^k}$ stores this indicator information.
- *Action*: The action space is denoted by a matrix $A_{U_k \times N_f^k}$ which indicates the exchange of contents. When the content f_l in user device U_i is exchanged with the content f_m in user device U_j , then $A_{U_i, f_l} = -1, A_{U_j, f_m} = -1$ and $A_{U_i, f_m} = 1, A_{U_j, f_l} = 1$. In the case of no change, the value is zero. So the next state will be $S_{U_k \times N_f^k}^{t+1} = S_{U_k \times N_f^k}^t + A_{U_k \times N_f^k}^t$. We assume that each user device can store C/C_s number of contents. So when the agent decides to get rid of specific content, then that content should be cached at a different user device.
- *Reward*: The reward is defined under two cases. When the desired content is cached in the local storage, then the reward is a max social tie because we can assume that every user has the maximum social tie with itself intuitively ($R = \max_social_tie$). However, in case

that the requested content is stored at the one-hop neighbor, then the reward is represented by $R = (Social\ tie) - (distance)$. The total reward is the summation of all reward to service requested contents to all users.

In our problem, the purpose is finding the optimal policy for the distribution of all contents to users. Basically, Q-learning experiences the cases through episodes and updates its estimates which resemble Bellman equations. This updates can be denoted by:

$$Q(s, a) \leftarrow (1 - l)Q(s, a) + l \left[r + \gamma \max_{a'} Q(s', a') \right] \quad (7)$$

In this equation, $Q(s, a)$ denotes Q-value when agent acts a in the state s . When the Q-value is high, it means doing α when the state s is highly preferable. γ is the discount factor indicating the value decrease over time and $0 < \gamma < 1$. l means learning rate so when it is high, the agent reflects the most recent Q-value greatly in order to update Q-value. In short, the Q-value is updated with the summation of the current reward value and the largest Q-value among the actions that can be taken thereafter with discount factor γ . The Q-learning agent explores the environment for a new experience at a rate of ϵ .

The overall procedure of Q-learning based caching scheme is given by Algorithm 1. We input D2D graph G_D and social graph G_S^k to output the optimal policy π^* . After initializing the environment and setting the base station, we define the state space and the action space in (line 3). For each iteration, the agent chooses random action with the probability of ϵ to explore the environment (line 5 - line 6). Otherwise, the agent selects the action which has the largest Q value for the current state (line 8 - line 9). With new state and calculated reward, the agent update Q table (line 10 - line 14).

Algorithm 1 Q-learning based Socially Aware Caching Model

Input :	G_D, G_S^k	
Output :	Optimal caching distribution policy π^*	
1	Initialize environment parameters: $\delta, C_N, T_{N_f^k \times T}$	
2	Set base station with $P_T^B, W_B, \sigma_B^2, \alpha$	
3	Define $S_{U_k \times N_f^k}, A_{N_f^k \times N_f^k}$	
4	for episode in range(n_episode) do	
5	if rand(0,1) > ϵ then	▷ Exploration
6	action \leftarrow rand_choice($A_{N_f^k \times N_f^k}$)	
7	else	▷ Exploitation
8	get Q-table[state]	
9	action \leftarrow argmax(Q-table[state])	
10	update new_state, reward	
11	$Q_1 \leftarrow$ Q-table[state][action]	
12	$Q_2 \leftarrow$ reward + $\gamma * \max(Q_table[next_state])$	
13	add $l * (Q_2 - Q_1)$ to Q-table[state][action]	
14	state \leftarrow next_state	
15	if done then	
16	Break	

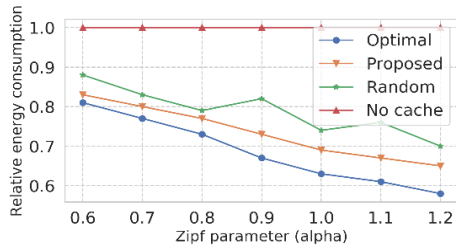


Figure 2 Relative energy consumption of different caching schemes

IV. SIMULATION RESULT

In this section, we simulate the proposed energy efficient Q-learning based caching scheme. We used Python 3.5 to make a simulation and NetworkX library[11] to organize D2D graph and social graph on the following environment. We considered 5G networks environment including 1 MBS and 3 mobile users and each user can cache 2 contents in their storage. We simulated the proposed algorithm for 1000 contents in the content catalog for a social community. In this environment, there are 90 states and 15 actions. The parameters related to the energy consumption are configured by referring to [8]. We compare the proposed solution with the random caching scheme. In the random caching scheme, the MBS distribute contents to all users randomly. For the calculation of the energy consumption over MBS and mobile devices, we assume the log of each users' content request which is followed by Zipf distribution is already given. In Fig. 2, the relative energy consumption of proposed caching and conservative caching methods is represented. With the no-caching scheme as an upper bound and the optimal caching scheme as a lower bound, the proposed caching outperforms random caching. From Fig. 3, we observe cache hit ratio under different Zipf parameter α . When the skewness α is getting larger, the probability to request popular content is increased. Generally, the proposed caching shows greater cache hit ratio compared with the random caching.

V. CONCLUSION AND FUTURE WORK

In this paper, we studied the influence of socially aware caching on MBS and mobile user devices and proposed the scheme of Q-learning based content distribution to users in the same social community under the constraint of each users' cache capacity. The proposed solution aims to maximize the energy efficiency over MBS and mobile devices, and we achieved high performance in terms of the cache hit ratio and energy consumption. For future work, we can extend the proposed solution on the multi-social community aware environment and analyze the latency of the proposed algorithm.

ACKNOWLEDGMENT

This work was partially supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2019-0-01287, Evolvable Deep

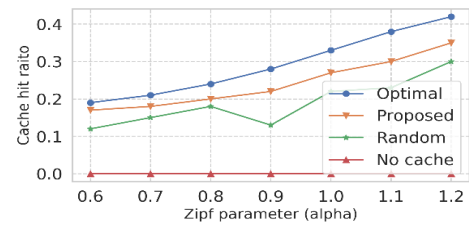


Figure 3 Cache Hit Ratio under different Zipf parameter α

Learning Model Generation Platform for Edge Computing), and by IITP grant funded by the Korea government (MSIT) (No.2015-0-00567, Development of Access Technology Agnostic Next-Generation Networking Technology for Wired-Wireless Converged Networks) *Dr. CS Hong is the corresponding author.

REFERENCES

- [1] Accessed: May 14, 2019. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>
- [2] B. Bai, L. Wang, Z. Han, W. Chen, and T. Svensson, "Caching based socially-aware D2D communications in wireless content delivery networks: a hypergraph framework," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 74-81, Aug. 2016.
- [3] M. N. Soorki, W. Saad, M. H. Manshaei, and H. Saidi, "Social Community-Aware Content Placement in Wireless Device-to-Device Communication Networks," [Online]. Available: arXiv:1808.02637
- [4] E. M. Daly and M. Haahr, "Social Network Analysis for Information Flow in Disconnected Delay-Tolerant MANETs," *IEEE Transactions on Mobile Computing*, vol. 8, no. 5, pp. 606-621, May 2009.
- [5] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: evidence and implications," in *IEEE Conference on Computer Communications (INFOCOM)*, New York, USA, pp. 126-134, Mar. 1999.
- [6] Y. Yang, Y. Zhang, L. Dai, J. Li, S. Mumtaz, and J. Rodriguez, "Transmission capacity analysis of relay-assisted device-to-device overlay/underlay communication," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 389-389, Feb. 2017.
- [7] M. Chen, Y. Hao, L. Hu, K. Huang, and V. K. N. Lau, "Green and Mobility-Aware Caching in 5G Networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8347-8361, Dec. 2017.
- [8] B. Chen and C. Yang, "Energy costs for traffic offloading by cache-enabled d2d communications," *IEEE Wireless Conference and Networking Conference*, Doha, Qatar, pp. 1-6, Apr. 2016.
- [9] D. Liu and C. Yang, "Energy Efficiency of Downlink Networks With Caching at Base Stations," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 907-922, Mar. 2016.
- [10] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no.3-4, pp. 279-292, 1992.
- [11] A. A. Hagberg, D. A. Schult and P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX," in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Gael Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11-15, Aug 2008
- [12] K. Thar, N. H. Tran, T. Z. Oo, C. S. Hong, "DeepMEC: Mobile Edge Caching Using Deep Learning," *IEEE Access*, vol. 6, no. 1, pp. 78260-78275, Dec. 2018.
- [13] M. Chen, W. Saad, C. Yin, and M. Debbah, "Echo State Networks for Proactive Caching in Cloud-Based Radio Access Networks With Mobile Users," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp.3520-3535, June 2017.