

User Preference Learning Based Edge Caching for Fog Radio Access Network

Yanxiang Jiang, *Senior Member, IEEE*, Miaoli Ma, Mehdi Bennis, *Senior Member, IEEE*,
Fu-Chun Zheng, *Senior Member, IEEE*, and Xiaohu You, *Fellow, IEEE*

Abstract—In this paper, the edge caching problem in fog radio access network (F-RAN) is investigated. By maximizing the overall cache hit rate, the edge caching optimization problem is formulated to find the optimal policy. Content popularity in terms of time and space is considered from the perspective of regional users. We propose an online content popularity prediction algorithm by leveraging the content features and user preferences, and an offline user preference learning algorithm by using the online gradient descent (OGD) method and the follow the (proximally) regularized leader (FTRL-Proximal) method. Our proposed edge caching policy not only can promptly predict the future content popularity in an online fashion with low complexity, but also can track the content popularity with spatial and temporal popularity dynamic in time without delay. Furthermore, we design two learning based edge caching architectures. Moreover, we theoretically derive the upper bound of the popularity prediction error, the lower bound of the cache hit rate, and the regret bound of the overall cache hit rate of our proposed edge caching policy. Simulation results show that the overall cache hit rate of our proposed policy is superior to those of the traditional policies and asymptotically approaches the optimal performance.

Index Terms—F-RAN, edge caching, cache hit rate, content popularity, user preference.

Manuscript received April 15, 2018, revised September 18, 2018, and accepted November 2, 2018. Part of this work has been presented at the 6th IEEE GLOBECOM Workshop on Emerging Technologies for 5G and Beyond Wireless and Mobile Networks (ET5GB), Singapore, December, 2017. The associate editor coordinating the review of this paper and approving it for publication was T. He.

This work was supported in part by the Natural Science Foundation of Jiangsu Province under grant BK20181264, the Research Fund of the State Key Laboratory of Integrated Services Networks (Xidian University) under grant ISN19-10, the Research Fund of the Key Laboratory of Wireless Sensor Network & Communication (Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences) under grant 2017002, the National Basic Research Program of China (973 Program) under grant 2012CB316004, and the U.K. Engineering and Physical Sciences Research Council under Grant EP/K040685/2.

Y. Jiang is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China, and the Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, 865 Changning Road, Shanghai 200050, China (e-mail: yxjiang@seu.edu.cn).

M. Ma and X. You are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: milies10@126.com, xhyu@seu.edu.cn).

M. Bennis is with the Centre for Wireless Communications, University of Oulu, Oulu 90014, Finland (e-mail: mehdi.bennis@oulu.fi).

F. Zheng is with the School of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen 518055, China, and the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China. (e-mail: fzheng@ieee.org).

I. INTRODUCTION

With the continuous and rapid proliferation of various intelligent devices and advanced mobile application services, wireless networks have been suffering from an unprecedented data traffic surge in recent years. Although centralized cloud caching and computing in cloud radio access network (C-RAN) can provide reliable and stable service for end users during off-peak periods [1], [2], ever-increasing mobile data traffic brings tremendous pressure on C-RAN with capacity-limited fronthaul links and centralized baseband unit (BBU) pool, which may cause communication interruptions or traffic congestions especially at peak traffic moments. The main reason is that as various social applications become more and more popular, data traffic over fronthaul links surges with a lot of redundant and repeated information, which further worsens the fronthaul constraints. In this case, a feasible solution is to shift a small amount of resources such as communications, computing, and caching to network edge, and enable most of the frequently requested contents being served locally. At this point, fog radio access network (F-RAN) as a complementary network architecture was proposed, which can effectively reduce fronthaul load by placing most popular contents closer to the requesting users and extending traditional cloud computing paradigm to the network edge [3]–[5]. Up until now, F-RAN has attracted more and more attention from researchers and engineers. In F-RAN, traditional access points are turned into fog access points (F-APs) equipped with limited caching and computing resources, where edge caching is a key component to improve the performance of F-RAN. Due to storage constraints and fluctuating spatio-temporal traffic demands, however, F-APs face a myriad of challenges. For example, how, what and when to strategically store contents in their local caches in order to achieve a higher cache hit rate?

Traditional caching policies such as first in first out (FIFO) [6], least recently used (LRU) [7], least frequently used (LFU) [8] and their variants [9] have been widely applied in wired networks, where there are abundant caching and computing resources and the served area is usually very large. However, these traditional caching policies may not be applied well in wireless networks due to the characteristics of edge nodes such as smaller coverage areas and limited caching resources, and they may suffer major performance degradation since they may not be able to predict future content popularity correctly. Most of the existing works on edge caching in wireless networks assumed that the content popularity is already known or

subject to a uniform distribution, and then focused on exploring the optimal caching policy under the above assumptions [10]–[12]. The edge caching problem was formulated as a many-to-many matching game between small base stations and service providers' servers in [10], it was converted to an approximation facility location problem in [11], and the successful transmission probability was maximized to obtain a local optimal caching and multicasting design in a general region in [12]. However, these studies are inconsistent with the reality. Generally, content popularity is not as traceable as a unified distribution no matter what kind of caching policy is applied.

By considering the time-varying nature of contents, recent works have turned to exploring sophisticated edge caching policies by learning future content popularity. In [13], the content popularity matrix was estimated through transfer learning by leveraging user-content correlation and information transfer between time periods. In [14], the training time for transfer learning was analyzed for obtaining a better estimation performance. Nevertheless, the content popularity matrix remains typically to be large, which needs a great deal of calculation in the estimation. Meanwhile, the transfer learning approach has a poor performance for the case of low information correspondence ratio. In [15], the cache content placement problem was modeled as a contextual multi-arm bandit problem and an online policy was presented to learn the content popularity. This policy learns the content popularity independently across contents whereas ignores the content similarity and the impact of user preference on content popularity, thereby resulting in high training complexity and slow learning speed. A low complexity online policy was proposed in [16], where content popularity was learned based on the assumption that the expected popularities of similar contents are similar. It performs well for video caching but may be ineffective for other types of content caching. However, all the above studies assume that the content popularity remains unchanged for a certain time period and the content library is stationary, and ignore to consider spatial changes of content popularity, and therefore cannot truly reflect the changes of content popularity. In real communications scenarios, both the coverage area of an edge node and the number of users that it can serve are limited. The change of content popularity in the time and space dimensions is real-time. Due to the continuous emergence of new contents, the content library in the cloud content center in the time dimension must change. Moreover, the set of current users in a specific region or scenario may change dynamically over time due to user mobility, and the content popularity may thus fluctuate too. Furthermore, due to the randomness of user requests, the content popularity will be dynamically changing over time. In practice, different users may have different degrees of interest, i.e., different user preferences, in the same content. Correspondingly, the request possibilities for the same content among different regions or scenarios gathered by users with different user preferences are different. This will result in a content popularity difference for the same content among different regions or scenarios. These changing factors make it impossible to measure the content popularity merely through a unified distribution or a

simple prediction. The small changes of the content popularity directly affect the caching decisions. Therefore, it is necessary to explore spatial and temporal dynamic of content popularity and track the dynamic in a timely manner for the ensurance of continuously caching popular contents, the achievement of optimal caching decisions, and the maximization of cache hit rate.

Motivated by the aforementioned discussions, our main contributions are summarized below.

- 1) We put forward a new idea of content popularity prediction. Unlike the static approach, we consider content popularity in terms of time and space from the perspective of regional users and propose an online content popularity prediction algorithm, which can predict the future content popularity of a certain region in an online fashion without any restriction on content types and track the popularity change in real time.
- 2) We propose an offline user preference learning algorithm, which can discover the user's own preference through its historically requested information. By monitoring the average prediction error in real time, it can be initiated automatically for relearning of user preference and continuous offline learning can thus be avoided.
- 3) We design two learning based edge caching architectures for F-RAN. They differ in that the offline user preference learning functionality is transferred from the F-APs in the first architecture to the smart user equipments (UEs) in the second one. Specifically, we introduce cooperative caching among regional F-APs and make more efficient usage of limited caching resources. Besides, we introduce a new module to enable regular monitoring of regional users by considering the impact of user mobility on cache decisions.
- 4) We analyze the performance of our proposed edge caching policy. We first derive the upper bound of the popularity prediction error of our proposed online content popularity prediction algorithm, and reveal the sub-linear relationship between the cumulative prediction error and the total number of content requests. We then derive the regret bound of the overall cache hit rate of our proposed edge caching policy, and show through theoretical analysis that our proposed policy has the capability to achieve the optimal performance asymptotically.
- 5) We validate our theoretical results with real data. Simulation results show that our proposed edge caching policy can predict the content popularity with high precision, and track the real local popularity in real time with spatial and temporal popularity dynamic. Simulation results also show the superior performance of our proposed policy in comparison with the other traditional policies and verify its asymptotical optimality.

The rest of this paper is organized as follows. In section II, the system model is described. Our proposed edge caching policy including online content popularity prediction and offline user preference learning is presented in Section III. The two learning based edge caching architectures are described in



Fig. 1. Illustration of the edge cache scenarios in F-RAN.

Section IV. The performance analysis is provided in Section V. Simulation results are shown in Section VI. Final conclusions are drawn in Section VII.

II. SYSTEM MODEL

A. Edge Caching Scenarios in F-RAN

We consider the edge caching scenarios in F-RAN as illustrated in Fig. 1. Large amounts of F-APs with limited storage capacity are deployed in different scenarios, for example, a town with dense crowds and moderate user mobility, a park with relatively dense crowds and high user mobility, or a stadium with ultra-dense crowds and relatively low user mobility. Due to different user density and user mobility in different scenarios, there exist differences in the distribution of popular contents among different scenarios. In order to ensure a stronger scenario suitability of the caching policy, we propose the following edge caching design rule: The F-APs deployed in different scenarios should regularly monitor the users in consideration of user mobility, and set a monitoring cycle matched with the characteristics of different scenarios.

In each edge caching scenario, according to the corresponding design criteria (for example, location), the F-APs with close distance can cooperate and belong to the same region, which are also called regional F-APs. The users in the same region, also called regional users, can request contents of interest from their associated regional F-APs if the contents are stored in the local caches or from neighboring F-APs in the other regions or the cloud content center through fronthaul links otherwise. Then, we focus on the caching policy for a single region, namely the regional caching.

B. Edge Caching Problem Formulation

We consider the edge caching problem in a specific region served by M F-APs, which constitute an F-AP set $\mathcal{M} = \{1, 2, \dots, m, \dots, M\}$. It is assumed that the regional users can fetch the contents of interest from the local caches of the M F-APs. Without loss of generality, we assume that all the contents have the same size¹ and each F-AP has the

same storage capacity and can store up to φ contents from the content library $\mathcal{F} = \{1, 2, \dots, f, \dots, F\}$, which may vary over time and is located in the cloud content center. Without loss of generality, assume $M\varphi \ll F$. By considering user mobility and according to the preset monitoring period, the F-APs monitor the users in the specific region during discrete time periods $t = 1, 2, \dots, T$, where T is set to be a finite time horizon. Let U_t denote the number of regional users served by the M F-APs during the t th time period with $U_{\min} \leq U_t \leq U_{\max}$, and $\mathcal{U}_t = \{1, 2, \dots, U_t\}$ the set of regional users monitored by the M F-APs. It is assumed that the regional users remain unchanged during the considered time period. In the way of cache content placement, for description convenience, we adopt the partition-based content placement method in [18], where each content is separated into M equal-sized subfiles, and the F-APs store its different subfiles.^{2 3 4} We remark here that more sophisticated content placement methods can be adopted.

Let D_t denote the number of requests during the t th time period, $D = \sum_{t=1}^T D_t$ the overall number of requests in the finite time horizon T , and $req_t = \{req_{t,1}, req_{t,2}, \dots, req_{t,d}, \dots, req_{t,D_t}\}$ the set of requests which come in sequence during the t th time period. The request $req_{t,d}$ can be further expressed as $req_{t,d} = \langle f(d), t(d), \mathbf{x}(d) \rangle, \forall 1 \leq t \leq T, \forall 1 \leq d \leq D_t$, where $f(d) \in \mathcal{F}$ denotes the requested content, $t(d)$ the requesting time, and $\mathbf{x}(d) \in \mathbb{R}^N$ the feature vector with dimension N describing the features of the requested content. Take movie as an example: $\mathbf{x}(d)$ may include features like the movie rating, the movie type, the keyword frequency of movie critics, etc. Without loss of generality, we normalize the various dimensions of $\mathbf{x}(d)$ and set $\mathbf{x}(d) \in [0, 1]^N$.

During the t th time period, for each arriving request $req_{t,d}$, the regional F-APs first check whether $f(d)$ has been stored locally. Let $\theta_{t,d}(f(d)) \in \{0, 1\}$ denote the cache status of $f(d)$ at the requesting time $t(d)$, where $\theta_{t,d}(f(d)) = 1$ if $f(d)$ is stored locally, and $\theta_{t,d}(f(d)) = 0$ otherwise. If $f(d)$ has been stored in the local caches, a cache hit happens and the requesting user can then be served locally. Otherwise, a cache miss happens and $f(d)$ will be fetched from neighboring F-

²Note that it does incur the cost for transferring data among the regional F-APs.

³The M regional F-APs can indeed be treated as a single cache by adopting the simple partition-based content placement method. In this regard, the intelligent caching policies on single cache such as [13]–[16] can be applied here, where transfer learning, contextual multi-arm bandit, and trend-aware online learning were employed, respectively. However, just as stated previously, these intelligent caching policies have made strict assumptions and ignored to consider spatial dynamic of content popularity.

⁴On the other hand, however, the considered M regional F-APs are not simply treated as a single cache. In comparison with the single-cache setting, the cooperative-cache setting here has the following advantages: 1) By separating each content into M equal-sized subfiles and exploiting cooperation, the transmission delay can be decreased. Specifically, for video content, users may just want to watch part of it. Correspondingly, only some of the subfiles will need to be transmitted. Therefore, the transmission delay can be further decreased. 2) Furthermore, the coverage area of multiple F-APs and the number of their served users are relatively large. Correspondingly, popular contents can be more “concentrated” (i.e. in the coverage area). Therefore, a higher cache hit rate can be achieved. 3) Besides, more contents can be cached in the cooperative F-APs in a specific region, and unnecessary caching redundancy can be avoided.

¹Note that contents with different sizes can always be split into data segments of the same size, and each data segment can then be considered as a “content”. This is a common practice in real world systems. Here we follow the same assumption and justification as in [15] and [17].

APs in the other regions or the cloud content center, and a caching decision will be further made to determine whether to store $f(d)$ locally. If the F-APs decide to store $f(d)$ and replace one of the existing contents in the local caches, denoted by $f_{\text{old}} \in \{f | \theta_{t,d}(f) = 1, \forall f \in \mathcal{F}\}$, a cache update happens. Then, the cache status will be updated according to the following rule

$$\theta_{t,d+1}(f) = \begin{cases} 0, & \text{if } f = f_{\text{old}}, \\ 1, & \text{else if } f = f(d), \\ \theta_{t,d}(f), & \text{else if } f \in \mathcal{F} \setminus \{f(d), f_{\text{old}}\}. \end{cases} \quad (1)$$

In addition, a caching decision that $f(d)$ will not be stored may be made, and then the cache status will remain unchanged, i.e., $\theta_{t,d+1}(f) = \theta_{t,d}(f), \forall f \in \mathcal{F}$.

For convenience, we use $\boldsymbol{\theta}_{t,d} = [\theta_{t,d}(1), \theta_{t,d}(2), \dots, \theta_{t,d}(f), \dots, \theta_{t,d}(F)]^T$ to indicate the cache status of all the contents at the requesting time $t(d)$. Generally, an edge caching policy can be represented by a function $\Phi : (\boldsymbol{\theta}_{t,d}, \mathbf{x}(d), \mathcal{U}_t) \mapsto \boldsymbol{\theta}_{t,d+1}$, which maps the current cache status vector, the current feature vector, and the set of regional users to the next cache status vector. After a request $req_{t,d}$ is served, the cache status vector should be updated according to the edge caching policy. We use the overall cache hit rate to evaluate the caching performance, which is defined as the number of cache hits over the whole requests during the finite time horizon T as follows

$$\begin{aligned} \mathcal{H}(\Phi) &= \frac{\sum_{t=1}^T \sum_{d=1}^{D_t} \theta_{t,d}(f(d))}{\sum_{t=1}^T D_t} \\ &= \frac{1}{D} \sum_{t=1}^T \sum_{d=1}^{D_t} \theta_{t,d}(f(d)). \end{aligned} \quad (2)$$

Then, the corresponding edge caching optimization problem⁵ can be expressed as follows [22]

$$\begin{aligned} &\max_{\Phi} \mathcal{H}(\Phi), \\ \text{s.t. } &\theta_{t,d}(f) \in \{0, 1\}, \text{ for } 1 \leq d \leq D_t, 1 \leq t \leq T, \forall f \in \mathcal{F}, \\ &\boldsymbol{\theta}_{t,d}^T \boldsymbol{\theta}_{t,d} \leq M\varphi, \text{ for } 1 \leq d \leq D_t, 1 \leq t \leq T. \end{aligned} \quad (3)$$

Our objective in this paper is to find the optimal edge caching policy by maximizing the overall cache hit rate over the finite time horizon T with the limited total cache size $M\varphi$. For convenience, a summary of major notations is presented in Table I.

⁵Note here that the focus of this paper is content popularity prediction and user preference learning, which is one of the most important parts of edge caching or F-RAN. To highlight this issue, we simplify the content placement and adopt the partition-based content placement method and use the cache hit rate as the optimization objective, which help us concentrate on the investigation of content popularity prediction and user preference learning. If more sophisticated content placement methods and other optimization objectives are considered, the corresponding optimization problem will have more F-RAN-specific parameters or communication-related parameters. Actually, the investigation results concerning content popularity prediction in this paper can be readily extended to deal with non-simplified scenarios such as [19]–[21] by using the predicted content popularity rather than the assumed Zipf distribution.

III. THE PROPOSED USER PREFERENCE LEARNING BASED EDGE CACHING POLICY

In order to maximize the cache hit rate, we propose a novel edge caching policy which includes an online content popularity prediction algorithm and an offline user preference learning algorithm. The proposed policy can continuously cache popular contents based on the content features and user preferences.⁶

A. Policy Description

The detailed edge caching policy is shown in Algorithm 1. The considered M F-APs serving in the region set a fixed monitoring period and periodically monitor the current user set in the region. During the t th time period, the M F-APs first obtain the current user set \mathcal{U}_t . For each arriving request $req_{t,d}$ from the user $u \in \mathcal{U}_t$, its request information will then be recorded. Meanwhile, the features of $f(d)$ are extracted and recorded. The recorded data will be used to train or update the user preference model in order to improve the prediction precision of the content popularity. Let $\mathcal{G}_{t,d} = \{f | \theta_{t,d}(f) = 1, \forall f \in \mathcal{F}\}$ denote the set of current contents in the local cache. It will be explored to see whether $f(d)$ has already been stored locally. Just as stated in the previous section, the corresponding caching decision will be made to determine whether $f(d)$ should be cached and which stored content should be removed from the storage space of the M F-APs when $f(d)$ needs to be cached.

In order to make an optimal caching decision, the feature vector $\mathbf{x}(d)$ and the vectors of the well-trained user preference model parameters of all the users $\{\mathbf{w}_u | \forall u \in \mathcal{U}_t\}$ are extracted to predict the popularity $P_{t,d}$ of the requested content $f(d)$. In addition, considering that the content popularity will change over time, in order to track popularity changes, we let $P_{t,f}^{\text{cur}}$ denote the current popularity of the caching content $f \in \mathcal{G}_{t,d}$ after it is requested (also called the residual request rate). We know that users may have a certain request delay on the same content. In order to ensure timely and reasonable cache update, we propose to select the content with the characteristics of the smallest content popularity P^{smallest} and relatively earlier initial cache time $t_{f^{\text{smallest}}}$,⁷ denoted by f^{smallest} , as the content to be removed from the local caches. In order to locate f^{smallest} quickly, we propose to reserve a priority queue $Q_{t,d}$ that stores the caching contents along with their current content popularity $P_{t,f}^{\text{cur}}$ and their initial caching time t_f for $f \in \mathcal{G}_{t,d}$. The elements of $Q_{t,d}$ are sorted in sequence when the request $req_{t,d}$ arrives, whose top element is composed of f^{smallest} , $t_{f^{\text{smallest}}}$ and P^{smallest} . A caching decision is made by comparing the predictive popularity $\hat{P}_{t,d}$ and P^{smallest} . If $\hat{P}_{t,d}$ is larger

⁶Note that the adaptive caching scheme recently proposed in [23] also considers content popularity prediction based on content features and users' behavior. However, it uses an extreme-learning machine (ELM) neural network to estimate the content popularity, and more addresses the optimization of the number of neurons, the construction and selection of content features. Besides, its definition of content popularity is also different with ours.

⁷Note that here we propose to select the content with the smallest content popularity in the local caches. Furthermore, if there exist multiple contents that have the same smallest content popularity in the local caches, we propose to select the content with the smallest content popularity that has the longest cache time, i.e., the content that has been cached earliest.

TABLE I. Summary of major notations.

M, \mathcal{M}	Number of regional F-APs, set of the M regional F-APs
t, T	Discrete time periods, finite time horizon
φ	Cache size of each F-AP
$U_t, U_{\max}, U_{\min}, \mathcal{U}_t$	Number of regional users during the t th time period, maximum U_t , minimum U_t , Set of the U_t regional users
D_t, D	Number of requests during the t th time period, number of overall requests in the finite time horizon T
$req_t, req_{t,d}$	Set of requests coming in sequence during the t th time period, the d th request during the t th time period
$f, f(d), f_{\text{old}}, f^{\text{smallest}}$	Content, the d th requested content, content to be removed, content with the smallest popularity in the current local cache
$t(d), t_f, t_{f^{\text{smallest}}}$	The d th requesting time, initial caching time of the content f , initial caching time of the content f^{smallest}
$\mathcal{G}_{t,d}, P_{t,f}^{\text{cur}}, P_{t,f}^{\text{smallest}}$	Set of current contents in the local cache at the requesting time $t(d)$, current popularity of the caching content $f \in \mathcal{G}_{t,d}$ after it is requested, the smallest content popularity
$Q_{t,d}$	Priority queue that stores the caching contents and information sequentially
\mathcal{F}, F	Content library, size of content library
$\mathbf{x}(d), \mathbf{x}^{(k)}, N$	Feature vector of the d th requested content, feature vector of the k th sample, dimension of feature vector
$\theta_{t,d}, \boldsymbol{\theta}_{t,d}$	Cache status of the requested content $f(d)$ at the requesting time $t(d)$, vector of cache status of all the contents at the requesting time $t(d)$
Φ, Φ^*	Edge caching policy, the optimal edge caching policy
$\mathcal{H}_t(\Phi), \mathcal{H}_t(\Phi^*)$	Cache hit rate of Φ during the t th time period, the optimal cache hit rate during the t th time period
$\mathcal{H}(\Phi), \mathcal{H}(\Phi^*)$	Overall cache hit rate of Φ in the finite time horizon T , the optimal overall cache hit rate in the finite time horizon T
$\mathbf{w}_u, \mathbf{w}_u^{(k)}, \mathbf{w}_u^*$	Vector of user preference model parameters of the user u , vector of user preference model parameters of the user u for the k th iteration, vector of the optimal user preference model parameters
$\hat{p}_{t,u,d}, p_{t,u,d}$	Predicted possibility that the user u requests $f(d)$ at $t(d)$ during the t th time period, real possibility that the user u requests $f(d)$ at $t(d)$ during the t th time period
$\hat{P}_{t,d}, P_{t,d}, \hat{P}'_{t,d'}, P'_{t,d'}$	Predicted popularity of $f(d)$ at $t(d)$, real popularity of $f(d)$ at $t(d)$, predicted popularity with respect to $P'_{t,d'}$, popularity of the d' th most popular content when it is requested firstly during the t th time period
$y(d), y^{(k)}$	Category label of the d th requested content, category label of the k th sample
$K_{t,u,d}, K$	Cumulative number of samples for the user u from the last model update to the time when the request $req_{t,d}$ arrives, number of collected samples for offline user preference learning
$\ell(\mathbf{w}_u, \mathbf{x}(d), y(d)), \xi_{t,u,d}$	Logistic loss for the user u at $t(d)$, average logistic loss for the user u at $t(d)$
$\gamma, \eta^{(k)}, \sigma^{(k')}$	Predefined threshold, non-increasing learning-rate schedule, parameter that has certain relationship with $\eta^{(k)}$
$\mathbf{g}^{(k)}, \mathbf{g}^{(1:k)}$	Gradient vector of the logistic loss of the k th sample with respect to \mathbf{w}_u , sum of the gradient vectors of the logistic loss of the previous k samples
$\lambda_1, \lambda_2; \alpha, \beta; \tau_u$	Regularization parameters with positive values; adjusting parameters; sufficiently small constant with a positive value
$\mathcal{E}, L_d(\mathbf{w}_u)$	Convex set of the optimization problem in (12), a sequence of convex loss functions
F_t	number of requested different contents during the t th time period
$R(D)$	Regret of the overall cache hit rate for the total D requests in the finite time horizon T

than P^{smallest} , the existing content f^{smallest} will be replaced by $f(d)$, the initial caching time of $f(d)$ will be recorded, and the current popularity of $f(d)$ and the priority queue $Q_{t,d}$ will be updated accordingly. After that, a cache update process is completed. Otherwise, nothing will be done to the local caches. The key here, obviously, is to obtain $\hat{P}_{t,d}$, which will be described in the next subsection.

B. Online Content Popularity Prediction

In this subsection, we propose an online content popularity prediction algorithm based on the content features and user preferences. During the t th time period, for each requesting user $u \in \mathcal{U}_t$, the requested content can be classified into a preferred category or an unpreferred one for this user based on its user preference. Generally, a user prefers to request contents of its preferred category. The problem of whether a user will request a certain content can be converted into a simple two-category one. We use the sigmoid function to approximate the correspondence between the feature vector and the category label of the requested content [24], and construct a logistic

regression model to approximate the user preference model. For the arriving request $req_{t,d}$, it is characterized by the feature vector $\mathbf{x}(d)$. Let $y(d)$ denote the corresponding category label with $y(d) = 1$ if the requested content is the preferred category of the user and $y(d) = 0$ otherwise. Let $p_{t,u,d}$ denote the possibility that the user $u \in \mathcal{U}_t$ requests the content $f(d)$ at the requesting time $t(d)$ during the t th time period. Specifically, we assume that a user will not have a second request to the same content⁸. If the user u has already requested $f(d)$ previously, then $\hat{p}_{t,u,d} = 0$. Otherwise, $p_{t,u,d}$ can be predicted based on the following user preference model

$$\hat{p}_{t,u,d} = p_{\mathbf{w}_u}[y(d) = 1 | \mathbf{x}(d)] = \frac{1}{1 + e^{-(\mathbf{w}_u^T \mathbf{x}(d))}}. \quad (4)$$

⁸Note that the above assumption is for all the contents and has a fair impact on the content popularity. By using this assumption, neither the popular contents can become unpopular ones nor the unpopular contents can become the popular ones. Furthermore, in some ways, the content which a user has requested repeatedly can be obtained directly from this user's own cache and does not need to be obtained repeatedly from the corresponding F-APs.

Algorithm 1 The proposed edge caching policy

```

1: procedure EDGE_CACHING( $req_{t,d}$ )
2:   for  $t = 1, 2, \dots, T$ , do
3:     The considered  $M$  F-APs monitor  $\mathcal{U}_t$ ;
4:     for  $d = 1, 2, \dots, D_t$ , do
5:       Record the request information of  $req_{t,d}$ ; Read the
       set of the current caching content  $\mathcal{G}_{t,d}$ ;
6:       if  $f(d)$  has been stored locally, then
7:         The users are served locally;  $P_{t,f(d)}^{\text{cur}} = P_{t,f(d)}^{\text{cur}} -$ 
          $1/U_t$ ;
8:       else
9:         Fetch  $f(d)$  from the cloud content center or the
         neighboring F-APs in the other regions;
10:      Extract  $\mathbf{x}(d)$  and  $\{\mathbf{w}_u | \forall u \in \mathcal{U}_t\}$ ;  $\hat{P}_{t,d} \leftarrow$ 
       Predict( $\mathbf{x}(d), \{\mathbf{w}_u | \forall u \in \mathcal{U}_t\}$ );
11:      Sort  $Q_{t,d}$  based on  $P_{t,f}^{\text{cur}}$  and  $t_f$  for  $f \in \mathcal{G}_{t,d}$ ; Get
        $P^{\text{smallest}}, f^{\text{smallest}}$  from the top of  $Q_{t,d}$ ;
12:      if  $\hat{P}_{t,d} > P^{\text{smallest}}$ , then ▷ Cache update
13:        Remove the top element from  $Q_{t,d}$ ;  $t_{f(d)} =$ 
         $t(d)$ ;  $P_{t,f(d)}^{\text{cur}} = \hat{P}_{t,d} - 1/U_t$ ;
14:        Insert  $\langle P_{t,f(d)}^{\text{cur}}, f(d), t_{f(d)} \rangle$  into  $Q_{t,d}$ ; Re-
        place  $f^{\text{smallest}}$  by  $f(d)$ .
15:      end if
16:    end for
17:  end for
18: end procedure

```

Furthermore, the regional content popularity $P_{t,d}$ can be predicted by using the average request possibility from the regional users for $f(d)$ as follows

$$\hat{P}_{t,d} = \frac{1}{U_t} \sum_{u=1}^{U_t} \hat{p}_{t,u,d}. \quad (5)$$

After that, if the requested content is determined to be stored into the local cache or it has been stored locally and is requested again, according to our previous assumption that a user will not have a second request to the same content, its current popularity at the requesting time can be calculated respectively as follows

$$P_{t,f(d)}^{\text{cur}} = \hat{P}_{t,d} - 1/U_t, \quad P_{t,f(d)}^{\text{cur}} = P_{t,f(d)}^{\text{cur}} - 1/U_t. \quad (6)$$

Besides, if the requested content is determined not to be stored locally and is requested by another regional user again, its current popularity at the next requesting time can be predicted by (5) since we have set the possibility that the previous requesting user will request this content again to zero. In this way, our proposed online content popularity algorithm can track the popularity changes in real time.

To measure the prediction performance, we introduce the logistic loss $\ell(\mathbf{w}_u, \mathbf{x}(d), y(d))$ for the user u , which is defined as the negative log-likelihood of $y(d)$ given $p_{\mathbf{w}_u}(y(d) | \mathbf{x}(d))$ and can be expressed as follows

$$\begin{aligned} \ell(\mathbf{w}_u, \mathbf{x}(d), y(d)) &= -y(d) \log p_{\mathbf{w}_u}(y(d) | \mathbf{x}(d)) \\ &\quad - (1 - y(d)) \log[1 - p_{\mathbf{w}_u}(y(d) | \mathbf{x}(d))]. \end{aligned} \quad (7)$$

Since user preference may change over time, we need to capture the moment when the user preference changes and update the user preference model in real time. For this purpose,

Algorithm 2 The online content popularity prediction algorithm

```

1: procedure PREDICT( $\mathbf{x}(d), \{\mathbf{w}_u | \forall u \in \mathcal{U}_t\}$ )
2:   for  $u \in \mathcal{U}_t$ , do
3:     if The user  $u$  has requested the content, then
4:        $\hat{p}_{t,u,d} = 0$ ;
5:     else
6:       Obtain the well-trained  $\mathbf{w}_u$  from the  $M$  F-APs;
        $\hat{p}_{t,u,d} = 1/(1 + e^{-(\mathbf{w}_u^T * \mathbf{x}(d))})$ ; Observe the category label  $y(d)$ 
       of  $f(d)$ .
7:     end if
8:     Get  $(\mathbf{x}^{(K_{t,u,d})}, y^{(K_{t,u,d})})$ ;  $K_{t,u,d} = K_{t,u,d-1} + 1$ ;
      $\xi_{t,u,d} = \xi_{t,u,d-1} + \frac{1}{K_{t,u,d}} \ell(\mathbf{w}_u, \mathbf{x}^{(K_{t,u,d})}, y^{(K_{t,u,d})})$ ;
9:     if  $\xi_{t,u,d} \geq \gamma$ , then
10:       $\mathbf{w}_u \leftarrow \text{Learn}(\{(\mathbf{x}^{(k)}, y^{(k)})\}_{k=1}^{K_{t,u,d}})$ .
11:    end if
12:  end for
13:  return  $\frac{1}{U_t} \sum_{u=1}^{U_t} \hat{p}_{t,u,d}$ .
14: end procedure

```

we collect the samples $\{(\mathbf{x}^{(k)}, y^{(k)})\}_{k=1}^{K_{t,u,d}}$ and monitor the prediction performance in real time, where $\mathbf{x}^{(k)}$ and $y^{(k)}$ denote the feature vector and category label of the k th sample, respectively, and $K_{t,u,d}$ the cumulative number of samples for the user u from the last model update to the time when the request $req_{t,d}$ arrives during the t th time period. Then, the average logistic loss for the user u can be expressed as follows

$$\xi_{t,u,d} = \frac{1}{K_{t,u,d}} \sum_{k=1}^{K_{t,u,d}} \ell(\mathbf{w}_u, \mathbf{x}^{(k)}, y^{(k)}). \quad (8)$$

Let γ denote a predefined threshold with $0 \leq \gamma \leq 1$. When $\xi_{t,u,d}$ exceeds γ , the update of the user preference model will be initiated.

The detailed description of our proposed online content popularity prediction algorithm is shown in Algorithm 2. Note that our proposed algorithm not only can predict content popularity in an online fashion, but also can determine when to update the user preference model proactively. We also remark here that the time complexity of our proposed algorithm when making prediction for one content is $\mathcal{O}(U_t)$.

C. Offline User Preference Learning

With the rapid development of social networks and various multimedia applications, each user will access a large amount of contents everyday. In order to provide users with more intelligent services and enhance the quality of experience, it is necessary and feasible to establish the independent user preference model for each user. When the update of the user preference model is initiated, we assume that there are K samples collected for the considered user and extracted from the recorded data, which are denoted by $\{(\mathbf{x}^{(k)}, y^{(k)})\}_{k=1}^K$. Based on the collected samples, we propose to learn the user preference model parameters iteratively by minimizing the logistic loss of each sample as follows

$$\mathbf{w}_u^{(k+1)} = \arg \min_{\mathbf{w}_u} \left(\ell(\mathbf{w}_u, \mathbf{x}^{(k)}, y^{(k)}) \right), \quad k = 1, 2, \dots, K, \quad (9)$$

where $\mathbf{w}_u^{(k+1)}$ denotes the vector of the learned user preference model parameters of the k th iteration for the user u . By using the online gradient descent (OGD) method [25], the solution of the above optimization problem can be obtained through the iteratively updated model parameters as follows

$$\mathbf{w}_u^{(k+1)} = \mathbf{w}_u^{(k)} - \eta^{(k)} \mathbf{g}^{(k)}, \quad k = 1, 2, \dots, K, \quad (10)$$

where $\mathbf{g}^{(k)} = \nabla_{\mathbf{w}_u} \ell(\mathbf{w}_u, \mathbf{x}^{(k)}, y^{(k)}) = [p_{\mathbf{w}_u}(y^{(k)} | \mathbf{x}^{(k)}) - y^{(k)}] \mathbf{x}^{(k)}$ denotes the gradient vector of the logistic loss of the k th sample with respect to \mathbf{w}_u , and $\eta^{(k)}$ a non-increasing learning-rate schedule with $\sigma^{(k')}$ satisfying $\sum_{k'=1}^k \sigma^{(k')} = 1/\eta^{(k)}$. Then, we have the following theorem.

Theorem 1: The solution of the optimization problem in (9) can also be obtained by solving the following equivalent optimization problem

$$\begin{aligned} \mathbf{w}_u^{(k+1)} = \arg \min_{\mathbf{w}_u} & \left(\left(\mathbf{g}^{(1:k)} \right)^T \mathbf{w}_u \right. \\ & \left. + \frac{1}{2} \sum_{k'=1}^k \sigma^{(k')} \left\| \mathbf{w}_u - \mathbf{w}_u^{(k')} \right\|_2^2 \right), \quad k = 1, 2, \dots, K, \end{aligned} \quad (11)$$

where $\mathbf{g}^{(1:k)} = \sum_{k'=1}^k \mathbf{g}^{(k')}$ denotes the sum of the gradient vector of the logistic loss of the previous k samples.

Proof: Please see Appendix A. \square

Due to the sparse and unbalanced data and high-dimensional feature vector, there may exist over-fitting and high computational complexity problems [26]. Inspired by the follow the (proximally) regularized leader (FTRL-Proximal) method [27], which is an online optimization method based on the OGD method, the $L1$ - and $L2$ -regularization terms are introduced simultaneously in the optimization problem in (11) in order to avoid the above mentioned potential problems whereas obtain the optimal model parameters. The introduction of the $L1$ -regularization term is beneficial for realizing feature selection and producing sparse model, while the introduction of the $L2$ -regularization term is conducive to the smooth solution of the corresponding optimization problem. Correspondingly, the model parameters can be updated according to the previous samples by solving the following optimization problem

$$\begin{aligned} \mathbf{w}_u^{(k+1)} = \arg \min_{\mathbf{w}_u} & \left\{ \left(\mathbf{g}^{(1:k)} - \sum_{k'=1}^k \sigma^{(k')} \mathbf{w}_u^{(k')} \right)^T \mathbf{w}_u \right. \\ & + \frac{1}{2} (\lambda_2 + \sum_{k'=1}^k \sigma^{(k')}) \left\| \mathbf{w}_u \right\|_2^2 + \lambda_1 \left\| \mathbf{w}_u \right\|_1 \\ & \left. + \frac{1}{2} \sum_{k'=1}^k \sigma^{(k')} \left\| \mathbf{w}_u^{(k')} \right\|_2^2 \right\}, \quad k = 1, 2, \dots, K, \end{aligned} \quad (12)$$

where λ_1 and λ_2 denote the regularization parameters with positive values, $\|\cdot\|_1$ and $\|\cdot\|_2^2$ the $L1$ -norm and $L2$ -norm, respectively.

It can be readily seen from (12) that the last item in the right hand side of (12), i.e., $1/2 \sum_{k'=1}^k \sigma^{(k')} \left\| \mathbf{w}_u^{(k')} \right\|_2^2$, is irrespective with \mathbf{w}_u . Let $\mathbf{z}^{(k)} = \mathbf{g}^{(1:k)} - \sum_{k'=1}^k \sigma^{(k')} \mathbf{w}_u^{(k')}$. Then, an iterative relationship between $\mathbf{z}^{(k)}$ and $\mathbf{z}^{(k-1)}$ can

be established as follows

$$\mathbf{z}^{(k)} = \mathbf{z}^{(k-1)} + \mathbf{g}^{(k)} - \left(\frac{1}{\eta^{(k)}} - \frac{1}{\eta^{(k-1)}} \right) \mathbf{w}_u^{(k)}, \quad (13)$$

which implies that we only need to store $\mathbf{z}^{(k-1)}$ after using the last sample for learning. Correspondingly, the optimization problem in (12) can be further expressed as follows

$$\begin{aligned} \mathbf{w}_u^{(k+1)} = \arg \min_{\mathbf{w}_u} & \left\{ (\mathbf{z}^{(k)})^T \mathbf{w}_u + \lambda_1 \left\| \mathbf{w}_u \right\|_1 \right. \\ & \left. + \frac{1}{2} (\lambda_2 + \sum_{k'=1}^k \sigma^{(k')}) \left\| \mathbf{w}_u \right\|_2^2 \right\}, \quad k = 1, 2, \dots, K. \end{aligned} \quad (14)$$

We know that there exists a difference for the change rate of the weight of each feature dimension for the requested content, and the gradient value with respect to each feature dimension can reflect this change rate. Therefore, different learning rates are preferred for different feature dimensions. Define

$$\mathbf{g}^{(k)} = [g_1^{(k)}, g_2^{(k)}, \dots, g_n^{(k)}, \dots, g_N^{(k)}]^T, \quad (15)$$

$$\mathbf{z}^{(k)} = [z_1^{(k)}, z_2^{(k)}, \dots, z_n^{(k)}, \dots, z_N^{(k)}]^T, \quad (16)$$

$$\mathbf{w}_u = [w_{u,1}, w_{u,2}, \dots, w_{u,n}, \dots, w_{u,N}]^T, \quad (17)$$

$$\mathbf{w}_u^{(k+1)} = [w_{u,1}^{(k+1)}, w_{u,2}^{(k+1)}, \dots, w_{u,n}^{(k+1)}, \dots, w_{u,N}^{(k+1)}]^T. \quad (18)$$

Let $\sum_{k'=1}^k \sigma_n^{(k')} = \frac{1}{\eta_n^{(k)}}$, where $\eta_n^{(k)} = \alpha / [\beta + \sqrt{\sum_{k'=1}^k (g_n^{(k')})^2}]$ denotes the learning-rate schedule of the n th feature dimension with α and β being the adjusting parameters which are chosen to yield good learning performance. Then, the optimization problem in (14) can be decoupled into the following N independent scalar minimization problems

$$\begin{aligned} w_{u,n}^{(k+1)} = \arg \min_{w_{u,n}} & \left\{ z_n^{(k)} w_{u,n} + \lambda_1 |w_{u,n}| \right. \\ & \left. + \frac{1}{2} (\lambda_2 + \sum_{k'=1}^k \sigma_n^{(k')}) w_{u,n}^2 \right\}, \quad n = 1, 2, \dots, N. \end{aligned} \quad (19)$$

It can be easily verified that the optimization problem in (19) is an unconstrained non-smooth one, where the second item $\lambda_1 |w_{u,n}|$ is non-differentiable at $w_{u,n} = 0$. Let $\eta = \partial |w_{u,n}| \Big|_{w_{u,n} = w_{u,n}^{(k+1)}}$ denote the partial differential of $|w_{u,n}|$ at $w_{u,n}^{(k+1)}$. Then, we have

$$\begin{cases} -1 < \eta < 1, & \text{if } w_{u,n}^{(k+1)} = 0, \\ \eta = -1, & \text{else if } w_{u,n}^{(k+1)} < 0, \\ \eta = 1, & \text{otherwise.} \end{cases} \quad (20)$$

Correspondingly, the optimal solution $w_{u,n}^{(k+1)}$ should satisfy the following relationship

$$\begin{aligned} z_n^{(k)} + \lambda_1 \eta + (\lambda_2 + \sum_{k'=1}^k \sigma_n^{(k')}) w_{u,n}^{(k+1)} = 0, \\ n = 1, 2, \dots, N. \end{aligned} \quad (21)$$

We have known previously that $\lambda_1 > 0$. Correspondingly, by classifying $z_n^{(k)}$ into three cases, i.e., $|z_n^{(k)}| < \lambda_1$, $z_n^{(k)} > \lambda_1$

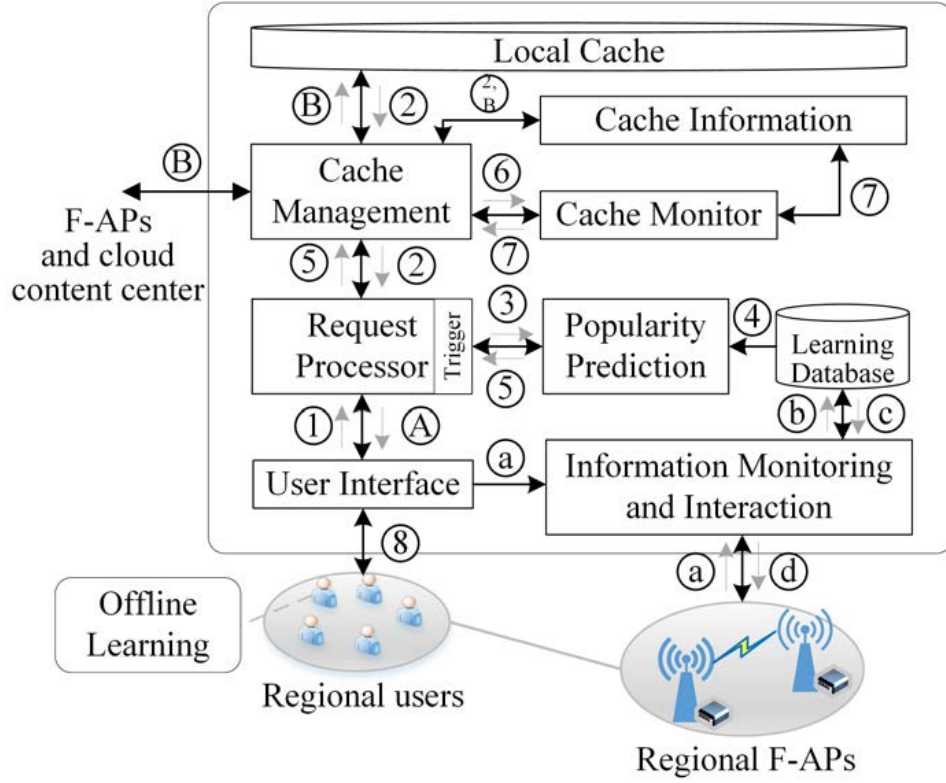


Fig. 4. Illustration of the learning based edge caching flowchart.

content based on the Learning Database. Note that the Offline Learning module will be initiated if the average prediction error is larger than the predefined threshold.

- The Offline Learning module is mainly responsible for learning the current user preference model parameters based on the collected information from the Feature Database and Request Database.
- The Data Updater module is mainly responsible for updating the content feature data, the requested content, and the requested time to the Feature Database and Request Database, and realizing the collection and update of the requested information.
- The Cache Information module is mainly responsible for storing and updating the current content popularity information, the initial cache time, and the cache content ID.
- The Cache Monitor module is mainly responsible for monitoring the cache information to capture the contents which need to be removed from the local cache.

The flowchart of our proposed first learning base edge caching architecture consists of five phases as illustrated in Fig. 2 and Fig. 4, and is presented below.

- (i) Initializing and periodic information monitoring
 - (a) The Information Monitoring and Interaction module periodically extracts the current user information of the serving F-AP and regional ones from their User Interface modules.
 - (b) This module regularly updates the collected regional user information to the Learning Database.
 - (c) This module extracts the current user information

of the serving F-AP from the Learning Database. (d) This module delivers the current user information of the serving F-AP to the regional ones.

- (ii) Direct local request responding
 - (1) The User Interface module delivers the user requesting information to the Request Processor module.
 - (2) The Request Processor module initiates a data updating procedure.
 - (3) The Data Updater module carries on numerical processing to the requested content feature and writes the processed feature data into the Feature Database, and updates the requested content information in the Request Database.
 - (4) If the requested content is stored locally, the Cache Management module delivers the stored content from the local cache to the Request Processor module.
 - (A) The Request Processor module serves the user request.
- (iii) Dynamic content caching and updating
 - (5) If the requested content is not stored locally, the Request Processor module triggers the Popularity Prediction module to make online content popularity prediction of the requested content.
 - (6) The Popularity Prediction module extracts the user preference model parameters and the requested content features from the Learning Database, and then predicts the popularity of the requested content.
 - (7) The Popularity Prediction module feeds back the predicted popularity of the requested content first to the Request Processor module, and then to the Cache Management module through the Request Processor module.
 - (8) The Cache Management module

triggers the Cache Monitor module to initiate monitoring of the cache content information. (9) The Cache Monitor module extracts the information of the content to be removed from the Cache Information module, and then feeds back the information to the Cache Management module. (10) The Cache Management module makes cache decision based on the feedback information, and notifies the Request Processor module to serve the user request. (11) The User Interface module broadcasts the local cached content information to the regional users.

- (iv) Cooperative caching and information synchronizing
 - (B) The Cache Management module executes the received cache decision of the current F-AP, and notifies the other regional F-APs to execute the same cache decision. If the requested content is to be cached, the Cache Management module retrieves the content from the neighboring F-APs in the other regions or the cloud content center, and then stores it locally by means of partition-based caching. The Cache Management module updates the cache information of the current F-AP, and synchronizes the cache information of the other regional F-APs.
- (v) Offline self-starting user preference learning
 - (C) The Popularity Prediction module initiates the Offline Learning module if the average prediction error cumulated under a user preference model is larger than the predefined threshold. (D) The Offline Model Update module retrieves the historical requested data of the considered user from the Feature Database and Request Database, generates a new training sample set, and then updates the user preference model parameters. (E) The Offline Model Update module refreshes the updated user preference model parameters to the Learning Database. (F) The Request Database releases the historical requested data of the considered user.

B. Learning Based Edge Caching Architecture (II)

For the proposed second architecture as illustrated in Fig. 3, part of the functionality, i.e., offline user preference learning, is transferred from the F-APs to the powerful and smart UEs. Therefore, both the signaling overhead among regional F-APs and the computational burden undertaken by the F-APs can be greatly reduced.

The corresponding processing flow of our proposed edge caching policy is presented in brief below. (a-d) The current F-AP is mainly responsible for monitoring the regional users in coordination with the regional F-APs and storing the corresponding user preference model parameters from the UEs into the Learning Database. (1-2, A) The current F-AP serves the requested user if it caches the requested content locally. (3-8) If the requested content is not stored in the local cache, the current F-AP predicts the content popularity, makes the corresponding caching decision, and broadcasts the caching information to the regional users. (B) The current F-AP notifies the other regional F-APs to execute the same cache decision and then updates the corresponding cache information.

We remark here that our proposed first architecture is more suitable for wireless communications scenarios including

intelligent F-APs and general UEs while our proposed second architecture is more suitable for wireless communications scenarios including general F-APs and intelligent UEs. With the rapid development of AI and smart UEs, the second architecture will show more advantages and dominate in future wireless communications. We also remark here that the signaling overhead can be further reduced by setting a cluster head for the clustered regional F-APs and the corresponding edge caching architecture is omitted here due to space limitation.

V. PERFORMANCE ANALYSIS

In this section, the performance of our proposed edge caching policy will be analyzed. Firstly, we derive the upper bound of the popularity prediction error of our proposed online content popularity prediction algorithm. Secondly, we derive the lower bound of the cache hit rate of our proposed edge caching policy. Finally, we derive the regret bound of the overall cache hit rate of our proposed edge caching policy.

A. The Upper Bound of the Popularity Prediction Error

Let \mathcal{E} denote the convex set of the optimization problem in (12) with $\mathcal{E} \in \mathbb{R}^N$, and define

$$W_u = \max_{\mathbf{w}_u, \mathbf{w}_{u'} \in \mathcal{E}} \|\mathbf{w}_u - \mathbf{w}_{u'}\|. \quad (23)$$

Let $L_d(\mathbf{w}_u)$ denote a sequence of convex loss functions, and define

$$G_u = \max_{\mathbf{w}_u \in \mathcal{E}, 1 \leq d \leq D_t, t=1,2,\dots,T} \|\nabla L_d(\mathbf{w}_u)\|. \quad (24)$$

Then, according to Corollary 1 in [28], the following relationship can be readily established for the optimization problem in (12) in the finite time horizon T

$$\sum_{t=1}^T \sum_{d=1}^{D_t} L_d(\mathbf{w}_u) - \sum_{t=1}^T \sum_{d=1}^{D_t} L_d(\mathbf{w}_u^*) \leq W_u G_u \sqrt{2D}, \quad (25)$$

where \mathbf{w}_u^* denotes the vector of the optimal user preference model parameters. Let τ_u denote a sufficiently small constant with a positive value, which can meet $\sum_{t=1}^T \sum_{d=1}^{D_t} L_d(\mathbf{w}_u^*) \leq \tau_u$. Define

$$W_{\max} = \max_{u \in \mathcal{U}, t=1,2,\dots,T} W_u, \quad (26)$$

$$G_{\max} = \max_{u \in \mathcal{U}, t=1,2,\dots,T} G_u, \quad (27)$$

$$\tau_{\max} = \max_{u \in \mathcal{U}, t=1,2,\dots,T} \tau_u. \quad (28)$$

Then, we have the following theorem.

Theorem 2: The expected popularity prediction error for the overall D requests in the finite time horizon T , i.e., $\mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{P}_{t,d} - P_{t,d}|$, can be upper bounded by $\frac{U_{\max}}{U_{\min}} (W_{\max} G_{\max} \sqrt{2D} + \tau_{\max})$.

Proof: Please see Appendix B. \square

It is clear that an upper bound exists for the cumulative prediction error of the content requested probability of one single user and that of the regional content popularity within a limited time periods. Specifically, by using the user preference

model that is obtained through self-learning, the upper bound of the cumulative prediction error of the content requested probability of one single user has a sub-linear relationship with the total number of requests D in the finite time horizon T . Similar relationship can be found for the cumulative prediction error of the content popularity in the considered region, which means that $\frac{1}{D} \mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{P}_{t,d} - P_{t,d}| \rightarrow 0$ as $D \rightarrow \infty$.

The above analytical results imply that the learned user preference model will asymptotically approach the real user preference model through sufficient learning with the collection of more user requesting information and the increased requests. Correspondingly, the proposed online content popularity prediction algorithm can make the content popularity prediction more accurate. On the other hand, the analytical results also reveal that the performance of our proposed policy can be improved with the increased requests. After a certain number of content requests, the prediction precision of the proposed policy can achieve the ideal value.

B. The Lower Bound of the Cache Hit Rate

In this subsection, we first show the cache hit rate of the optimal edge caching policy which knows the real popularities of all the contents and caches the most popular contents during each time period, and then derive the lower bound of the cache hit rate of the proposed edge caching policy and reveal their relationship.

1) The cache hit rate of the optimal edge caching policy:

In the ideal case, the optimal cache hit rate can be achieved by caching the most popular contents of the current time period based on the known content popularity. Let Φ^* denote the optimal edge caching policy. In practice, we note that the cache hit rate depends not only on the edge caching policy Φ but also on the degree of concentration of the same content requests. Generally speaking, a more concentrated content request process implies a higher potential cache hit rate. For the ease of analysis, we assume that the requests of the same content are concentrated in one time period. Let $P'_{t,d'}$ denote the content popularity of the d' th most popular content when it is requested firstly during the t th time period, and F_t the number of requested different contents during the t th time period. Then, by using the relationship in (5), the optimal cache hit rate during the t th time period can be calculated as follows

$$\mathcal{H}_t(\Phi^*) = \frac{\sum_{d'=1}^{M\varphi} U_t P'_{t,d'}}{\sum_{d'=1}^{F_t} U_t P'_{t,d'}} = \frac{\sum_{d'=1}^{M\varphi} P'_{t,d'}}{\sum_{d'=1}^{F_t} P'_{t,d'}}. \quad (29)$$

Note that $\mathcal{H}_t(\Phi^*)$ has no relation with the number of requested users during the current time period, and may vary with different time periods. Besides, we make no assumption about the popularity distribution, and $\mathcal{H}_t(\Phi^*)$ is just the cache hit rate that is calculated based on the real content popularity during the current time period.

2) The lower bound of the cache hit rate of the proposed edge caching policy: In practice, there always exist popularity prediction errors. Let $\hat{P}'_{t,d'}$ denote the predicted popularity

with respect to $P'_{t,d'}$. Let

$$\Delta P_t = \max_{d'=1,2,\dots,F_t} |\hat{P}'_{t,d'} - P'_{t,d'}|. \quad (30)$$

Then, we have the following theorem.

Theorem 3: During the t th time period, the achievable cache hit rate $\mathcal{H}_t(\Phi)$ can be lower bounded by $\mathcal{H}_t(\Phi^*) - M\varphi(\Delta P_t U_t + 2)/D_t$.

Proof: Please see Appendix C. \square

It is clear that a lower bound exists for the cache hit rate of the proposed edge caching policy during each time period. The analytical result from Theorem 3 gives us the insight that there exists a certain performance gap, i.e., $M\varphi(\Delta P_t U_t + 2)/D_t$, between the cache hit rate of our proposed edge caching policy and that of the optimal edge caching policy. The first component of the performance gap, i.e., $M\varphi\Delta P_t U_t/D_t$, is mainly caused by the popularity prediction error ΔP_t and principally determined by the accuracy of the learned user preference model. The second component of the performance gap, i.e., $2M\varphi/D_t$, is mainly caused by the operational mechanism of our proposed edge caching policy, which decides to cache a content only after but not before its first request (i.e., an initial cache miss will happen).

Although our proposed edge caching policy may result in an initial cache miss, it can actually avoid the extremely high computational load that may bring about to F-APs due to the need of continuously predicting all the content popularities otherwise. We point out here that this type of performance gap will gradually approach zero with the increased number of content requests, and hence the benefit outweighs the cost. Moreover, the analytical result from Theorem 3 also reveals that the overall performance gap will approach zero when the prediction error of the content popularity approaches zero, i.e., $\Delta P_t \rightarrow 0$, and the number of content requests during the t th time period is much larger than the overall cache size of all the F-APs in the considered region, i.e., $D_t \gg M\varphi$. Correspondingly, the cache hit rate of our proposed edge caching policy during the t th time period will approach that of the optimal edge caching policy if the above two conditions are satisfied.

C. The Regret Bound of the Overall Cache Hit Rate

In order to measure the performance loss of the proposed edge caching policy in comparison with the optimal one, we will analyze and bound the regret of the overall cache hit rate.

Let $\mathcal{H}(\Phi^*)$ denote the overall cache hit rate of the optimal edge caching policy in the finite time horizon T . Then, from (2) and (29), it can be calculated as follows

$$\mathcal{H}(\Phi^*) = \frac{\sum_{t=1}^T \sum_{d'=1}^{M\varphi} U_t P'_{t,d'}}{\sum_{t=1}^T \sum_{d'=1}^{F_t} U_t P'_{t,d'}} = \frac{\sum_{t=1}^T \mathcal{H}_t(\Phi^*) D_t}{\sum_{t=1}^T D_t}. \quad (31)$$

Let $\mathcal{H}(\Phi)$ denote the overall cache hit rate of the proposed edge caching policy in the finite time horizon T . Then, we have

$$\mathcal{H}(\Phi) = \frac{\sum_{t=1}^T \mathcal{H}_t(\Phi) D_t}{\sum_{t=1}^T D_t}. \quad (32)$$

Then, the regret of the overall cache hit rate of the proposed edge caching policy for the total D requests in the finite time horizon T can be defined as follows

$$R(D) = \mathbb{E}[\mathcal{H}(\Phi^*) - \mathcal{H}(\Phi)]. \quad (33)$$

Utilizing the analytical results from Theorem 2 and Theorem 3, we have the following theorem.

Theorem 4: The regret of the overall cache hit rate for the total D requests in the finite time horizon T , i.e., $R(D)$, can be upper bounded by $U_{\max} M\varphi \left[\frac{U_{\max}}{U_{\min}} (W_{\max} G_{\max} \sqrt{2D} + \tau_{\max}) + 2T/U_{\max} \right] / D$.

Proof: Please see Appendix D. \square

According to the above theorem, with the limited U_{\max} , $M\varphi$ and T , the following relationship can be naturally obtained: $\lim_{D \rightarrow +\infty} R(D) = 0$, which shows that the performance loss of our proposed edge caching policy can be gradually reduced to zero as the number of requests is increased, i.e., our proposed edge caching policy has the capability to achieve the optimal performance asymptotically. The reason is that the learned user preference model gradually approaches the real one with the increased request samples, which makes the prediction errors even smaller.

VI. SIMULATION RESULTS

To evaluate the performance of the proposed edge caching policy, we take movie content⁹ as an example and our main datasets are extracted from the MovieLens 200M Dataset [29], [30]. From the MovieLens, we choose the requesting dataset of the selected 30 users¹⁰ who request the contents from January 01, 2010 to October 17, 2016. The first part of the requesting dataset, whose requesting dates are from January 01, 2010 to December 31, 2015, is used for initializing the user preference, while the second part of the requesting dataset, whose requesting dates are from January 01, 2016 to October 17, 2016, is used for evaluating the performance. To simulate the content request process, we take the movie rating from a user as the request for this movie just as it is assumed in [15] and [16]. In our simulations, we randomly select 25 users from the 30 users as the fixed regional users while the remaining 5 users as the mobile users that randomly enter the region. Besides, we set the number of considered F-APs M to 3, the finite time horizon T to 6984 hours, the preset monitoring cycle to 1 hour¹¹ and the predefined threshold γ to 0.2, respectively.

⁹Other types of contents are also possible. However, due to the limitations of data acquisition, movie content is just taken here as an example.

¹⁰It takes a great deal of simulation works to learn user preference models for a large number of users, and this is the reason why we only select 30 users in our simulations. In practice, regional F-APs indeed serve a much larger number of users. With the increase of the users served by the regional F-APs, the aggregation of the contents requested by the users will become much higher. Accordingly, popular contents will be more concentrated whereas non-popular contents will be more dispersed. Correspondingly, it will have a larger cache hit rate to cache the popular contents with more served users.

¹¹Note that we do not explore the impact of different monitoring cycle settings on our proposed caching policy. The reason is that this is a specific regional scenario which we simulate by randomly selecting users. The impact of the monitoring cycle on the caching policy here has no reference significance to the actual scenario.

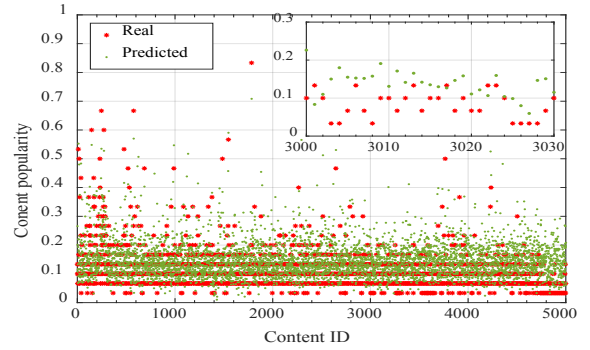


Fig. 5. Comparison of the predicted popularity and the real popularity in the finite time horizon T .

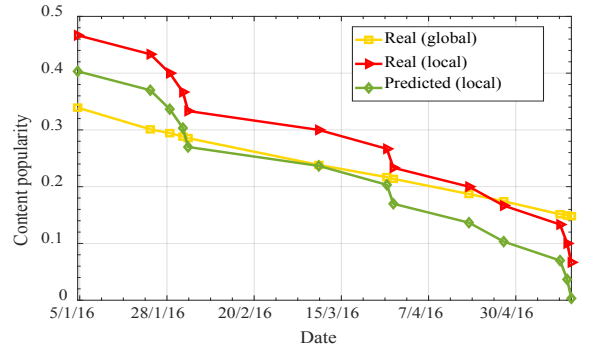


Fig. 6. Comparison of the predicted local popularity, the real local popularity, and the real global popularity for a certain requested content over time.

In Fig. 5, we show the predicted popularity by using our proposed edge caching policy in comparison with the real popularity at the moment when the contents are requested firstly by one of the regional users for the preceding 5000 contents in the finite time horizon T , where the content ID is marked according to the first requesting time of its representing content in chronological order. It can be observed that the error between the predicted popularity and the real popularity is very small. It can also be observed that the first requesting time of a regional popular content is random. Therefore, it is impractical for the existing edge caching policies to cache the most popular contents directly without consideration of the content requesting time and temporal popularity dynamic. It reveals the potential advantages of implementing a caching policy in conjunction with the content requesting time and real-time content popularity.

In Fig. 6, we show the predicted local popularity, the real local popularity, and the real global popularity over time for a randomly selected content that is requested by the regional users. It can be observed that both the global real popularity and the local real popularity decrease with time whereas the latter fluctuates slightly, which verifies that the global popularity cannot precisely reflect the temporal changes of the local popularity. It can also be observed that the predicted local popularity changes with the real local popularity in real time and the former approaches the latter, which reveals that our proposed policy can indeed track the real local popularity

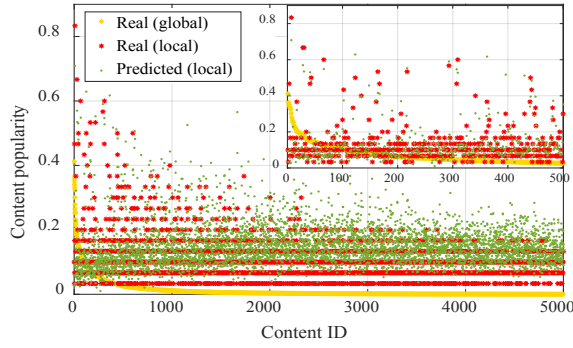


Fig. 7. Comparison of the predicted local popularity, the real local popularity for the considered region, and the real global popularity for all the regions.

changes without delay. This will guide the F-APs in a timely manner to clear the content that is no longer popular.

Without consideration of the requesting time difference of the users and the duration difference of the continuous requests for the same content, we analyze the spatial changes of content popularity. In Fig. 7, we show the predicted local popularity, the real local popularity for the considered region in comparison with the real global popularity for all the regions, where the content ID is marked according to the real global popularity of its representing content in descending order. It can be observed that most of the contents with the real local popularity larger than 0.2 have a content ID smaller than 2000, which reveals that most of the local popular contents originate from the global popular ones. This also reveals that when a user preference model cannot be well learned due to sparse data, the global popular contents can be selected as the user's initial requesting contents. We can observe that the contents with the content ID smaller than 500 generally have a larger real global popularity but a fluctuant real local popularity. We can also observe that the real global popularity approximately follows a typical Zipf distribution whereas the real local popularity does not. These observations reveal that the local popularity changes with the spatial popularity dynamic and does not necessarily follow the global popularity, and confirm the necessity of exploring the distribution of content popularity in a specific region.

In Fig. 8, we show the overall cache hit rate of our proposed policy with different $M\varphi$ in the finite time horizon T . Also included in Fig. 8 are the overall cache hit rates of the four baseline policies, i.e., the FIFO [6], LRU [7], LFU [8] policies, and the optimal policy with real content popularity. The total cache size $M\varphi$ increases from $1.5\%F = 60$ to $11.97\%F = 4800$ contents with $F = 40110$. It can be observed that the overall cache hit rates of all the considered policies are gradually increased with $M\varphi$. It can also be observed that the overall cache hit rate of our proposed policy gradually approaches the optimal performance and is apparently superior to those of the other three baseline policies for all the considered cache sizes. The reason is that the latter can not predict future content popularity. Instead, our proposed policy not only can predict the content popularity online, but also can track its changes in real time. Specifically, it can be

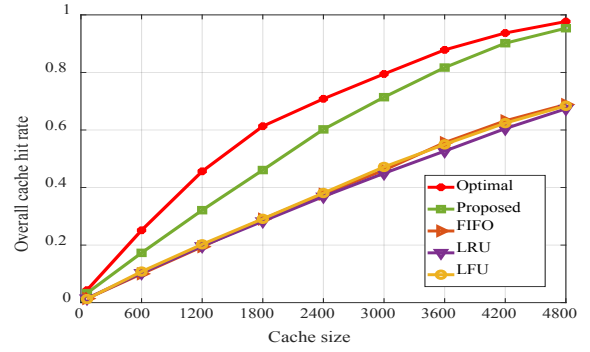


Fig. 8. The overall cache hit rate versus the total cache size ($M\varphi$) for the proposed policy and the baseline policies.

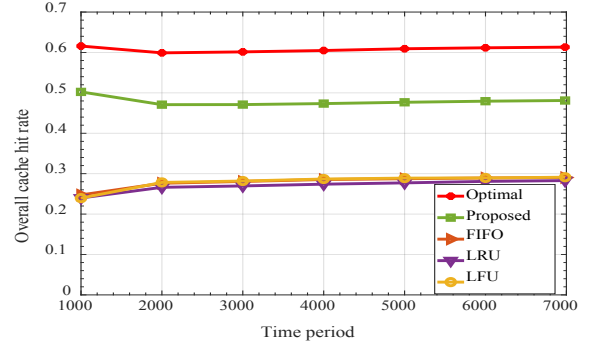


Fig. 9. The overall cache hit rate versus time period for the proposed policy and the baseline policies.

observed that our proposed policy only needs a cache size of approximately 2400 contents to achieve the cache hit rate of 0.6 whereas the other three baseline policies need a cache size of approximately 4200 contents.

In Fig. 9, we show the overall cache hit rates of our proposed policy and the four baseline policies until the current time period with $M\varphi = 1800$. It can be observed that the overall cache hit rate of our proposed policy follows consistently along with that of the optimal policy. The reasons are that both of them make caching decisions according to the content popularity, and that the distributions of the predicted popularity and the real one are consistent during every time period. It can also be observed that the changes of the overall cache hit rates of all the policies are different. The FIFO, LRU and LFU policies have low overall cache hit rates during the initial time periods due to the inevitable cold-start problem, whereas our proposed policy can cache the predicted popular contents according to the already-learned user preference model during the initial time period and then achieve a higher cache hit rate accordingly. After that, the overall cache hit rates of all the policies gradually increase with the time period. The reason is that caching decisions can be made more accurate with the increase of user requests.

VII. CONCLUSIONS

In this paper, we have proposed two edge caching architectures and a novel edge caching policy by learning user preference and predicting content popularity. Our proposed

policy can promptly detect the regional popular contents through online content popularity prediction, and store it to the local caches in real time. Specifically, we have proposed a self-starting offline user preference model updating mechanism by monitoring the average logistic loss in real time, which avoids frequent and blind training. Analytical results have shown that our proposed policy has the capability of asymptotically approaching the optimal performance. Simulation results have shown that our proposed policy achieves a better caching performance (i.e., overall cache hit rate) compared to the other traditional policies. Future work will explore the idea of on-device caching using distributed and low-latency machine learning.

APPENDIX

A. PROOF OF THEOREM 1

It can be readily seen that the objective function of the optimization problem in (11) is convex. Therefore, the iteratively updated model parameters can be obtained by setting the first order partial derivative of the corresponding objective function with respect to \mathbf{w}_u to zero as follows

$$\begin{aligned} & \partial \left(\left(\mathbf{g}^{(1:k)} \right)^T \mathbf{w}_u + \frac{1}{2} \sum_{k'=1}^k \sigma^{(k')} \left\| \mathbf{w}_u - \mathbf{w}_u^{(k')} \right\|_2^2 \right) / \partial \mathbf{w}_u \\ & = \mathbf{g}^{(1:k)} + \sum_{k'=1}^k \sigma^{(k')} \left(\mathbf{w}_u - \mathbf{w}_u^{(k')} \right) = 0. \end{aligned} \quad (34)$$

Correspondingly, the solution of the optimization problem in (11), i.e., $\mathbf{w}_u^{(k+1)}$, should satisfy the following relationship

$$\sum_{k'=1}^k \sigma^{(k')} \mathbf{w}_u^{(k+1)} = \sum_{k'=1}^k \sigma^{(k')} \mathbf{w}_u^{(k')} - \mathbf{g}^{(1:k)}. \quad (35)$$

Utilizing the relationship $\sum_{k'=1}^k \sigma^{(k')} = 1/\eta^{(k)}$, we can establish that

$$\frac{1}{\eta^{(k)}} \mathbf{w}_u^{(k+1)} = \sum_{k'=1}^k \sigma^{(k')} \mathbf{w}_u^{(k')} - \mathbf{g}^{(1:k)}. \quad (36)$$

When $k = 1$, from (36), we can readily establish: $\mathbf{w}_u^{(2)} = \mathbf{w}_u^{(1)} - \eta^{(1)} \mathbf{g}^{(1)}$. When $k = 2, 3, \dots, K$, replace k in (36) by $k - 1$. Then, we have

$$\frac{1}{\eta^{(k-1)}} \mathbf{w}_u^{(k)} = \sum_{k'=1}^{k-1} \sigma^{(k')} \mathbf{w}_u^{(k')} - \mathbf{g}^{(1:k-1)}. \quad (37)$$

From (36) and (37), we can readily establish that

$$\frac{1}{\eta^{(k)}} \mathbf{w}_u^{(k+1)} - \frac{1}{\eta^{(k-1)}} \mathbf{w}_u^{(k)} = \sigma^{(k)} \mathbf{w}_u^{(k)} - \mathbf{g}^{(k)}. \quad (38)$$

Exploiting the relationship $\sigma^{(k)} = (1/\eta^{(k)} - 1/\eta^{(k-1)})$, we can further establish that

$$\frac{1}{\eta^{(k)}} \mathbf{w}_u^{(k+1)} - \frac{1}{\eta^{(k-1)}} \mathbf{w}_u^{(k)} = \left(\frac{1}{\eta^{(k)}} - \frac{1}{\eta^{(k-1)}} \right) \mathbf{w}_u^{(k)} - \mathbf{g}^{(k)}. \quad (39)$$

Then, we have

$$\mathbf{w}_u^{(k+1)} = \mathbf{w}_u^{(k)} - \eta^{(k)} \mathbf{g}^{(k)}, \quad k = 2, 3, \dots, K. \quad (40)$$

According to the above analysis, we can certainly establish

$$\mathbf{w}_u^{(k+1)} = \mathbf{w}_u^{(k)} - \eta^{(k)} \mathbf{g}^{(k)}, \quad k = 1, 2, \dots, K. \quad (41)$$

It is obvious that the above solution of the optimization problem in (11) is the same as the solution of the optimization problem in (9), i.e., (10). This completes the proof.

B. PROOF OF THEOREM 2

We first analyze the upper bound of the expected prediction error of the content requested possibility of one single user for the overall D requests in the finite time horizon T . Without loss of generality, the convex loss function is chosen to be an absolute loss function, i.e., $L_d(\mathbf{w}_u) = |\hat{p}_{t,u,d} - p_{t,u,d}|$. Then, from (25), the following relationship can be readily established

$$\mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{p}_{t,u,d} - p_{t,u,d}| \leq W_u G_u \sqrt{2D} + \tau_u. \quad (42)$$

Furthermore, by using the relationship in (5), the expected popularity prediction error for the overall D requests in the finite time horizon T can be formulated as follows

$$\mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{P}_{t,d} - P_{t,d}| = \mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} \sum_{u=1}^{U_t} \frac{1}{U_t} |\hat{p}_{t,u,d} - p_{t,u,d}|. \quad (43)$$

By considering that $U_{\min} \leq U_t \leq U_{\max}$, the following inequation can be readily established

$$\mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{P}_{t,d} - P_{t,d}| \leq \frac{1}{U_{\min}} \mathbb{E} \sum_{u=1}^{U_{\max}} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{p}_{t,u,d} - p_{t,u,d}|. \quad (44)$$

Then, from (42), we can obtain

$$\mathbb{E} \sum_{t=1}^T \sum_{d=1}^{D_t} |\hat{P}_{t,d} - P_{t,d}| \leq \frac{U_{\max}}{U_{\min}} \left(W_{\max} G_{\max} \sqrt{2D} + \tau_{\max} \right). \quad (45)$$

This completes the proof.

C. PROOF OF THEOREM 3

During each time period, the proposed edge caching policy always tries to cache the $M\varphi$ most popular contents. In the ideal case, the contents with the $M\varphi$ largest real popularities, i.e., $\{P'_{t,1}, P'_{t,2}, \dots, P'_{t,M\varphi}\}$, will be cached. Due to the popularity prediction errors, the contents with the $M\varphi$ largest predicted popularities will however be cached in our proposed edge caching policy. Assume that the predicted popularities are sorted in descending order as follows: $\hat{P}'_{t,d'_1} \geq \hat{P}'_{t,d'_2} \geq \dots \geq \hat{P}'_{t,d'_f} \geq \dots \geq \hat{P}'_{t,d'_{F_t}}$. Obviously, $\{\hat{P}'_{t,d'_1}, \hat{P}'_{t,d'_2}, \dots, \hat{P}'_{t,d'_{M\varphi}}\}$ represent the $M\varphi$ largest predicted popularities. Then, the following relationship can be readily established

$$\sum_{f=1}^{M\varphi} \hat{P}'_{t,d'_f} \geq \sum_{d=1}^{M\varphi} \hat{P}'_{t,d'}. \quad (46)$$

According to the definition of ΔP_t , we have $|\hat{P}'_{t,d'} - P'_{t,d'}| \leq \Delta P_t$. Correspondingly, we have

$$\hat{P}'_{t,d'} \geq P'_{t,d'} - \Delta P_t, \quad \sum_{d'=1}^{M\varphi} \hat{P}'_{t,d'} \geq \sum_{d'=1}^{M\varphi} (P'_{t,d'} - \Delta P_t). \quad (47)$$

From (29), the following relationship holds

$$\sum_{d'=1}^{M\varphi} P'_{t,d'} = \mathcal{H}_t(\Phi^*) \sum_{d'=1}^{F_t} P'_{t,d'}. \quad (48)$$

Correspondingly, from (46), (47) and (48), we can readily obtain

$$\sum_{f=1}^{M\varphi} \hat{P}'_{t,d'_f} \geq \mathcal{H}_t(\Phi^*) \sum_{d'=1}^{F_t} P'_{t,d'} - \Delta P_t M\varphi. \quad (49)$$

According to the previous descriptions in Section II, the achievable cache hit rate $\mathcal{H}_t(\Phi)$ during the t th time period can be defined as follows

$$\mathcal{H}_t(\Phi) = \frac{1}{D_t} \sum_{d=1}^{D_t} \theta_{t,d}(f(d)). \quad (50)$$

We have assumed previously that the requests of the same content are concentrated in one time period. Therefore, the contents with the corresponding popularities $\{\hat{P}'_{t,d'_1}, \hat{P}'_{t,d'_2}, \dots, \hat{P}'_{t,d'_{M\varphi}}\}$ will be cached after the first content request with an initial cache miss happens during the t th time period. Then, we have

$$\sum_{d=1}^{D_t} \theta_{t,d}(f(d)) = \sum_{f=1}^{M\varphi} \left[U_t \hat{P}'_{t,d'_f} - 1 \right], \quad (51)$$

where $\lfloor \cdot \rfloor$ denotes the floor operation. Then, by using the relationships $D_t = \sum_{d'=1}^{F_t} U_t P'_{t,d'}$ and $\lfloor x \rfloor \geq x - 1$, we can further establish the following inequation

$$\mathcal{H}_t(\Phi) \geq \frac{\sum_{f=1}^{M\varphi} \hat{P}'_{t,d'_f} - 2/U_t}{\sum_{d'=1}^{F_t} P'_{t,d'}}. \quad (52)$$

By utilizing (49), the following relationship can be readily established

$$\mathcal{H}_t(\Phi) \geq \mathcal{H}_t(\Phi^*) - \frac{M\varphi}{D_t} (\Delta P_t U_t + 2). \quad (53)$$

This completes the proof.

D. PROOF OF THEOREM 4

From (31), (32), and (33), we have

$$R(D) = \mathbb{E} \frac{1}{\sum_{t=1}^T D_t} \sum_{t=1}^T [\mathcal{H}_t(\Phi^*) - \mathcal{H}_t(\Phi)] D_t. \quad (54)$$

By utilizing the analytical results from Theorem 3, the following relationship can be readily established

$$R(D) \leq \mathbb{E} \frac{M\varphi}{D} \sum_{t=1}^T (\Delta P_t U_t + 2). \quad (55)$$

According to the definition of ΔP_t , we have $\Delta P_t \leq \sum_{d=1}^{D_t} |\hat{P}'_{t,d} - P'_{t,d}|$. Exploiting the above relationship and

considering that $U_t \leq U_{\max}$, we can further obtain

$$R(D) \leq \mathbb{E} \frac{U_{\max} M\varphi}{D} \sum_{t=1}^T \left(\sum_{d=1}^{D_t} |\hat{P}'_{t,d} - P'_{t,d}| + \frac{2}{U_{\max}} \right). \quad (56)$$

By using the analytical results from Theorem 2, the following relationship can be readily established

$$R(D) \leq \frac{U_{\max} M\varphi}{D} \left[\frac{U_{\max}}{U_{\min}} \left(W_{\max} G_{\max} \sqrt{2D} + \tau_{\max} \right) + \frac{2T}{U_{\max}} \right]. \quad (57)$$

This completes the proof.

REFERENCES

- [1] M. Peng, Y. Li, Z. Zhao, and et al., "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar. 2015.
- [2] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. the IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.
- [3] E. Zeydan, E. Bastug, M. Bennis, and et al., "Big data caching for networking: moving from cloud to edge," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 36–42, Sept. 2016.
- [4] M. Peng, S. Yan, K. Zhang, and et al., "Fog-computing-based radio access networks: Issues and challenges," *IEEE Netw.*, vol. 30, no. 4, pp. 46–53, Jul. 2016.
- [5] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.
- [6] C. Aggarwal, J. Wolf, and P. Yu, "Caching on the world wide web," *IEEE Trans. Knowl. Data Eng.*, vol. 11, no. 1, pp. 94–107, Jan. 1999.
- [7] H. Ahlehagh and S. Dey, "Video caching in radio access network: Impact on delay and capacity," in *Proc. 2012 IEEE Wireless Commun. Netw. Conf.*, Apr. 2012, pp. 2276–2281.
- [8] X. Wang, M. Chen, T. Taleb, and et al., "Cache in the air: Exploiting content caching and delivery techniques for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–139, Feb. 2014.
- [9] M. Shafiq, A. Liu, and A. Khakpour, "Revisiting caching in content delivery networks," in *Proc. 2014 ACM Int. Conf. Measurement Modeling Comput. Syst.*, Jun. 2014, pp. 567–568.
- [10] K. Hamidouche, W. Saad, and M. Debbah, "Many-to-many matching games for proactive social-caching in wireless small cell networks," in *Proc. 12th Int. Symp. Modeling Opt. Mobile, Ad Hoc, Wireless Netw.*, May 2014, pp. 569–574.
- [11] K. Poularakis, G. Iosifidis, and L. Tassioulas, "Approximation algorithms for mobile data caching in small cell networks," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3665–3677, Oct. 2014.
- [12] B. Zhou, Y. Cui, and M. Tao, "Analysis and optimization of caching and multicasting in large-scale cache-enabled wireless networks," *IEEE Trans. Commun.*, vol. 15, no. 7, pp. 5101–5112, Jul. 2016.
- [13] E. Bastug, M. Bennis, and M. Debbah, "A transfer learning approach for cache-enabled wireless networks," in *Proc. 13th Int. Symp. Modeling Opt. Mobile, Ad Hoc, Wireless Netw.*, Jul. 2015, pp. 161–166.
- [14] B. Bharath, K. Nagananda, and H. Poor, "A learning-based approach to caching in heterogeneous small cell networks," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1674–1686, Apr. 2016.
- [15] S. Müller, O. Atan, M. Schaar, and et al., "Context-aware proactive content caching with service differentiation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1024–1036, Feb. 2017.
- [16] S. Li, J. Xu, M. Schaar, and et al., "Trend-aware video caching through online learning," *IEEE Trans. Multimedia*, vol. 18, no. 12, pp. 2503–2516, Dec. 2016.
- [17] —, "Popularity-driven content caching," in *Proc. IEEE INFOCOM 2016*, Apr. 2016, pp. 1–9.
- [18] Z. Chen, J. Lee, T. Quek, and et al., "Cooperative caching and transmission design in cluster-centric small cell networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3401–3415, May 2017.
- [19] X. Cui, Y. Jiang, X. Chen, and et al., "Graph-based cooperative caching in Fog-RAN," in *Proc. IEEE ICNC 2018 Workshops*, Hawaii, USA, Mar. 2018.

- [20] Y. Hu, Y. Jiang, M. Bennis, and et al., "Distributed edge caching in ultra-dense fog radio access networks: A mean field approach," in *Proc. IEEE VTC 2018 Fall*, Chicago, USA, Aug. 2018.
 - [21] A. Peng, Y. Jiang, M. Bennis, and et al., "Performance analysis and caching design in fog radio access networks," in *IEEE GLOBECOM 2018 Wrkshps. ET5GB*, Abu Dhabi, UAE, Dec. 2018.
 - [22] Y. Jiang, M. Ma, M. Bennis, and et al., "A novel caching policy with content popularity prediction and user preference learning in Fog-RAN," in *Proc. IEEE GLOBECOM 2017 Wrkshps. ET5GB*, Singapore, Dec. 2017.
 - [23] S. M. S. Tanzil, W. Hoiles, and V. Krishnamurthy, "Adaptive scheme for caching youtube content in a cellular network: Machine learning approach," *IEEE Access*, vol. 5, pp. 5870–5881, May 2017.
 - [24] V. Raykar, S. Yu, L. Zhao, and et al., "Supervised learning from multiple experts: Whom to trust when everyone lies a bit," in *Proc. 26th Annl*80] *Int. Conf. Machine Learning*, Jun. 2009, pp. 889–896.
 - [25] P. Bartlett, E. Hazan, and A. Rakhlin, "Adaptive online gradient descent," in *Proc. Neural Inf. Process. Syst. Conf.*, Jan. 2007, pp. 65–72.
 - [26] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: Datamining, inference, and prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
 - [27] H. MacMahan, G. Holt, D. Sculley, and et al., "Ad click prediction: A view from the trenches," in *Proc. ACM SIGKDD 19th Int. Conf. Knowl. Discovery Data Mining*, Apr. 2013, pp. 1222–1230.
 - [28] B. McMahan., "A unified view of regularized dual averaging and mirror descent with implicit updates," *ArXiv:1009.3240v2*, 2011.
 - [29] F. Harper and J. Konstan, "The movielens datasets: History and context," *ACM Trans. Interactive Intell. Syst.*, vol. 5, no. 4, pp. 1–19, Dec. 2015.
- <http://files.grouplens.org/datasets/movielens/ml-latest.zip>, Dec. 2016.