

# Energy-Optimal Edge Content Cache and Dissemination: Designs for Practical Network Deployment

Shao-Yu Lien, Shao-Chou Hung, Hsiang Hsu, and Der-Jiunn Deng

The authors offer energy-optimal edge content cache and dissemination designs for hot spot and rural areas. Exploiting Lyapunov optimization and optimum control theory, they demonstrate analytical foundations and principles of the proposed optimum designs. Engineering insights are also revealed through practical simulation investigations to extend the research frontier of energy-efficient edge content cache and dissemination.

## ABSTRACT

Cloud storage and computing, although it empowers ubiquitous and global information acquisition, analysis, and management, the latency performance and heavy traffic burdens at fronthaul/hackhaul links have emerged as critical concerns due to the scalability issue, which consequently drives the development of edge content cache and dissemination. Despite ample research on latency and traffic load reduction, an effective design of edge content cache and dissemination from the perspective of energy efficiency is imperative. Noting the fact that wireless service demands are inherently unbalanced over geographic regions, a practical cell planning strategy of operators' basic interests should deploy base stations massively in hot spot areas and sparsely in rural areas. A mobile user is therefore able to acquire desired contents through multiple base stations in a hot spot, while performing autonomous content cache and dissemination with other mobile users in rural areas. Taking this practical constraint into consideration, this article offers energy-optimal edge content cache and dissemination designs for hot spot and rural areas. Exploiting Lyapunov optimization and optimum control theory, we demonstrate analytical foundations and principles of the proposed optimum designs. Engineering insights are also revealed through practical simulation investigations to extend the research frontier of energy-efficient edge content cache and dissemination.

## INTRODUCTION

The paradigm of cloud storage and computing to collect and process data around the world at a logically unified cloud server has been demonstrated to empower ubiquitous knowledge acquisition and global information analysis/management [1]. With the aid of existing cellular infrastructure, this paradigm has been further extended to mobile users to access both location-based and non-location-based information in this decade. However, recent research has revealed severe inefficiency of cloud storage/computing due to scalability [2], and such inefficiency includes:

- Growing latency to access information stored/processed at databases whose physical locations could be anywhere on this planet

- Growing burdens on backhaul links of mobile networks to forward information between base stations and cloud servers
- Decreasing spectrum utilization at fronthaul links of mobile networks due to heavy burdens to forward information between base stations and mobile users

These thorny issues consequently drive the development of edge content cache and dissemination [3–5].

Caching frequently desired knowledge and location-based information at network edges (e.g., base stations or mobile users) has been regarded as a promising innovation to tackle the scalability issue so as to significantly alleviate information acquisition latency and traffic burdens at backhaul links [6]. In the meantime, network edges may further disseminate cached/obtained information/knowledge to other network edges. As information dissemination only occurs among network edges in physical proximity, spectrum could be fully reused in the spatial domain to considerably enhance the spectrum utilization at fronthaul links. Despite considerable discussions on the technical merits in terms of latency, traffic burdens at backhaul links, and spectrum utilization at fronthaul links, an optimum design from the perspective of minimizing overall energy consumption still remains open [7].

To minimize overall energy consumption, a crucial fact of network deployment is generally ignored. Eliminating all the potential coverage holes to offer ubiquitous network services is typically a primary objective in cell planning. However, this ideal strategy may largely harm overall energy consumption in mobile networks due to the fact that wireless service demands are not uniformly distributed over all geographic regions. In practice, network deployment of operators' utilitarian interests should treat hot-spot and rural areas with different strategies. In hot-spot areas with huge traffic demands, rich research results have revealed that more low-power base stations (or small cells), instead of a few macrocell base stations, should be massively deployed to facilitate low overall energy consumption. As a result, a mobile user located at overlapped coverage areas could possibly receive wireless services from multiple base stations. In this case, the

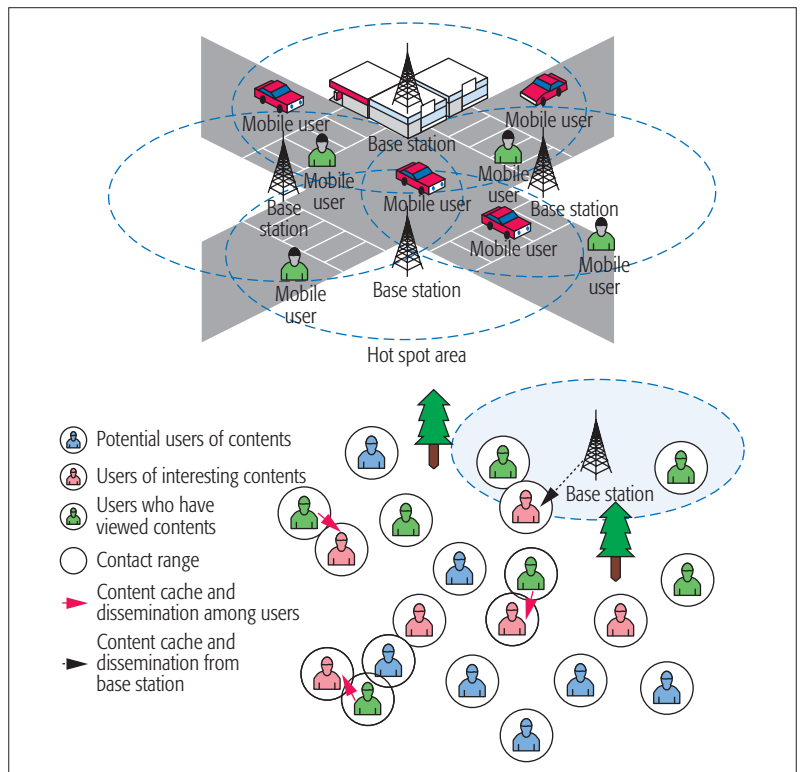
challenge of an energy-optimal content cache and dissemination design is to decide a minimum number of base stations caching different contents to service a mobile user such that the latency constraint of content access can be satisfied. On the other hand, in rural areas with very few traffic demands (e.g., a mountain district, a forest region, or an area with low population), sparse base station deployment is already satisfactory to avoid offering overcapacity of wireless services. In this case, mobile users autonomously caching and disseminating the contents obtained from base stations could be efficacious to supply wireless services in coverage holes. Nevertheless, the challenge of an energy-optimal design turns out to be optimally controlling the beginning of cache time at each mobile user and base station to launch further information dissemination.

Accordingly, the purpose of this article is to offer substantially energy-optimal designs for edge content cache and dissemination in practical network deployment. To this end, the application of the optimum control theory as well as Lyapunov drift theory together with social statistics and epidemic spreading to develop the proposed energy-optimal designs are elaborated. The provided knowledge thus paves substantial foundations to facilitate practical implementation of energy-optimal edge content cache and dissemination.

## PRINCIPLES OF EDGE CONTENT CACHE AND DISSEMINATION IN PRACTICAL NETWORK DEPLOYMENT

As aforementioned, from the operators' perspective, eliminating all the coverage holes may not be an efficient cell planning strategy, and a practical network deployment strategy is illustrated in Fig. 1. In hot-spot areas (e.g., urban areas, indoor environments of shopping malls or offices), operators are expected to deploy more base stations to supply more resources in both the fronthaul and backhaul links. Through caching location-based information or popular (frequently accessed) contents at these edge base stations, traffic loads at backhaul links and the core network can be relaxed. Since backlinks and the core network may involve a considerable number of switches, routers, and databases in creating connections from a mobile user to a remote cloud server, the amount of energy consumption to operate backhaul links and core network could be intractably large. Therefore, such an edge content cache framework could also alleviate the overall energy consumption in mobile networks through minimizing the utilization of backhaul links.

However, due to limited memory space, each base station may only cache a limited amount of information and knowledge, and cached information/knowledge can be provided to a mobile user if a mobile user requests it. In this case, to increase memory space to cache more information/knowledge, each base station may cache a part of information/knowledge that is identical with that at other base stations, while each base station may cache a part of information/knowledge that is disjoint from other base stations, as illustrated in Fig. 2. When these base stations are massively deployed, the coverage regions of multiple base stations could be overlapped, and a mobile user

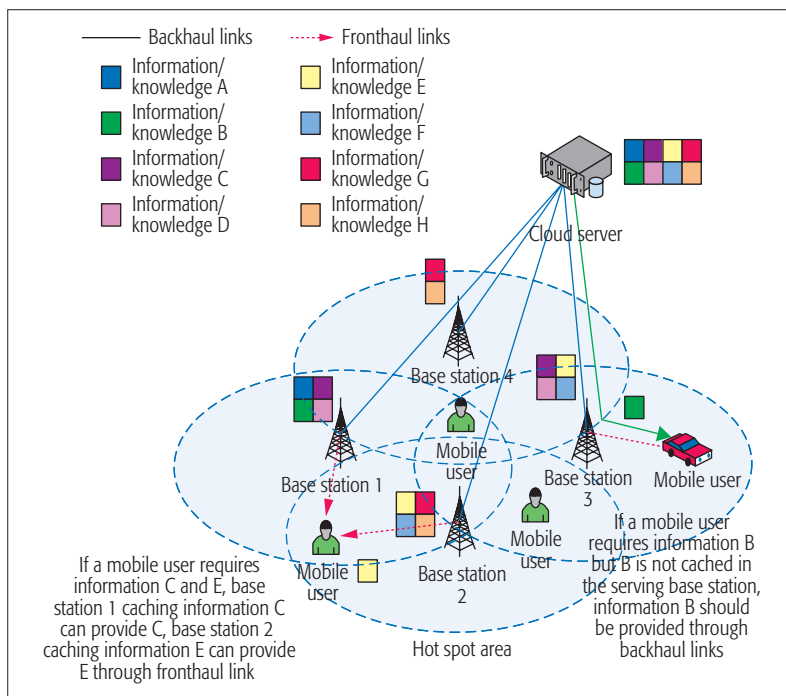


**Figure 1.** In practical network deployment of operators' utilitarian interests, base stations should be massively deployed in hot-spot areas, while sparse base station deployment should be adopted in rural areas.

located at overlapped coverage regions could receive wireless services from multiple base stations. As a result, if a mobile user requests a sort of information/knowledge while the requested information/knowledge is not cached in a base station, this mobile user can request information/knowledge from other base stations. Consequently, the more base stations a mobile user can connect to, the higher probability that a mobile user is able to obtain desired information/knowledge through fronthaul links (links between a mobile user and a base station). If, unfortunately, the requested information/knowledge is not cached at any base station to which a mobile user is able to connect, a base station should forward this request to the cloud service, and the requested information/knowledge is provided from the cloud server through backhaul links.

As aforementioned, utilizing backhaul links and the core network may consume intractably large energy. To minimize the latency to access the required information/knowledge, a straightforward strategy for a mobile user is to connect (send requests to) as many base stations as possible so as to maximize the probability to obtain required information/knowledge. However, this strategy may increase another sort of energy consumption at fronthaul links. An energy-optimal design is consequently that each mobile user only connects to a minimum number of base stations such that the latency constraint to obtain desired information/knowledge can be satisfied. An analytical design derived from the Lyapunov drift theory is introduced later.

On the other hand, in rural areas (e.g., mountain/forest regions with low population), an ener-



**Figure 2.** In hot-spot areas with massive base station deployment, a mobile user is able to connect to more than one base station, and each base station may cache different sets of information/knowledge.

gy-efficient design is to deploy only a limited number of base stations to save energy at fronthaul links. Nevertheless, it is likely that a mobile user is not able to connect to a base station to obtain desired information/knowledge all the time. To facilitate information/knowledge dissemination, a mobile user is able to cache contents when this mobile user is able to enjoy wireless services provided from a base station. Subsequently, when this mobile user moves out from the coverage region of a base station, this mobile user is able to provide cached contents to other mobile users in need, and these mobile users also cache the obtained contents to further provide to others. To minimize energy consumption at each mobile user, content dissemination only occurs when mobile users are in physical proximity. In this case, the behavior of information dissemination is similar to disease epidemics [8, 9], as illustrated in Fig. 1.

To boost content dissemination, each mobile user should cache as many contents as possible to offer ample opportunities for an out-of-network-coverage mobile user to obtain contents from a neighboring mobile user. However, this design may not be allowed due to limited memory space in each mobile user. This critical limitation further makes an appropriate beginning time instance for each mobile user to perform content cache a crucial issue. Since a mobile user is able to share contents only if contents have been cached, a mobile user performing content cache at a very early time may benefit content dissemination among mobile users. However, this scheme may spend much energy of a mobile user since a mobile user should share contents once a mobile user has obtained contents. Furthermore, when a mobile user has cached contents, the storage capacity of this mobile user cannot store

other contents. On the other hand, if all mobile users delay the beginning time to perform caching, contents cannot be fully disseminated among mobile users during a desired time duration. As a result, an energy-optimal design can be used to minimize the utilization of fronthaul links and the number of mobile users who want but are still waiting for contents during a desired time duration. An analytical design derived from the optimum control theory [10] is introduced later.

## ENERGY-OPTIMAL DESIGN IN HOT-SPOT AREAS

For each mobile user, the objective is to obtain desired information/knowledge on time either through connecting to multiple base stations to increase the probability of successfully accessing desired information/knowledge from base stations, or through connecting to only one base station to obtain desired information/knowledge from the cloud server. An energy design thus can be analytically formulated as an optimization, in which the objective is to minimize the average number of connected base stations for each mobile user over time, so as to minimize the energy consumption at fronthaul links. In the meantime, two constraints should be satisfied.

### Average Amount of Information/Knowledge Acquisition from the Cloud Server over Time Should Not Exceed a Pre-Defined Threshold:

Since an accurate analysis on the amount of energy consumption through utilizing backhaul links and core network may not be practically tractable, this constraint is equivalent to controlling energy consumption at backhaul links and the core network not to exceed a certain threshold.

### Queue Length of Each Mobile User Cannot Exceed a Certain Threshold over Time:

When a mobile user requires information/knowledge to process a task, this task stays in a queue, and this task can be removed from a queue only if the required information/knowledge can be obtained. If this constraint can be satisfied at a mobile user, we say that this mobile user is stabilized.

The goal of this design is consequently to decide the number of base stations that each mobile user is permitted to connect to at any time instant, and the amount of information/knowledge permitted to be downloaded from the cloud server at any time instant, so as to solve the optimization.

In this optimization, the objective and two constraints should be tackled asymptotically in the time domain through dynamically determining two parameters as stated above at each time instant. Such a problem is a sort of “dynamic optimization,” and in the literature, the Markov decision process (MDP) [11] and differential evolution (DE) are widely discussed as general methods, especially to strike the trade-off between resource utilization and delay. However, when the numbers of base stations and mobile users grow, the model of MDP and DE could become sophisticated, and the number of iterations could also exponentially increase. In fact, there are straightforward methods to satisfy the two constraints; for example, permitting a mobile user with an unacceptable queue length to connect to more base stations or download more information/knowledge from the

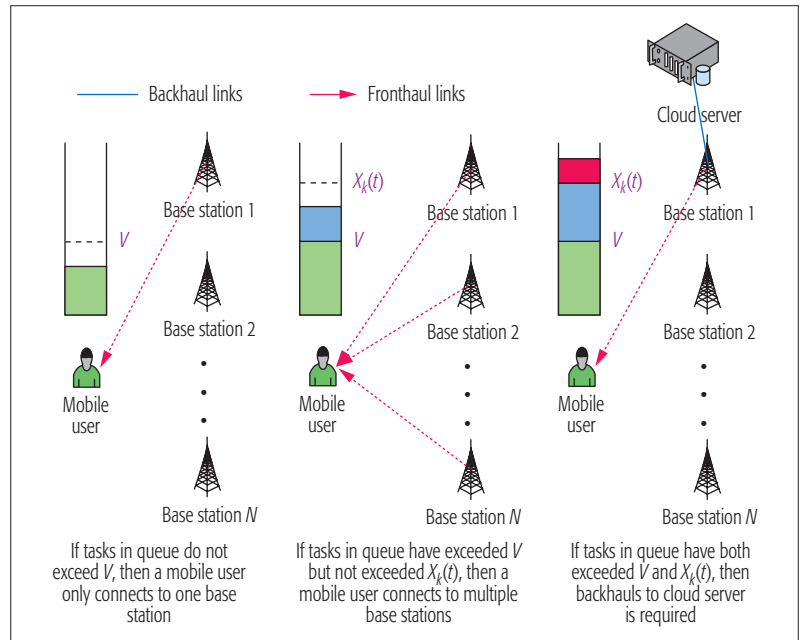


cloud server. Although this facilitates latency performance of information/knowledge acquisition for each mobile user, the result could largely deviate from the optimum solution. In other words, there is a trade-off between resource utilization and network performance. This engineering issue consequently motivates a recent innovation known as *Lyapunov optimization*.

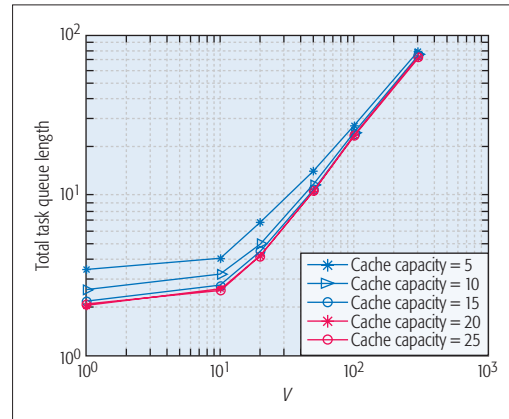
Lyapunov optimization, originating from *Lyapunov drift theory* [12], introduces the *drift-plus-penalty* concept to develop a dynamic control algorithm. Through properly defining a drift-plus-penalty function, the dynamic control algorithm maximizes the drift-plus-penalty function subject to the queue stability at each mobile user. To define a drift-plus-penalty function, two parameters are introduced: a virtual queue of each mobile user (denote  $X_k(t)$  as the virtual queue length of the  $k$ th mobile user at time  $t$ ) and a constant cost  $V$  to utilize multiple base stations. As a result, the penalty of utilizing backhaul links, the cost of connecting to multiple base stations (fronthaul links), and the queue length of each mobile user are jointly taken into account in a drift-plus-penalty function. When a drift-plus-penalty function can be analytically formulated, the energy-optimal design can be analytically reformulated as a new optimization to maximize the drift-plus-penalty function. Instead of analytically providing a drift-plus-penalty function for our design, an illustration as shown in Fig. 3 can offer engineering insights to comprehend the dynamic operations of the optimization.

As shown in Fig. 3, both  $X_k(t)$  and  $V$  can be conceptually regarded as particular thresholds. When the task queue length of the  $k$ th mobile user is below  $V$ , this mobile user only connects to a single base station. If this base station fortunately caches information/knowledge required by the mobile user, the task queue length of the mobile user may not grow; otherwise, the task queue length could increase. When the task queue length exceeds  $V$ , the mobile user connects to multiple base stations. In practice, there can be multiple  $V$  values associated with different numbers of base stations. When the task queue length exceeds a higher  $V$  value, the mobile user is able to connect to more base stations. If the task queue length continues growing to exceed  $X_k(t)$ , the mobile user is permitted to download information/knowledge from the cloud server through backhaul links.

In Figs. 4 and 5, simulation results of the optimum solutions in the proposed design using Lyapunov optimization are provided. To simulate the dynamic of the number of base stations, the number of connectable base stations at time  $t + 1$  is modeled as the difference between the number of connectable base stations at time  $t$  and the number of base stations that are out of network coverage to a mobile user at time  $t$ , plus the number of base stations newly become connectable to a mobile user at time  $t$ . The deployment (locations) of base stations follows a homogeneous Poisson point process with the density  $5 \times 10^{-5}$ . A mobile user is able to receive wireless services from a base station if the distance between a mobile user and a base station is less than 150 m, and the velocity of each mobile device is set to be 20 m/s.



**Figure 3.** If the task queue length does not exceed  $V$ , a mobile user only connects to one base station. If the task queue length has exceeded  $V$  but not exceeded  $X_k(t)$ , a mobile user connects to more than one base station. If the task queue length has exceeded both  $V$  and  $X_k(t)$ , backhauls to cloud server are required.



**Figure 4.** Simulation results of the total task queue length of all mobile users under different cache capacities and  $V$  values.

Figure 4 shows the total (summation of) task queue length of all mobile users under different cache capacities (i.e., amount of information/knowledge cached) denoted by  $M$  in each base station. 10 information/knowledge items with different popularities ranking from 1 to 10 are considered in this simulation, and each base station randomly chooses  $M$  from 10 information/knowledge items according to a Zipf distribution. We can observe from Fig. 4 that the total task queue length increases as  $V$  grows. When the cost to connect to multiple base stations (i.e.,  $V$ ) increases, each mobile user may avoid connecting to more base stations to alleviate the cost. As a result, a mobile user may not obtain required information/knowledge from base stations to lead to a long task queue length. On the other hand, a larger cache capacity in each base station is able to accommodate a variety of information/knowledge. Therefore, a mobile user may have a higher

Caching at the base station can be regarded as a special case of autonomous content cache among mobile users in which there is no inter-user content dissemination. As a result, both with and without autonomous content, cache and dissemination among mobile users can be treated as a general case to be solved using a common design.

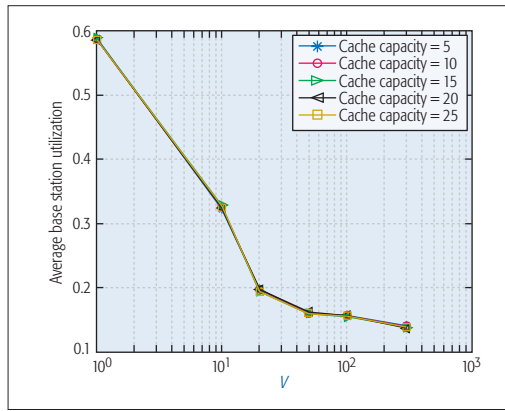


Figure 5. Simulation results of the average base station utilization under different  $V$  values.

chance to obtain required information/knowledge through only connecting to a few base stations. Consequently, a larger cache capacity in each base station facilitates a shorter task queue length in a mobile station.

In Fig. 5, the average base station utilization under different  $V$  values are shown. Similarly, when the cost to connect to multiple base stations increases, each mobile user may avoid connecting to more base stations to alleviate the cost, leading to low utilization at each base station.

### ENERGY-OPTIMAL DESIGN IN RURAL AREAS

To analyze the complex and opportunistic mechanisms of edge content cache and dissemination among mobile users, non-linear ordinary differential equations (ODEs) of epidemic spreading provide an adequate model capturing the random contacts and time dynamic relationship among mobile users. Particularly, the susceptible-infected-recovered (SIR) model receives considerable attention to be utilized as the state evolutionary equations for the control process [13]. To bridge the analogy between the SIR model and composition of mobile users in rural areas,  $S(t)$ ,  $I(t)$ , and  $R(t)$  are denoted as the susceptible, infected, and recovered population of mobile users at time  $t$ , respectively. The susceptible mobile users in  $S(t)$  denote the potential viewers of the contents, the infected mobile users in  $I(t)$  denote contagious users who need but still waiting for the contents, and the recovered mobile users in  $R(t)$  denote those who have been served with the contents and should no longer request the contents any more.

**Without Autonomous Content Cache and Dissemination among Mobile Users:** In this case, the only way that an infected mobile user becomes a recovered mobile user is to receive wireless service from a base station. Denote the time instant that a base station begins caching the contents from the cloud server as  $T_C$ . At the time prior to  $T_C$ , the wireless services provisioning from a base station to an infected mobile user relies on both backhaul and fronthaul links, as the contents are in fact provided from the cloud server. We denote the service rate provided at the fronthaul link at time  $t$  as  $u(t) = \kappa_1$  when  $t < T_C$ . After the time instant  $T_C$ , the base station has performed content cache, and therefore backhaul links are

not required. At this moment, the service rate provided at the fronthaul link at time  $t$  is denoted as  $u(t) = \kappa_2$  when  $t \geq T_C$ . Generally,  $\kappa_1 \leq \kappa_2$ , since the service rate at fronthaul links are subject to the service rate at backhaul links before caching takes place. However, when caching the contents at a base station, the service rate at fronthaul links increases due to the mitigation of traffic load on backhaul by caching.

**With Autonomous Content Cache and Dissemination among Mobile Users:** We also denote the time instant that a mobile user begins caching the contents as  $T_C$ . After the time instant  $T_C$ , in addition to obtaining the contents from a base station, an infected mobile user also acquires the contents from a recovered mobile user when they are within a contact region. In this case, an infected mobile user becomes a recovered mobile user.

When an infected mobile user contacts a susceptible mobile user within a contact region, there is a viral rate  $v$  that a susceptible mobile user interested in accessing the contents becomes an infected mobile user.

Caching at the base station can be regarded as a special case of autonomous content cache among mobile users in which there is no inter-user content dissemination. As a result, both with and without autonomous content, cache and dissemination among mobile users can be treated as a general case to be solved using a common design. In our energy-optimal design, there are two kinds of costs. The first kind of cost is related to the time duration of caching (i.e.,  $T_C$ ), since in a caching duration the store capacity for caching the contents cannot be used to cache other contents. The second kind of cost is  $u(t)$ , which also depends on  $T_C$ . If we delay  $T_C$ , although the storage capacity in base stations or mobile users can be vacated, backhaul links could be utilized for a long time to largely consume energy at backhaul links and the core network. Besides the costs, an efficient design of caching also depends on the number of recovered mobile users. Initially, there is no recovered user, causing scarce sharing and inefficient caching. Nevertheless, the number of recovered mobile users should increase in time and hence enhance the content sharing. The above concerns thus render the optimum control of  $T_C$  a critical issue.

Targeting at deriving an energy-optimal design, an optimization can be formulated with the facilitation of the optimum control theory [10] to jointly consider the costs and efficiency of epidemic content dissemination (i.e., the cumulative number of infected mobile users  $S(t)$  within a desired time duration  $T_f$  should be minimized),

$$T_C^* = \arg \min_{T_C} \int_0^{T_f} I(t)^\beta + \frac{1}{\alpha} u(t)^\alpha dt \quad (1)$$

where  $\alpha \geq 0$  and  $\beta \geq 0$  are two weighting parameters for  $u(t)$  and  $I(t)$ , respectively. A larger  $\alpha$  indicates a larger cost of using fronthaul/backhaul links.  $\beta$  represents the requirement in system dynamics, in which a larger  $\beta$  and therefore larger  $I(t)^\beta$  indicates fewer unserved users in the system before the expiration of  $T_f$  are desired. Equation 1 can thus be solved through utilizing the general methodology provided by the optimum control theory.

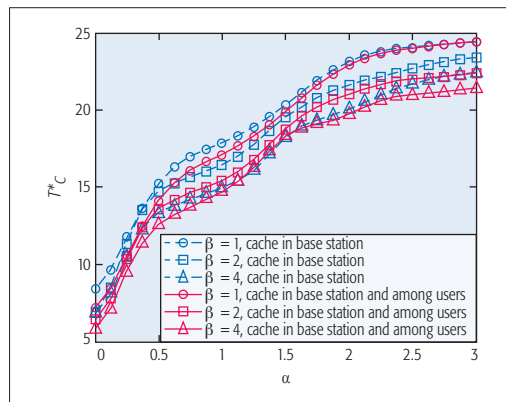
In Fig. 6, the simulation results of the optimum TC under different  $\alpha$  and  $\beta$  are demonstrated. Pertaining to this simulation setup,  $N$  nodes are deployed to travel in a square area with wrap-around condition via the L  vy walk mobility model to capture human moving behavior [14]. The step size and pause time are accounted for by a power law distribution with a negative exponent. We set the step size exponent to be 1.5 and the pause time exponent to be 1.38, to fit the real trace-based data collected in [15]. As shown in Fig. 6, the optimal caching time increases with the growth of  $\alpha$ . If there are many contents circulating in the network, a longer time is needed to decide whether to cache the viral contents in order to optimize the usage of storage capacity at the base station and mobile users. Nevertheless, since the number of mobile users who have viewed the contents grows by time, late caching time implies a better chance to efficaciously utilize caching among the mobile users for sharing. However, as we wish the number of mobile users waiting for the content to be as small as possible, the optimal caching time becomes early to handle the requirement. These two factors form a trade-off in designing energy-optimal edge content cache and dissemination in rural areas.

## CONCLUSION

In this article, analytical foundations of energy-optimal designs for edge content cache and dissemination in hot-spot and rural areas are offered. In hot-spot areas, with the facilitation of Lyapunov drift theory and information/knowledge cache in base stations, each mobile user only receives wireless services from a minimum amount of base stations such that the task queue length can be stabilized. The proposed design consequently minimizes energy consumption at fronthaul links and controls energy consumption at backhaul links. On the other hand, in rural areas, with the facilitation of optimum theory, content cache in both base stations and mobile users, and the SIR model for epidemic content sharing, the method to derive the optimum cache time is elaborated. The proposed design therefore strikes the optimum trade-off between the utilization (and hence energy consumption) of fronthaul/backhaul links and the efficiency of content dissemination (i.e., minimizing the number of unserved mobile users within a desired time duration). The proposed designs are particularly tailored for practical cell planning strategy, to supply relevant knowledge in implementation and research.

## REFERENCES

- [1] G. Pallis, "Cloud Computing: The New Frontier of Internet Computing," *IEEE Internet Comp.*, vol. 14, no. 5, Sept. 2010, pp. 70–73.
- [2] S.-C. Hung et al., "Architecture Harmonization Between Cloud Radio Access Networks and Fog Networks," *IEEE Access*, vol. 3, Dec. 2015, pp. 3019–34.
- [3] W. Wang et al., "Edge Caching at Base Stations with Device-to-Device Offloading," *IEEE Access*, vol. 5, Mar. 2017, pp. 6399–6410.
- [4] D. Liu et al., "Caching at the Wireless Edge: Design Aspects, Challenges, and Future Directions," *IEEE Commun. Mag.*, vol. 54, no. 9, Sept. 2016, pp. 22–28.
- [5] D.-J. Deng et al., "Latency Control in Software-Defined Mobile-Edge Vehicular Networking," *IEEE Commun. Mag.*, vol. 55, no. 8, Aug. 2017, pp. 87–93.
- [6] H. Hsu and K.-C. Chen, "A Resource Allocation Perspective on Caching to Achieve Low Latency," *IEEE Commun. Lett.*, vol. 20, no. 1, Jan. 2016, pp. 145–48.



**Figure 6.** Optimum cache time  $T_C^*$  under different cost configurations in terms of  $\alpha$  and  $\beta$ , where 1000 mobile users are deployed to a 10,000 m<sup>2</sup> area with  $T_f = 150$ ,  $\kappa_1 = 0.1$ ,  $\kappa^2 = 0.2$ ,  $v = 1$ , and  $u(t) = 0.15$  when content cache and dissemination is allowed among mobile users.

- [7] C. Fang et al., "A Survey of Energy-Efficient Caching in Information-Centric Networking," *IEEE Commun. Mag.*, vol. 52, no. 11, Nov. 2014, pp. 122–29.
- [8] D. J. Daley, J. Gani, and J. M. Gani, *Epidemic Modeling: an Introduction*, Cambridge Univ. Press, 2001, vol. 15.
- [9] L. Weng, F. Menczer, and Y.-Y. Ahn, "Virality Prediction and Community Structure in Social Networks," *Scientific Reports*, vol. 3, 2013.
- [10] D. E. Kirk, *Optimal Control Theory: An Introduction*, Courier Corp., 2012.
- [11] E. Stevens-Navarro, Y. Lin, and V. W. S. Wong, "An MDP-Based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Network," *IEEE Trans. Vehic. Tech.*, vol. 57, no. 2, Mar. 2008, pp. 1243–54.
- [12] Y. Cui et al., "A Survey on Delay-Aware Resource Control for Wireless Systems — Large Deviation Theory, Stochastic Lyapunov Drift, and Distributed Stochastic Learning," *IEEE Trans. Info. Theory*, vol. 58, no. 3, Mar. 2012, pp. 1677–1701.
- [13] S. Meyn, *Control Techniques for Complex Networks*, Cambridge Univ. Press, 2008.
- [14] I. Rhee et al., "On the Levy-Walk Nature of Human Mobility," *IEEE/ACM Trans. Net.*, vol. 19, no. 3, 2011, pp. 630–43.
- [15] S. Kim, C.-H. Lee, and D. Y. Eun, "Superdiffusive Behavior of Mobile Nodes and Its Impact on Routing Protocol Performance," *IEEE Trans. Mobile Comp.*, vol. 9, no. 2, 2010, pp. 288–304.

## BIOGRAPHIES

SHAO-YU LIEN is with the Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan. He received the IEEE Communications Society Asia-Pacific Outstanding Paper Award 2014, the Scopus Young Researcher Award (issued by Elsevier) 2014, the URSI AP-RASC 2013 Young Scientist Award, and the IEEE ICC 2010 Best Paper Award. His research interests include LTE Pro, 5G New Radio, cyber-physical systems, and configurable networks.

SHAO-CHOU HUNG received his Ph.D. from the Department of Electrical Engineering at National Taiwan University in 2017. He is currently a senior algorithm engineer at MediaTek. His research focuses on statistical learning, dynamic optimal control, and its applications on 5G New Radio and V2X vehicular networks.

HSIANG HSU received his B.S. degrees in electrical engineering and mathematics, and M.Eng. in communication engineering from National Taiwan University. He is currently pursuing a Ph.D. degree at SEAS, Harvard University. His research interests include resource allocation, optimization, and machine learning in complex networks.

DER-JIUNN DENG joined the National Changhua University of Education (NCUE) as an assistant professor in the Department of Computer Science and Information Engineering in August 2005 and became a Distinguished Professor in August 2016. In 2012 and 2015, he received the Outstanding Faculty Research Award of NCUE. His research interests include multimedia communication and wireless networks.

The proposed design strikes the optimum trade-off between the utilization of fronthaul/backhaul links and the efficiency of content dissemination. The proposed designs are particularly tailored for the practical cell planning strategy, to supply relevant knowledge in implementation and research.