IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received April 14, 2018, accepted May 29, 2018, date of publication June 6, 2018, date of current version June 26, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2843182

Towards Near Optimal WiFi Offloading With Uncertain Contact Duration

CHAO DONG[1], (Member, IEEE), ZHIMIN LI[2], YUBEN QU[3], QIHUI WU[1], (Senior Member, IEEE), SHAOJIE TANG[4], (Member, IEEE), AND ZHEN QIN[5]

[1]College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[2]Institute of China Electronic Equipment System Engineering Company, Beijing 100141, China
[3]Xi'an Research Institute of High Technology, Xi'an 710025, China
[4]Naveen Jindal School of Management, The University of Texas at Dallas, Richardson, TX 75080, USA
[5]College of Communications Engineering, Army Engineering University of PLA, Nanjing 210014, China

Corresponding author: Zhimin Li (zhiminleee1989@gmail.com)

This work was supported in part by the National NSF of China under Grant 61472445, Grant 61631020, Grant 61702525, and Grant 61702545, and in part by the NSF of Jiangsu Province under Grant BK20140076.5.

ABSTRACT Due to the simplicity of implementation, user-initiated Wi-Fi offloading becomes more and more popular, and naturally the benefits of users become the main optimization goal. We notice the inter-contact and intra-contact durations could be uncertain in reality by reason of the user mobility and network dynamics. The two uncertain durations can cause great impact on the benefit of users; however, they were either ignored or simply assumed to be deterministic in most previous works. In this paper, for the first time, we study Wi-Fi offloading problem with uncertain contact durations. The aim is to guarantee the benefit of users (delay and payment) without damaging operator's benefit (amount of the offloaded traffic) at the same time. We propose a multi-armed bandit (MAB)-based online offloading scheme (MABOO) to solve the problem and prove the near-optimality of MABOO in terms of the utility theoretically. Extensive simulations show that MABOO always approaches the optimal scheme and achieves higher utility as well as offloads more traffic compared with the minimal payment and on-the-spot-offloading schemes.

INDEX TERMS Traffic offloading, MAB, WiFi, contact duration.

I. INTRODUCTION

Nowadays, with the development of the innovative technologies, *e.g.*, 5G, IoT and mobile edge computing, the smart devices become more and more popular and bandwidth-hungry applications dominate the mobile data traffic gradually. Hence, the mobile cellular networks are often heavily loaded due to the explosive growth of mobile data traffic. According to Cisco's report [1], the global mobile data traffic has reached 3.7 exabytes per month by the end of 2015, and is expected to increase nearly 8-fold between 2015 and 2020. To address this issue, increasing the network capacity is the most effective method. However, due to the high cost and long construction period, the mobile cellular networks capacity grows at an extremely slow pace. Mobile traffic offloading, which uses complementary communication technologies such as WiFi [2], [3] or Device-to-Device (D2D) [4]–[6] to offload the traffic originally transmitted over cellular networks, seems to be a cost-effective and timely manner. Considering the popularization of WiFi networks, most of the mobile data traffic is likely to be offloaded using WiFi [7].

There are two main WiFi offloading approaches, namely, operator-initiated and user-initiated offloading. For the former, through collecting the information of traffic flows and alternative WiFi networks, the operators decide which and when the traffic should be offloaded to WiFi networks, with the hope of offloading traffic as much as possible [8]–[10]. Compared with operator-initiated offloading, the user-initiated approach is easy to implement and becomes more and more popular these years. It is natural to assume that users are selfish, meaning that they always pick a network technology for their own benefit. There have been some works to study how to guarantee the users' benefits preferentially. Cheung and Huang [11] proposed delay-aware WiFi offloading and network selection (DAWN) to achieve a good tradeoff between the user payment and its QoS characterized by the file transfer deadline. Wang and Wu [12] studied the opportunistic decision-making problem with a data utility delay model considering transmission cost and delay. Besides, Cheng *et al.* [13] proposed to offload cellular traffic of vehicular users through carrier-WiFi networks to maximize the utility of users and operators which is a combination of service payment and satisfaction of users.

From the above works, we can see that both the delay and payment have important influence on the benefit of users when using WiFi offloading. Considering both the user mobility and network dynamics, for a user, the duration between the two successive usages of WiFi service and the duration of utilizing an uninterrupted WiFi service, are two critical factors for WiFi offloading. Here we refer to the former as inter-contact duration and the latter as intra-contact duration. On one hand, the users would like to postpone the data transmission in the inter-contact duration to wait for WiFi offloading opportunities and save payment. On the other hand, if the inter-contact duration is too long or the contact time duration is too short, the benefit loss caused by the delay will be higher than the payment saved from WiFi offloading. Therefore, for the mobile users, a practical way to make the offloading decision is to take both contact durations with WiFi networks into consideration. In reality, the users usually do not know the exact contact durations which have been proven to be uncertain [14], [15]. To verify this, we analyze DieselNet traces [16] and find that both contact durations do not follow some common distributions, *e.g.*, exponential, Pareto and Gaussian, *etc*. Unfortunately, both contact durations were either ignored or simply assumed to be deterministic when making WiFi offloading decisions in most previous works.

In this paper, we study WiFi offloading problem with uncertain contact durations. To achieve the double-wins of both users and operators, we pursue to guarantee the benefit of users (delay and payment) without damaging operator's benefit (amount of the offloaded traffic) at the same time. The challenge mainly comes from the uncertainty of the two contact durations. To deal with this challenge, we model the WiFi offloading problem as a non-stochastic Multi-armed Bandit (MAB) [17] problem without assuming any specific distribution on both contact durations. An inter-contact duration combined with its following intra-contact duration is considered as an arm of the gamble machine, and a round is defined as the sum of an inter-contact duration and its following intra-contact duration. To capture the profits of both users and operators, we define a utility function as the difference between the payment saved from WiFi offloading and the cost of waiting for WiFi offloading opportunities. Then, we propose MABOO which is a MAB-based Online Offloading scheme. Based on the utilities of different contact duration pairs, MABOO sequentially chooses the inter-contact and intra-contact durations jointly at each round and decides the corresponding transmission policy, which can maximize the total utility in the long run. In conclusion, the main contributions of this paper are summarized as follows:

- We consider the effect of uncertain contact duration on WiFi offloading which aims at guaranteeing the benefit of users without damaging operator's benefit at the same time. To the best of our knowledge, we are the first to study WiFi offloading problem with uncertain inter-contact and intra-contact durations.
- We propose MABOO, which makes WiFi offloading decisions adaptively with MAB method to deal with the uncertain contact durations. We theoretically prove that MABOO can achieve near optimal performance in terms of the utility.
- We evaluate the performance of MABOO through extensive simulations. The results show that compared with the minimal payment offloading and on-the-spot-offloading schemes, MABOO achieves higher utility all the time and can offload more data traffic in most cases under the appropriate settings of the user requirements, the operator's pricing strategy and the data rate of both networks.

The remainder of this paper is organized as follows. The related works on WiFi offloading is reviewed in Section II. The system model and problem formulation are presented in Section III. We propose the MABOO scheme in details in Section IV and evaluate its performance in Section V. Finally, we conclude our work in Section VI.

II. RELATED WORKS

To tackle with the explosive growth of mobile data traffic, traffic offloading technology which mainly includes D2D-based and WiFi-based has drawn more and more attention these years. For D2D-based traffic offloading, except the works [4]–[6] from academia, in 2010, the 3rd Generation Partnership Project (3GPP) proposed Proximity Services (Prose) [18] to define relevant usage models and derive technical requirements for D2D within 3GPP LTE networks [19]. Note that except LTE-based D2D communications, ProSe can also be used to establish LTE-assisted WiFi-based D2D communications. Leveraging LTE, ProSe is able to provide client discovery function and set up WiFi link between the clients to offload cellular traffic. For WiFi-based traffic offloading, in 2011, 3GPP proposed ANDSF [20] in the Evolved Core Packet (ECP) to assist Mobile User (MU) in discovering non-3GPP access networks (*e.g.*, WiFi) and provide MU with rules and operator policies to connect to the non-3GPP access networks. As [21], most WiFi-based offloading schemes can be embedded in ANDSF to improve the performance of traffic offloading. With the development of ANDSF and the aggregation of LTE and WiFi, *e.g.*, LTE and WiFi Aggregation (LWA) and LTE WLAN Radio Level Integration with IPsec Tunnel (LWIP) in 3GPP Release 13, various WiFi-based offloading schemes from academia were suggested in the past few years.

At first, to alleviate the severe traffic congestion situation, the main concern of WiFi offloading was to offload as much traffic originally transmitted over mobile cellular networks as possible to WiFi networks [22]–[29]. Lee *et al.* [29] showed that WiFi networks can offload about 65% of cellular data traffic without any delayed transmission and the gain can raise 29% more when 1 hour or longer deadline is allowed. However, ignoring the benefits of users is out of place, for example, the users may not be willing to use WiFi offloading because the delay or the benefit from the delay is not satisfactory [9]. A time-dependent pricing scheme which encourages users to delay their traffic from the higher-price to lower-price time zone was proposed in [30]. Cheung and Huang [11] considered delay-aware WiFi offloading and showed that the optimal transmission policy exhibits a threshold structure in terms of both the remaining time and file size. They also demonstrated that WiFi offloading may not be a desirous option when user has a tight deadline constraint. In [8], an auction based incentive framework for downlink mobile traffic offloading was proposed. The users were proposed to send bids including both the delay it can tolerate and the discount the user wants to obtain for that delay, and the provider bought the delay tolerance from the users. In [13], two game-theory based WiFi offloading schemes were proposed to offload traffic of vehicular users through carrier-WiFi networks. In the auction game-based scheme, the operator sells WiFi access opportunities and the user submits the bid to buy WiFi access opportunity when the utility is positive. Comparatively, in the congestion game-based scheme, all users make offloading decisions based on utilities of other users and their own satisfaction. Furthermore, the utility decay model that considering both the delay and cost was adopted and the optimal downloading strategies were analyzed under the exponential and Gaussian distributions in vehicular environments in [12]. The transmission data size was assumed to be exact the same each time the user meets and contacts with a WiFi AP.

Although above works have studied how to guarantee the benefits of both users and operators at the same time from different viewpoints, most of them either simply ignored one of the contact durations or only assumed that they follow known distributions. In fact, due to the mobility of users and uncertainty of the network environments, the inter-contact and intra-contact durations are limited and uncertain, while this characteristic is considered in this work. In the following, we will introduce how to handle the uncertain contact durations with MAB method [17].

III. MODEL AND PROBLEM FORMULATION

In this section, we will introduce the system model firstly, then define a utility function and present the problem formulation in details. Table 1 provides a list of the major notations used in this paper.

A. SYSTEM MODEL

We assume that a MU needs Internet service to transmit its data. As shown in Fig. 1, a MU is moving in the coverage of the mobile cellular networks, *i.e.*, the cellular service is assumed always available to the MU. Occasionally, the MU may move into the coverage of WiFi APs, for example, in an office building or a coffee shop. At this time, the MU can select to use WiFi networks to access Internet. Defining p_c as the usage price of the mobile cellular networks, which is always higher than p_w (the usage price of the WiFi networks).

TABLE 1. A list of major notations.

Notation	Definition
t_E	inter-contact duration
t_A	intra-contact duration
ψ_i	the strategy MU decides on round i
a	the action MU chooses depending on ψ_i
U_a	the benefit achieved from action a in one round
P_a	the payment saved from action a in one round
C_a	the cost waiting for WiFi for action a in one round
$H(t)$	satisfaction function with delay t
q_c	the data rate of mobile cellular networks
q_w	the data rate of WiFi networks
p_c	the usage price of mobile cellular networks
p_w	the usage price of WiFi networks
D^i	the data size at the beginning of round i
D^i_R	the remaining data size after using cellular in round i
$s_{\psi_i}(i)$	the reward gained from strategy ψ_i at round i
$S_\emptyset(T)$	accumulative reward after T round with a static scheme $\emptyset$
$\hat{S}_{\hat{\emptyset}}(T)$	accumulative reward after T round with an online scheme $\hat{\emptyset}$
R_T	the regret after T rounds

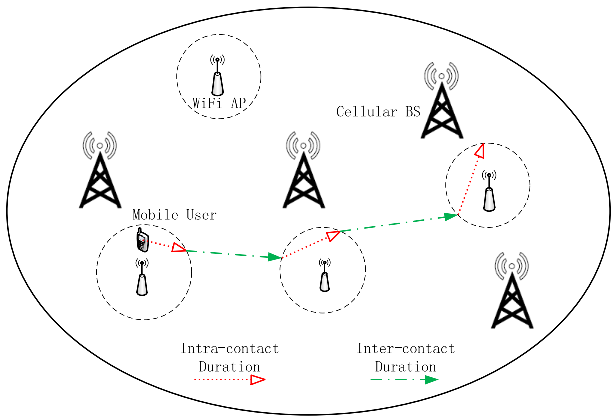


FIGURE 1. System model.

It is noted that this pricing strategy encourages the MU to offload its data traffic to WiFi networks so as to save data service payment and alleviate the traffic pressure of the cellular networks [9].

Since the WiFi service is location-dependent and not always available, the MU perhaps needs to wait for a period to use WiFi networks to access the Internet. We define this waiting time as inter-contact duration (t_E), *i.e.*, from the time the MU quits the service of previous WiFi AP to the time it begin to use WiFi service again. We also define intra-contact duration as t_A which denotes usage time of WiFi service when the MU enters the coverage of a WiFi AP. On one hand, the longer the inter-contact duration, perhaps more data traffic needs to be offloaded. Hence the MU can achieve more payment saving and the traffic pressure can be alleviated for operators. But on the other hand, waiting for WiFi offloading also brings the cost including payment. For example, the satisfaction will degrade when the MU waits too much time, *i.e.*, the inter-duration is too long. Here we introduce a satisfaction function $H(t)$ [8], which is a monotonously decreasing function with delay t, and represents the price that

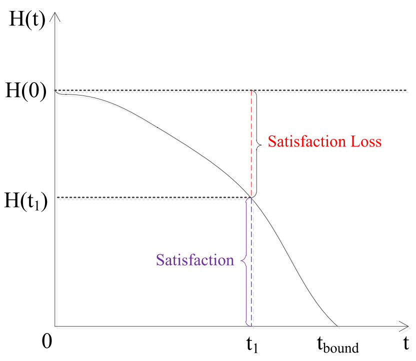


FIGURE 2. Satisfaction function.

the MU is willing to pay for the data with delay t. Fig. 2 shows an example of the satisfaction function $H(t)$, where t_{bound} is the upper bound of the MU's delay tolerance and depends on the decay speed of user satisfaction with delay t. Once the delay reaches the bound, the MU's satisfaction becomes zero, indicating that the MU is not willing to pay for the data. We can note that the highest user satisfaction $H(0)$ can be reached when delay $t = 0$. Without loss of generality, we define $H(0) = p_c D$ where D is the size of the data, that is to say, the MU chooses to transmit the whole data immediately using cellular networks and incurs no waiting delay. With delay t_1, the MU is only willing to pay $H(t_1)$, and $H(0) - H(t_1)$ is the satisfaction loss caused by delay t_1. Meanwhile, it is reasonable to think that the satisfaction loss is negligible as long as the MU can achieve Internet access all the time. Therefore, when the relation between two durations is not appropriate, the payment saving may be smaller than the cost and WiFi offloading perhaps is not a good choice for the users. In conclusion, for the user-initiated WiFi offloading, the MU should make offloading decisions based on its saved payment and cost which are relevant to the inter-contact and intra-contact durations.

Previous works either ignored one of both durations or treated them following known distributions, *e.g.*, assuming the inter-contact duration follows the exponential distribution or intra-contact duration follows Pareto distribution. In fact, because of the user mobility and network dynamics, both the inter-contact and intra-contact durations are uncertain. For example, when a MU arrive a new street, it does not know when it will meet and use a WiFi AP. Even for the old places, the network condition is dynamic which decide how the MU access the Internet. Hence, we consider the uncertainty of both contact durations jointly and do not assume any specific distributions in our work. In the following we formulate the WiFi offloading problem into a MAB-based decision making problem.

B. PROBLEM FORMULATION

Without loss of generality, we assume that time are slotted, and the MU needs to make a decision about how to access Internet when it quits the service of a WiFi AP. In other words, the decision round is defined as the duration which including an inter-contact duration and its following intra-contact duration. As mentioned above, the MU makes the offloading decision depending on both inter-contact and intra-contact durations. Let $\psi_i =< t_E^i, t_A^i >$ represent a strategy which is the inter-contact and intra-contact duration pair that the MU decides at the beginning of round i, we also assume that the two durations are bounded as $t_E^i \in \{1, \ldots, E_{max}\}$ and $t_A^i \in \{1, \ldots, A_{max}\}$, where E_{max} represents the longest inter-contact duration and A_{max} represents the longest intra-contact duration. Hence, we have $\psi_i \in \Psi = \{1, \ldots, E_{max}\} \times \{1, \ldots, A_{max}\}$ and $|\Psi| = E_{max} \times A_{max}$.

The strategy $\psi_i =< t_E^i, t_A^i >$ decides the action a of MU in round i. Specifically, we define $a \in \mathcal{A} = \{0, 1, 2\}$, where $a = 0$ means that the MU decides to use cellular networks before the next decision round, $a = 1$ means that the MU decides to use cellular networks first and switch to WiFi networks as long as a WiFi AP is available, and $a = 2$ represents the MU does not access Internet until a WiFi AP is available.

Next, we define a utility function U_a which mainly represents the benefit the MU obtains from the transmission action a in one round. It is the difference between the payment saving (P_a) when using WiFi offloading (intra-contact duration) and the cost (C_a) when waiting for WiFi offloading (inter-contact duration):

$$U_a = P_a - C_a. \tag{1}$$

Then, the utility can be calculated according to different transmission actions as follows.

1) $a = 0$. This action means that the MU will use cellular networks to transmit the data throughout this round and ignore the potential WiFi opportunities. The reasons for MU to select this action are mainly as follows. First, the inter-contact duration is perhaps too long, which leads to high cost of waiting for WiFi service. Second, perhaps the data rate of WiFi networks is not desired or the usage price of WiFi service is too high, which brings little payment saving. For $a = 0$, because the MU does not use WiFi service and can access the Internet all the time, both the saved payment from WiFi offloading and the cost for waiting the WiFi opportunities are 0, *i.e.*, $P_0 = 0$ and $C_0 = 0$. Hence, we define the utility U_0 for this transmission action as:

$$U_0 = 0. \tag{2}$$

2) $a = 1$. This action means that the MU will use cellular transmission in the inter-contact duration and then use WiFi offloading to transmit the remaining data in the following intra-contact duration. When the gap of data rate and usage price for both network services is reasonable relatively, the MU tends to select this action. Due to the MU can use network service all the time, as mentioned before, the loss of satisfaction can be ignored. Hence, the cost C_1 is the payment that the MU

use the cellular network in t_E^i:

$$C_1 = p_c \min\{q_c t_E^i, D^i\}, \tag{3}$$

where q_c denotes the transmission rate of cellular network. Then $q_c t_E^i$ denotes the data size that MU k transmits with cellular service and D^i represents the data size for user k at the beginning of the round i. Next, the saved payment from WiFi offloading for transmission action $a = 1$ is:

$$P_1 = (p_c - p_w) \min\{q_w t_A^i, D_R^i\}, \tag{4}$$

where q_w denotes the rate of the WiFi network, $q_w t_A^i$ denotes the data size that the MU transmits using WiFi service, and $D_R^i = \max\{D^i - q_c t_E^i, 0\}$ represents the remaining size of data for the MU after using cellular service in round i. Hence, the utility U_1 for action $a = 1$ can be computed as:

$$U_1 = (p_c - p_w) \min\{q_w t_A^i, D_R^i\} - p_c \min\{q_c t_E^i, D^i\}. \tag{5}$$

3) $a = 2$. This action means that the MU will delay its transmission in the inter-contact duration until WiFi service is available. When the MU can tolerate high waiting delay, or compared with WiFi service, either the data rate of cellular is too low or the usage price is too high, the MU may select this action. As there is no payment for using cellular networks in the inter-contact duration, the cost for this action is the user satisfaction loss from waiting for the WiFi offloading opportunities:

$$C_2 = H(0) - H(t_E^i). \tag{6}$$

The saved payment can be computed with the same method as the MU uses WiFi service in the intra-contact duration.
Then the utility U_2 for this transmission action can be computed as:

$$U_2 = (p_c - p_w) \min\{q_w t_A^i, D_R^i\} - (H(0) - H(t_E^i)). \tag{7}$$

Finally the MU will decide the action that maximizes the benefit, *i.e.*, the expected utility in one round:

$$a = \arg\max_{a \in \mathcal{A}} (U_a). \tag{8}$$

Note that we treat the payment for using cellular service in inter-contact duration as cost, hence the utility represents the benefit of both users and operators. In reality, at the end of round i, MU k can compute the reward, *i.e.*, actual utility, according to the happened transmission procedure. Let $s_\psi(i)$ denote the reward gained from a strategy ψ at round i and T denote the number of decision rounds during the whole data transmission procedure. For a static offloading scheme $\emptyset$ which uses fixed strategy ψ at each round, the accumulative reward up to round T can be represented as:

$$S_\emptyset(T) = \sum_{i=1}^{T} s_\psi(i). \tag{9}$$

For the online offloading scheme $\hat{\emptyset}$ which selects different strategy ψ_i at round i, the accumulative reward up to round T can be represented as:

$$\hat{S}_{\hat{\emptyset}}(T) = \sum_{i=1}^{T} s_{\psi_i}(i). \tag{10}$$

In our study, we aim to design an online offloading scheme $\hat{\emptyset}$ which can maximize the expected utility:

$$\max_{\hat{\emptyset}} \frac{\sum_{i=1}^{T} s_{\psi_i}(i)}{T}. \tag{11}$$

For MAB-based problem, the regret is utilized to evaluate the performance of the scheme. Defining the optimal static offloading scheme is the one which can obtain biggest accumulative reward among all the static schemes. As in [31], the regret after T rounds is defined as the difference between the accumulative reward achieved by the optimal static offloading scheme $\emptyset_o$ and the one achieved by the proposed online scheme $\hat{\emptyset}$:

$$R_T = S_{\emptyset_o}(T) - \hat{S}_{\hat{\emptyset}}(T). \tag{12}$$

A strategy whose average regret per round $\frac{R_T}{T} \to 0$ with probability 1 when $T \to \infty$ is a zero-regret strategy. Because the accumulative reward of the optimal static scheme is fixed, to maximize the expected utility, our objective is to design an online scheme $\hat{\emptyset}$ with regret as small as possible.

IV. DESIGN AND ANALYSIS

In this section, we develop a MAB-based Online Offloading (MABOO) scheme under uncertain inter-contact and intra-contact durations, which makes the WiFi offloading decision through MAB in an online fashion. First, we overview MABOO and present its design in detail. Then, we theoretically analyze the performance of MABOO.

A. OVERVIEW

Without assuming any known stochastic distribution for inter-contact and intra-contact durations, we propose MABOO to guarantee the benefit of users without loss of operator's benefit at the same time. To be specific, MABOO utilizes MAB [17] to choose the inter-contact and intra-contact duration pair at each round, which can achieve almost optimal utility in the long run. The key idea of MABOO is as follows. Initially, we guess an optimal strategy (*i.e.*, offloading decision). In the following rounds, with some specific probability, we execute the previously guessed strategy, otherwise, we try some other strategies in the whole strategy set. Based on the feedback, *i.e.*, the exact achieved utility during the current round, our guess can be adjusted dynamically.

Algorithm 1 MAB-Based Online Offloading Scheme (MABOO)

Parameters:

γ - Tradeoff between exploration and exploitation parameter, $\gamma > 0.5$

β - Strategy gain estimation error parameter, $\beta > 0$

η - Learning speed parameter, $\eta > 0$

Process:

1: Set all the strategies with the same weight, set $w_{t_E,t_A}(0) \leftarrow 1$, for all $1 \le t_E \le E_{max}$, $1 \le t_A \le A_{max}$ and $W(0) = E_{max} \times A_{max}$;

2: **for** round $i = 1$ to T **do**

3: Calculate the probability distribution of different inter-contact and intra-contact duration pairs for all $1 \le t_E \le E_{max}$, $1 \le t_A \le A_{max}$ according to

$$p_{t_E,t_A}(i) = (1-\gamma)\frac{w_{t_E,t_A}(i-1)}{W(i-1)} + \frac{\gamma}{E_{max}A_{max}}$$

4: Randomly select an inter-contact and intra-contact duration pair $\psi_i =< t_E^i, t_A^i >$ from Ψ according to the above probability distribution $p_{t_E,t_A}(i)$;

5: Decide the transmission action a_i for round i according to the utility function with ψ_i,

$$a_i = \arg\max_{a^j \in \mathcal{A}}(U_{a^j}), \quad \forall j \in \{0, 1, 2\}$$

6: Get the scaled reward $s_{t_E^i,t_A^i}(i) \in [0, 1]$ based on the actual transmission procedure at the end of this round,

$$s_{t_E^i,t_A^i}(i) = U'_{a_i}$$

7: Calculate the virtual reward $s'_{t_E,t_A}(i), \forall t_E, t_A$,

$$s'_{t_E,t_A}(i) = \begin{cases} \frac{s_{t_E,t_A}(i)+\beta}{p_{t_E,t_A}(i)}, & \text{if } t_E = t_E^i, t_A = t_A^i \\ \frac{\beta}{p_{t_E,t_A}(i)}, & \text{otherwise.} \end{cases}$$

8: Update the strategy weights $w_{t_E,t_A}(i)$ and the sum weight $W(i)$

$$w_{t_E,t_A}(i) = w_{t_E,t_A}(i-1)\exp\left(\eta s'_{t_E,t_A}(i)\right), \forall t_E, t_A,$$

$$W(i) = \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} w_{t_E,t_A}(i)$$

9: **end for**

B. DESIGN

The MABOO scheme is presented in Algorithm 1. Like most previous studies based on MAB [17], [32], we also employ a parameter γ to tradeoff between the *exploitation* and *exploration* in the problem of choosing the inter-contact and intra-contact duration pair. γ is a relatively small parameter, whose value is mainly determined by the number of rounds T. The meaning of *exploitation* and *exploration* here is explained in the following. At the beginning, we choose a random inter-contact and intra-contact duration pair from Ψ by setting equal weighs for all inter-contact and intra-contact duration pairs as illustrated in step 1, since we have no idea about the relationship between the gain and contact duration initially.

In the following rounds, with probability $1 - \gamma$, we will *exploit* the strategy used in the previous round. The exploitation is able to guarantee an almost optimal performance if the previous strategy is also almost optimal. On the other hand, we will *explore* new inter-contact and intra-contact duration pairs with probability γ, *i.e.*, choosing each inter-contact and intra-contact duration pair with equal probability $\frac{1}{E_{max}A_{max}}$. The exploration is also critical in the sense that it can eventually improve MABOO to approach the optimal solution.

The calculation of $p_{t_E,t_A}(i)$ in step 3 represents the aforementioned tradeoff, where the two parts in the left hand side denotes the exploitation and exploration, respectively. In step 4 and 5, the transmission is executed according to the transmission action determined by the selected inter-contact and intra-contact duration pair, whose scaled reward is calculated in step 6. To compensate the reward of the strategies that are unlikely to be chosen, MABOO adopts the virtual reward $s'_{t_E,t_A}(i)$ through adding the actual reward by β as shown in step 7. In step 8, we update all the weights for the corresponding strategies.

Worth noting that the parameter β is used to control the bias in the estimation of the contact duration pair reward $s'_{t_E,t_A}(i)$ and η is to control the learning speed. The values of γ, β and η are very important for the performance of MABOO, which will be discussed in the next subsection.

C. THEORETICAL ANALYSIS

In this subsection, we will study the performance of MABOO theoretically by analyzing its regret.

Theorem 1: The expect regret R_T of MABOO in round T is bounded as

$$R_T \le 6\sqrt{TLlnL} \tag{13}$$

with probability $1 - \epsilon$ for any $\epsilon \in (0, 1)$, when $\gamma = \sqrt{LlnL/T}$, $\beta = \sqrt{ln(L/\epsilon)/LT}$, $\eta = \sqrt{lnL/4TL}$, where $L = I_{max}D_{max}$.

Proof: Please find the detailed proof in Appendix A. ■

According to the above regret analysis, we can prove that MABOO is asymptotically reward optimal as shown in the following theorem.

Theorem 2: MABOO is asymptotically reward optimal when T is sufficiently large.

Proof: According to Eq. (13), we have the average regret per round as $R_T/T = 6\sqrt{LlnL}/\sqrt{T}$. When $T \to \infty$, according to Theorem 1, the per round regret $R_T/T \to 0$ with probability $1 - \epsilon$ for any $\epsilon \in (0, 1)$. In other words, when T is sufficiently large, MABOO is asymptotically reward optimal. ■

Remark: As shown in Theorem 2, the asymptotical optimality of MABOO is greatly influenced by the value of T.

To ensure the asymptotical optimality, let the average regret per round $\frac{R_T}{T}$ be no more than δ, where δ is a small value, *e.g.*, $\delta = 0.01$. Based on the conclusion of Theorem 2, we should solve the following inequality: $6\sqrt{L \ln L}/\sqrt{T} \leq \delta$, whose solution is $T \geq \frac{36\, L \ln L}{\delta^2}$. In other words, the number of rounds for such convergence is proportional to $\frac{1}{\delta^2}$.

V. PERFORMANCE EVALUATION

In this section, through extensive simulations, we evaluate the performance of MABOO by comparing it with three benchmark schemes in terms of the utility (accumulative reward, in \$) and the total offloaded traffic. We first introduce the evaluation setting and then present the evaluation results.

A. EVALUATION SETTING

The simulation evaluation is conducted based on the DieselNet traces [16] which were collected during Fall 2007 and we use the trace set between October 22 to November 16 2007 in the evaluation. In DiesleNet traces, the time when a WiFi connection between the MU and AP starts and the duration of the connection were recorded. Hence we can get the inter-contact and intra-contact durations through pre-processing on the original traces. Next, we use the traces including 34 MUs connecting with 301 APs. Different from location-dependent WiFi service, we assume that the cellular networks is always available.

We compare MABOO with other three benchmark offloading schemes as follows. Because the utility is zero, we do not consider the scheme of selecting $a = 0$ always in the evaluation.

- Optimal static scheme. To evaluate how MABOO approaches the optimal performance, based on the full information of both durations, we use brute-force searching method to get the optimal static offloading scheme which maximizes the accumulative reward.
- Minimal Payment Offloading (MPO) [12]. In MPO, to save more payment, the MU always postpones its data transmission to wait for the potential WiFi offloading opportunities until the satisfaction decreases to zero, then it uses the cellular networks to complete the remaining transmission. MPO can be treated as always selecting $a = 2$ and brings the user satisfaction loss into consideration. That is to say, ''always waiting for WiFi'' is a special case of MPO when the delay is lower than t_{bound}.
- On-The-Spot Offloading (OTSO) [29]. In OTSO, the MU uses WiFi service whenever the WiFi networks is available and users cellular networks when WiFi is unavailable. It is easy to see that OTSO is to select $a = 1$ all the time.

We evaluate the performance in terms of utility and total offloaded traffic which is represented by the percentage of offloaded traffic to the whole data traffic. The size of data is randomly chosen from 100 MB to 500 MB in each round. To make the results more convincing, we consider three different factors in the evaluation. First, the user satisfaction degrades more quickly when the delay sensitivity of a user becomes bigger which indicates the user is more delay intolerable. We use the satisfaction function $H(t) = H(0) - at^b$ where a is the parameter of delay sensitivity and b represents the manner of the satisfaction decay which is selected from 0.8 to 1.2 randomly. For example, $b > 1$, $b = 1$ and $b < 1$ mean the satisfaction decay in concave, linear and convex manners respectively. Second, the usage price of WiFi and cellular networks has the important impact on offloading decision. For convenience, we fix the usage price of the mobile cellular networks as \$6/Gbyte [11] and change that of the WiFi networks from \$0/Gbyte to \$1/Gbyte. As mentioned before, the usage price of WiFi networks is usually cheaper than that of cellular networks. Third, the data rate of both networks is critical for the user's selection. We assume the cellular data rate is 5 Mbps and change the data rate of WiFi networks from 2 Mbps to 50 Mbps in the evaluation. For each set of parameter choices, we run the simulations for 500 times to get reasonable results as possible. Because that each point in the following figures is the average result of 500 times simulations, it can be treated as the reflection of some probability and the line represents the trend of probability change.

B. EVALUATION RESULTS

1) DELAY SENSITIVITY

In this evaluation, we compare the performance of four offloading schemes under different delay sensitivity parameter a which is varied from 0.1 to 1. We set the usage price of WiFi networks as \$0.1/Gbyte and the WiFi data rate as 20 Mbps.

As shown in Fig. 3a and Fig. 3b, both the utility and total offloaded traffic decreases with the increase of delay sensitivity parameter a for all schemes except OTSO. MABOO achieves better performance than MPO and OTSO schemes and the maximum utility improvement is about 45.5% and 205.1% respectively. For MABOO, when the delay sensitivity is small, the MU tends to choose action $a = 2$ to achieve more payment saving. But when the delay sensitivity becomes larger, the MU perhaps cannot tolerate the satisfaction loss brought by the waiting delay, it will tend to choose action $a = 1$ in more cases and the utility decreases since more little payment saving can be achieved. We also observe from Fig. 3b, the total offloaded traffic by MABOO decreases with the increase of delay sensitivity. For OTSO scheme, because the MU uses WiFi service whenever the WiFi networks is available and can access Internet all the time, the delay sensitivity has no impact on its offloading decision and user satisfaction. Hence, the utility and the total offloaded traffic are unchanged with different delay sensitivity. We also note that the utility of MPO scheme becomes smallest when the delay sensitive increases. Under this scenario, compared with the other schemes, the MPO scheme brings more satisfaction loss due to it always postpones the transmission to wait for WiFi service until the satisfaction

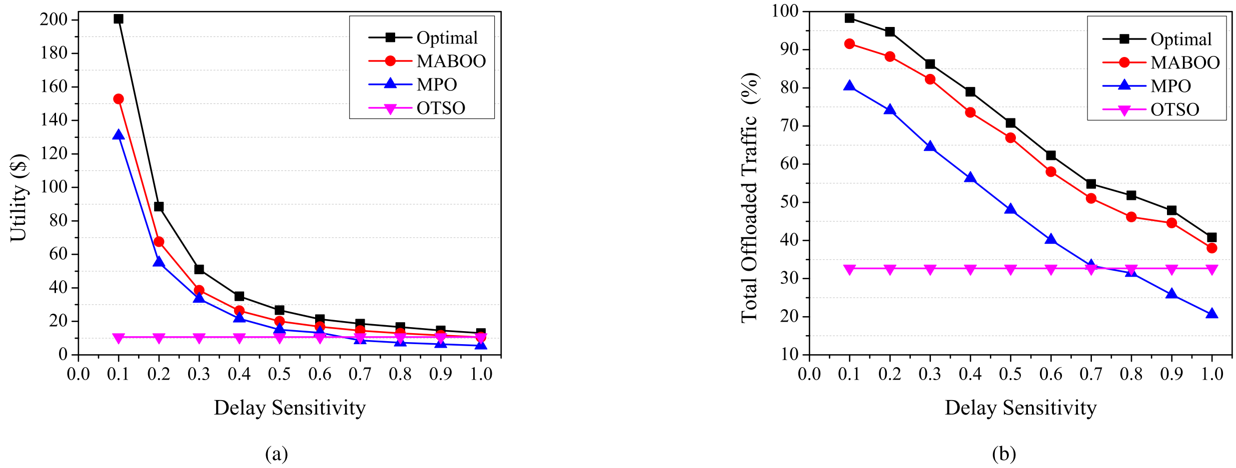


FIGURE 3. **Performance comparison under different delay sensitivity. (a) Utility versus delay sensitivity. (b) Offloaded traffic versus delay sensitivity.**

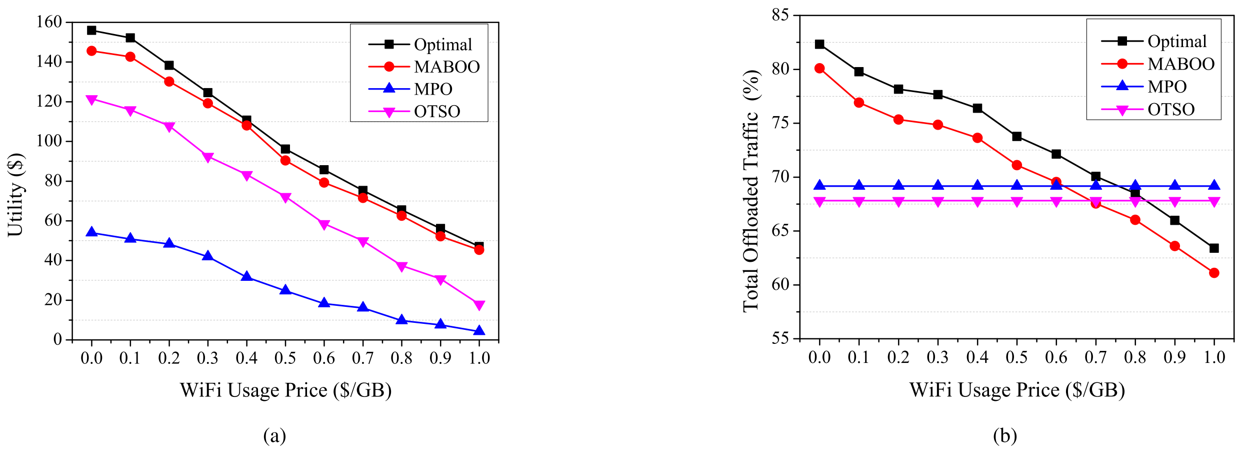


FIGURE 4. **Performance comparison under different WiFi usage price. (a) Utility versus WiFi usage price. (b) Offloaded traffic versus WiFi usage price.**

decreases to zero, and hence it has smaller payment saving because it will use cellular networks all the time as long as the satisfaction has become zero.

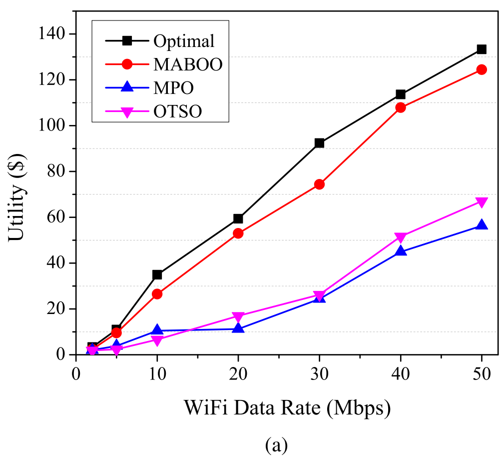
2) USAGE PRICE

In this evaluation, we fix the delay sensitivity parameter $a = 0.4$ and the data rate of WiFi networks as 20 Mbps.

As shown in Fig. 4a, the utility of all four schemes decrease when the usage price of WiFi networks changes from \$0/Gbyte to \$1/Gbyte. MABOO achieves almost the same utility with the optimal scheme, and on average it achieves 368.3% and 46.8% higher utility than MPO and OTSO schemes respectively. The reason is that the MU with MABOO can transmit its data with the most appropriate network service under different relation between the usage price of both networks. For MABOO, when the WiFi usage price is 0 or small, the MU is likely to choose action $a = 2$ to achieve more payment saving. While with the increase of the WiFi usage price, in most cases, the MU tends to choose action $a = 1$ even $a = 0$ due to the payment saving becomes more little and hence the utility decreases. From Fig. 4b, the total offloaded traffic decreases for MABOO and optimal schemes, while it is unchanged for MPO and OTSO schemes. For the former schemes, when the usage price of WiFi networks increases, although there is still huge gain and the MU prefers to offload the traffic using WiFi service, the payment saving from offloading will decrease, and the probability for MU selecting WiFi also goes down. This results in the decrease of the total offloaded traffic. For MPO and OTSO schemes, they do not consider the usage price when making an offloading decision, hence the total offloaded traffic keeps unchanged. We also note that MABOO offloads less traffic when the usage price of WiFi becomes higher than a threshold, *e.g.*, \$0.7/Gbyte in our study. This is because when WiFi service becomes expensive, it is beneficial to select cellular networks for the MUs.

3) DATA RATE

In this evaluation, we compare the performance of four schemes with changing data rate of WiFi networks from 2



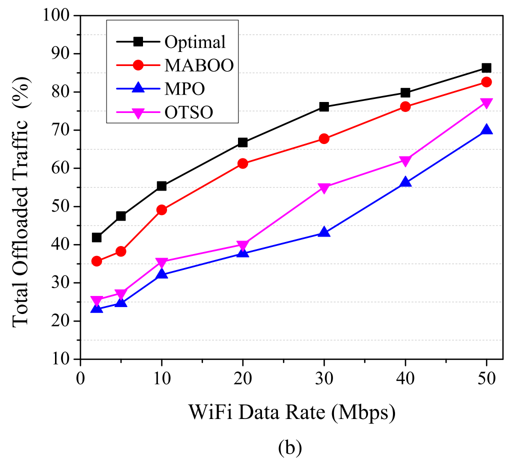


FIGURE 5. Performance comparison under different WiFi data rate. (a) Utility versus WiFi data rate. (b) Offloaded traffic versus WiFi data rate.

Mbps to 50 Mbps. We fix the usage price of both networks and set the delay sensitivity parameter $a = 0.4$.

The performance of different schemes is shown in Fig. 5. On average, MABOO achieves 55.3% and 59.7% higher utility than MPO and OTSO schemes respectively. When the data rate of WiFi networks is small, the MU with MABOO tends to take action $a = 1$ even $a = 0$ according to the relation between the payment saving from WiFi offloading and the cost of using cellular in the inter-contact duration. However, when the data rate of WiFi networks increases, the MU more tends to use WiFi service in the intra-contact duration, *i.e.*, $a = 1$ or $a = 2$, which can be proven by the result that the total offloaded traffic also increases when the data rate of WiFi networks increases. Meanwhile, due to the payment saving from WiFi offloading increases, the utility of MABOO increases with the increase of WiFi data rate.

4) APPLICABILITY OF MABOO

In reality, the biggest challenge for MABOO is its applicability when the MU experiences short trips. From Theorem 2, we prove that MABOO is asymptotically reward optimal when T is sufficiently large, and next we analyze the theoretical convergence speed of MABOO, which has the relation of quadratic power with δ, *i.e.*, $\frac{1}{\delta^2}$, where δ is an upper bound of the average regret per round. Therefore, theoretically, when the MU increases δ, the convergence time T can be greatly shortened with higher speed. Nevertheless, this is only the theoretical result when there is no any prior information. In fact, the MU can acquire some prior information in advance. As in [33], the MU can acquire some valuable information from ANDSF server, *e.g.*, historical contact durations, which can be used to speed up the convergence of MABOO. That is to say, it is just like that the MU has run MABOO for a lot of rounds in advance. Accordingly, these historical data is beneficial for the MU, especially when it usually experiences short but some similar trips. In addition, except the historical contact durations, the information about the wireless and geographical environment along the new trips is also useful for speeding up the convergence of MABOO.

In addition, as mentioned above, delay performance is an important user benefit. Then, how to identify delay-tolerant applications is critical for the applicability of MABOO. For well-known delay-tolerant applications, *e.g.*, email, MU can identify them through default port number. However, for some private applications, their port numbers are specified by the programmer or allocated randomly by the Operating System (OS), which indicates there are no default prot numbers for them. Fortunately, these applications can use Type of Service (ToS) field in the IP packet header to declare their performance requirements, and then MU can identify delay-tolerant applications through the ToS field of the received IP packet.

In conclusion, by using MAB to handle the uncertain inter-contact and intra-contact durations, MABOO achieves higher utility all the time and can offload more mobile data traffic at the same time in most cases. The double-win situation can be reached when the relation among the delay sensitivity of users, pricing strategy of the operators, and the data rate of WiFi and cellular networks is appropriate.

VI. CONCLUSION

In this paper, we have focused on the effect of uncertain inter-contact and intra-contact durations jointly on the WiFi offloading performance. We hope to pursue the double-wins situation under which both the benefit of users and operators are guaranteed. We modeled the problem with MAB method and proposed MABOO which can make the offloading decision online. We have proven that the performance of MABOO is near optimal which was also evaluated by extensive simulations. For the future work, we will take more factors into consideration to reflect the benefit of users, *e.g.*, the energy consumption with different network access techniques.

APPENDIX

PROOF OF THEOREM 1

Proof: Define $G_{t_E,t_A}(T) = \sum_{i=1}^{T} s_{t_E,t_A}(i)$, $G'_{t_E,t_A}(T) = \sum_{i=1}^{T} s'_{t_E,t_A}(i)$ as the actual gain and virtual gain for strategy (t_E, t_A) up to round T respectively. And the total gain up to round T of the chosen strategy sequence $(t_E^i, t_A^i)_{i=1,2,\ldots,T}$ is as follows: $\hat{G}(T) = \sum_{i=1}^{T} s_{t_E^i, t_A^i}(i)$.

Define $W(i) = \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} w_{t_E,t_A}(i)$. Since $w_{t_E,t_A}(i) = \prod_{i'=0}^{i-1} \frac{w_{t_E,t_A}(i'+1)}{w_{t_E,t_A}(i')} = \prod_{i'=1}^{i} e^{\eta s'_{t_E,t_A}(i')}$, we have $W(T) = \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} \prod_{i'=1}^{T} e^{\eta g'_{t_E,t_A}(i')} = \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} e^{\eta G'_{t_E,t_A}(T)}$ and $W(0) = E_{max}A_{max}$. We derive the bound of the regret R_T by using the quantity $\ln \frac{W(T)}{W(0)}$. For simplicity, we define $L = E_{max}A_{max}$.

For the lower bound, according to the definitions, we have

$$\ln \frac{W(T)}{W(0)} = \ln \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} e^{\eta G'_{t_E,t_A}(T)} - \ln L \geq \eta \max_{(t_E,t_A)\in\Psi} G'_{t_E,t_A}(T) - \ln L. \quad (14)$$

For the upper bound,

$$\begin{aligned} \ln \frac{W(i)}{W(i-1)} &= \ln \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} \frac{w_{t_E,t_A}(i)}{W(i-1)} \\ &= \ln \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} \frac{w_{t_E,t_A}(i-1)}{W(i-1)} e^{\eta s'_{t_E,t_A}(i)} \\ &\leq \ln\{\sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} \frac{w_{t_E,t_A}(i-1)}{W(i-1)} [\eta s'_{t_E,t_A}(i) + \eta^2 s'^2_{t_E,t_A}(i)]\} \\ &= \ln\{1 + \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} \frac{p_{t_E,t_A}(i)}{1-\gamma} [\eta s'_{t_E,t_A}(i) + \eta^2 s'^2_{t_E,t_A}(i)]\} \\ &\leq \frac{\eta}{1-\gamma} \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} p_{t_E,t_A}(i) s'_{t_E,t_A}(i) \\ &\quad + \frac{\eta^2}{1-\gamma} \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} p_{t_E,t_A}(i) s'^2_{t_E,t_A}(i). \end{aligned} \quad (15)$$

On the one hand, we have

$$\sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} p_{t_E,t_A}(i) s'_{t_E,t_A}(i) = s_{t_E^i,t_A^i}(i) + L\beta. \quad (16)$$

On the other hand,

$$\sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} p_{t_E,t_A}(i) s'^2_{t_E,t_A}(i) = (1+\beta) \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} s'_{t_E,t_A}(i), \quad (17)$$

which is due to $p_{t_E,t_A}(i) s'_{t_E,t_A}(i) \leq s_{t_E,t_A}(i) + \beta \leq 1 + \beta$.

Thus,

$$\ln \frac{W(i)}{W(i-1)} \leq \frac{\eta}{1-\gamma}(s_{t_E^i,t_A^i} + L\beta) + \frac{\eta^2(1+\beta)}{1-\gamma} \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} s'_{t_E,t_A}(i). \quad (18)$$

Summing for $i = 1, 2, \ldots, T$, we have the following inequality:

$$\begin{aligned} \ln \frac{W(T)}{W(0)} &\leq \frac{\eta}{1-\gamma}(\hat{G}(T) + L\beta T) \\ &\quad + \frac{\eta^2(1+\beta)}{1-\gamma} \sum_{t_E=1}^{E_{max}} \sum_{t_A=1}^{A_{max}} G'_{t_E,t_A}(T) \\ &\leq \frac{\eta}{1-\gamma}(\hat{G}(T) + L\beta T) \\ &\quad + \frac{\eta^2(1+\beta)}{1-\gamma} L \max_{(t_E,t_A)\in\Psi} G'_{t_E,t_A}(T). \end{aligned} \quad (19)$$

Combining the lower bound and the upper bound, we have

$$\frac{\eta}{1-\gamma}(\hat{G}(T) + L\beta T) + \frac{\eta^2(1+\beta)}{1-\gamma} L \max_{(t_E,t_A)\in\Psi} G'_{t_E,t_A}(T) \geq \eta \max_{(t_E,t_A)\in\Psi} G'_{t_E,t_A}(T) - \ln L. \quad (20)$$

That is also

$$\hat{G}(T) \geq (1 - \gamma - \eta(1+\beta)L) \max_{(t_E,t_A)\in\Psi} G'_{t_E,t_A}(T) - \frac{1-\gamma}{\eta} \ln L - L\beta T. \quad (21)$$

For any fixed $u > 0$ and $v > 0$, according to the Chernoff bound, we have $Pr[G_{t_E,t_A}(T) > G'_{t_E,t_A}(T) + u] \leq e^{-uv} E[e^{v[G_{t_E,t_A}(T) - G'_{t_E,t_A}(T)]}]$. Let $v = \beta$ and $u = \frac{\ln \frac{L}{\epsilon}}{\beta}$. Then, $Pr[G_{t_E,t_A}(T) > G'_{t_E,t_A}(T) + \frac{1}{\beta} \ln \frac{L}{\epsilon}] \leq \frac{\epsilon}{L}$.

Applying this bound, with probability at least 1 - ϵ, we have

$$\hat{G}(T) \geq (1 - \gamma - \eta(1+\beta)L)[\max_{(t_E,t_A)\in\Psi} G_{t_E,t_A}(T) - \frac{1}{\beta} \ln \frac{L}{\epsilon}] - \frac{1-\gamma}{\eta} \ln L - L\beta T. \quad (22)$$

By doing some transpositions and using the fact that $\max_{(t_E,t_A)\in\Psi} G_{t_E,t_A}(T) \leq T$, with probability 1 - ϵ, we have

$$\begin{aligned} \max_{(t_E,t_A)\in\Psi} G_{t_E,t_A}(T) - \hat{G}(T) &\leq (\gamma + \eta(1+\beta)L)T \\ &\quad + (1 - \gamma - \eta(1+\beta)L)\frac{1}{\beta} \ln \frac{L}{\epsilon} + \frac{1-\gamma}{\eta} \ln L + L\beta T \\ &\leq \gamma T + 2\eta T L + \frac{1}{\beta} \ln \frac{L}{\epsilon} + \frac{1-\gamma}{\eta} \ln L + L\beta T. \end{aligned} \quad (23)$$

If we set $\gamma = \sqrt{\frac{L \ln L}{T}}$, $\eta = \sqrt{\frac{\ln L}{4TL}}$, and $\beta = \sqrt{\frac{1}{LT} \ln \frac{L}{\epsilon}}$, then

$$\max_{(t_E,t_A)\in\Psi} G_{t_E,t_A}(T) - \hat{G}(T) \leq 6\sqrt{TL \ln L} \quad (24)$$

with probability 1 - ϵ. ■

REFERENCES

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update (2015–2020)*, Cisco, San Jose, CA, USA, 2016.
[2] F. Rebecchi, M. D. de Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, "Data offloading techniques in cellular networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 580–603, 2nd Quart., 2015.

[3] R. Maallawi, N. Agoulmine, B. Radier, and T. B. Meriem, "A comprehensive survey on offload techniques and management in wireless access and core networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1582–1604, 3rd Quart., 2015.

[4] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, Nov. 2014.

[5] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in LTE-advanced networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1923–1940, 4th Quart., 2015.

[6] P. Mach, Z. Becvar, and T. Vanek, "In-band device-to-device communication in OFDMA cellular networks: A survey and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1885–1922, 4th Quart., 2015.

[7] *Mobile Data Offload and Onload: Wi-Fi, Small Cell and Carrier-Grade Strategies 2013–2017*, Juniper, Sunnyvale, CA, USA, 2013.

[8] X. Zhuo, W. Gao, G. Cao, and S. Hua, "An incentive framework for cellular traffic offloading," *IEEE Trans. Mobile Comput.*, vol. 13, no. 3, pp. 541–555, Mar. 2014.

[9] J. Lee, Y. Yi, S. Chong, and Y. Jin, "Economics of WiFi offloading: Trading delay for cellular capacity," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1540–1554, Mar. 2014.

[10] H. Park, Y. Jin, J. Yoon, and Y. Yi, "On the economic effects of user-oriented delayed Wi-Fi offloading," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2684–2697, 2016.

[11] M. H. Cheung and J. Huang, "DAWN: Delay-aware Wi-Fi offloading and network selection," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1214–1223, Jun. 2015.

[12] N. Wang and J. Wu, "Opportunistic wifi offloading in a vehicular environment: Waiting or downloading now?" in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, Apr. 2016, pp. 1–9.

[13] N. Cheng, N. Lu, N. Zhang, X. S. Zhang, X. Shen, and J. W. Mark, "Opportunistic WiFi offloading in vehicular environment: A game-theory approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 1944–1955, Jul. 2016.

[14] X. Zhuo, Q. Li, W. Gao, G. Cao, and Y. Dai, "Contact duration aware data replication in delay tolerant networks," in *Proc. IEEE ICNP*, Vancouver, BC, Canada, 2011, pp. 236–245.

[15] Z. Lu, X. Sun, and T. La Porta, "Cooperative data offloading in opportunistic mobile networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, Apr. 2016, pp. 1–9.

[16] A. B. B. N. Levine and A. Venkataramani, "Enabling interactive applications for hybrid networks," in *Proc. ACM Mobicom*, San Francisco, CA, USA, Sep. 2008, pp. 70–80.

[17] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The non-stochastic multiarmed bandit problem," *SIAM J. Comput.*, vol. 32, no. 1, pp. 48–77, 2002.

[18] *Technical Specification Group Services and System Aspects; Study on Architecture Enhancements to Support Proximity-Based Services (Prose) (Release 10)*, document 3GPP TR 23.703, 2010.

[19] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, "Cellular traffic offloading onto network-assisted device-to-device connections," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 20–31, Apr. 2014.

[20] *Architecture Enhancement for Non-3GPP Accesses (Release 11)*, document 3GPP TR 23.402, 2011.

[21] H. Desta and K. Rudiger, "Study on performance-centric offload strategies for LTE networks," in *Proc. IEEE WMNC*, Dubai, UAE, Apr. 2013, pp. 1–10.

[22] B. Han, P. Hui, V. S. A. Kumar, M. V. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Trans. Mobile Comput.*, vol. 11, no. 5, pp. 821–834, May 2012.

[23] C. Wong, S. Sen, and S. Ha, "Offering supplementary wireless technologies: Adoption behavior and offloading benefits," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 1061–1069.

[24] S. Dimatteo, P. Hui, B. Han, and V. O. Li, "Cellular traffic offloading through wifi networks," in *Proc. IEEE MASS*, Valencia, Spain, Oct. 2011, pp. 192–201.

[25] Y. Choi, H. W. Ji, J.-Y. Park, H.-C. Kim, and J. Silvester, "A 3w network strategy for mobile data traffic offloading," *IEEE Commun. Mag.*, vol. 49, no. 10, pp. 118–123, Oct. 2011.

[26] B. Han, P. Hui, V. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular traffic offloading through opportunistic communications: A case study," in *Proc. ACM CHANTS*, Chicago, IL, USA, 2010, pp. 31–38.

[27] F. Mehmeti and T. Spyropoulos, "Is it worth to be patient? Analysis and optimization of delayed mobile data offloading," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, Apr. 2014, pp. 2364–2372.

[28] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting mobile 3G using WiFi," in *Proc. ACM MobiSys*, San Francisco, CA, USA, 2010, pp. 209–222.

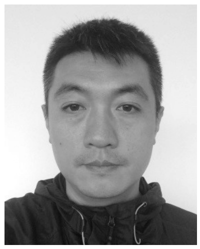
[29] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile data offloading: How much can WiFi deliver?" *IEEE/ACM Trans. Netw.*, vol. 21, no. 2, pp. 536–550, Apr. 2013.

[30] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, "Tube: time-dependent pricing for mobile data," *ACM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 247–258, 2012.

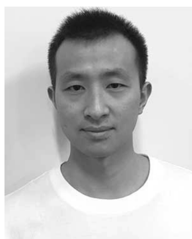
[31] Y. Qu, C. Dong, S. Tang, C. Chen, and C. Tian, "Towards near optimal network coding for secondary users in cognitive radio networks," in *Proc. IEEE SECON*, Singapore, 2014, pp. 522–530.

[32] Y. Qu, C. Dong, D. Niu, H. Wang, and C. Tian, "A two-dimensional multiarmed bandit approach to secondary users with network coding in cognitive radio networks," *Math. Problems Eng.*, vol. 2015, Jul. 2015, Art. no. 672837.

[33] K. Dae Sun, N. Yujin, K. Yoshinori, and Y. Hidetoshi, " Efficient ANDSF-assisted Wi-Fi control for mobile data offloading," in *Proc. IEEE IWCMC*, Sardinia, Italy, Jul. 2013, pp. 343–348.



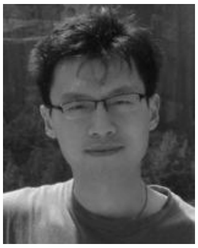
CHAO DONG (M'04) received the Ph.D. degree in communication engineering from the PLA University of Science and Technology, China, in 2007. From 2008 to 2011, he held a post-doctoral position with the Department of Computer Science and Technology, Nanjing University, China. He is currently a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include D2D communications, UAVs swarm networking, and anti-jamming network protocol. He is a member of ACM and IEICE.



ZHIMIN LI received the B.S. degree in information engineering, the M.S. degree in communication and information systems, and the Ph.D. degree in information and communications engineering from the PLA University of Science and Technology, in 2010, 2013, and 2017, respectively. He is currently an Engineer with the Institute of China Electronic Equipment System Engineering Company, Beijing, China. His research interests include mobile data offloading, D2D communications, and wireless networking.

YUBEN QU received the B.S. degree in mathematics and applied mathematics from Nanjing University in 2009, and the M.S. degree in communication and information systems and the Ph.D. degree in computer science and technology from the PLA University of Science and Technology, in 2012 and 2016, respectively. He is currently a Research Assistant with the Xi'an Research Institute of High Technology, China. His research interests include mobile edge computing, D2D communications, crowdsensing, and network coding.

QIHUI WU received the B.S. degree in communications engineering, and the M.S. and Ph.D. degrees in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1994, 1997, and 2000, respectively. From 2003 to 2005, he was a Post-Doctoral Research Associate with Southeast University, Nanjing, China. From 2005 to 2007, he was an Associate Professor with the Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China, where he served as a Full Professor from 2008 to 2016. In 2011, he was an Advanced Visiting Scholar with the Stevens Institute of Technology, Hoboken, USA. Since 2016, he has been a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests span the areas of wireless communications and statistical signal processing, with emphasis on system design of software defined radio, cognitive radio, and smart radio.



SHAOJIE TANG (M'07) received the B.S. degree in radio engineering from Southeast University, China, in 2006, the Ph.D. degree from the Department of Computer Science, Illinois Institute of Technology, in 2012. He is currently an Assistant Professor with the Naveen Jindal School of Management, The University of Texas at Dallas. His main research interests focus on wireless networks (including sensor networks and cognitive radio networks), social networks, security and privacy, and game theory. He has served on the Editorial Board of the *Journal of Distributed Sensor Networks*. He also served as TPC member of a number of conferences such as ACM MobiHoc, IEEE ICNP, and IEEE SECON.

ZHEN QIN received the B.S. degree in information engineering from Liaocheng University in 2017. She is currently a Graduate Student with the Army Engineering University of PLA. Her research interests include traffic offloading and D2D communications.