



Satisfaction-based Dynamic Bandwidth Reallocation for multipath mobile data offloading

Nuntanut Bhooanusas, Sok-Ian Sou^{*}, Kai-Chun Cheng

Department of Electrical Engineering, Institute of Computer and Communication Engineering, National Cheng Kung University, Tainan 70101, Taiwan

ARTICLE INFO

Keywords:

LTE
WiFi offloading
User satisfaction
Payment
MPTCP

ABSTRACT

Mobile data offloading is a key solution to mitigate heavy traffic load through cellular networks to the WiFi networks. To motivate the mobile users to transmit data over the cellular/WiFi networks under the control of the operator smartly, many incentive pricing frameworks have been proposed with the execution of dynamic bandwidth reallocation schemes to improve network efficiency. In this paper, a Satisfaction-based Dynamic Bandwidth Reallocation (SDBR) scheme is proposed to improve user satisfaction in term of the blocking probability and the user payment in data offloading environments. In SDBR, the operator can efficiently reallocate the bandwidth to accommodate more potential users; the overall operator revenue therefore increases as well. To demonstrate the effectiveness of our proposed SDBR scheme, we first assess the blocking probability with and without serving users having WiFi connections and discuss the benefits of the dynamic bandwidth allocation. Then, we elaborate a satisfaction-based pricing scheme for data downloading to motivate both mobile operator and users to work with SDBR, which brings the increment of operator overall revenue and reduction of individual user payment. Finally, simulation results show that our proposed SDBR scheme outperforms the four existing bandwidth allocation schemes in both time-based and volume-based data transmission.

1. Introduction

The explosive growth of data traffic over cellular networks is amongst one of the most engaging challenges for mobile operators. According to a recent Cisco report [1], the total number of global mobile subscribers is likely to increase from 5.1 billion (66% of population) in 2018 to 5.7 billion (71% of population) by 2023. There is growing concern that the massive increase in mobile communications may lead to a serious overload of existing cellular networks. Accordingly, *WiFi Offloading* [2,3] has been proposed to alleviate the traffic load over cellular networks, in which some amount of cellular data traffic belonging to a mobile user equipment (UE) is transmitted over a WiFi network.

Meanwhile, the Internet Engineering Task Force (IETF) has proposed the MultiPath Transmission Control Protocol (MPTCP) standard to integrate more TCP connections simultaneously transmitted over multiple physical paths in the same mobile device. Under the support of a stable cellular connection, when a user has WiFi connection, it is more efficient to use MPTCP to transmit data over both cellular and WiFi networks instead of simply using TCP either over cellular networks or WiFi networks. Such WiFi offloading approach based on MPTCP,

referred to as *multipath offloading* in this paper, is more efficient because it naturally enables the UE to transparently transmit mobile data through the cellular network along with the WiFi network. As a result, it not only alleviates data traffic congestion over the cellular network, but also improves the quality of service (QoS) such as reducing the data transmission time.

Several Dynamic Bandwidth Reallocation (DBR) schemes have been proposed to maintain the QoS level of a whole given service [4–8]. Under the operation of multipath offloading, the users can enjoy high speed connections only when they have WiFi connections, while some users without WiFi connection experience the poor network condition. To take a balance of the transmission quality, by a DBR scheme, the mobile operator reclaims some amount of cellular bandwidth from users with WiFi connection and then redistribute such reclaimed bandwidth to users with only cellular connection. As a result, the QoS level of all user categories is reasonably balanced.

In multipath offloading, it is essential to encourage users to connect with WiFi as much as possible. In order to avoid selfish WiFi Access points (APs) refusing user requests from cellular users, various incentive frameworks [9–12] have been proposed to persuade WiFi APs and/or users to accept data offloading from cellular networks under

^{*} Corresponding author.

E-mail addresses: nutt.nuntanut@gmail.com (N. Bhooanusas), sisou@mail.ncku.edu.tw (S.-I. Sou), jkaikai9420@gmail.com (K.-C. Cheng).

a pricing strategy such as ‘Paying for offloading’ [9]. To motivate the mobile users to smartly transmit data over the cellular/WiFi networks under the control of the operator, many incentive pricing frameworks have been proposed with the execution of dynamic bandwidth reallocation schemes to improve network efficiency. Under such an approach, a win-win situation is created in which WiFi APs are motivated by revenues from the cellular service providers, whereas cellular customers are willing to spend less on the data downloading that WiFi offloading is possible.

Several DBR approaches have been proposed for data offloading in cellular and WiFi networks, mostly regarding to the network performance such as the transmission latency and the deadline miss ratio. As shown in [13], it suggests that the performance of the bandwidth allocation can be evaluated by blocking probability, which is seldom to be discussed. A larger value of the blocking probability explicitly implies that more users are unsatisfied with the cellular networks. It is worth to note that the service price one is willing to pay to the operator is highly related to the service satisfaction that he/she actually gets. In [13], both *Equal Sharing* policy (i.e., allowing all users to share the pool of cellular bandwidth in the same BS) and the *Guaranteed Bit Rate* (GBR) policy (i.e., admitting a limited number of users to the network with fixed cellular bandwidth) have their own merits. Despite the policy enforcement, the QoS performance for the user also depends on whether the user connects with a WiFi network and the quality of the associated WiFi link. Although it seems that many users connect with both cellular and WiFi networks at the same time nowadays, the traffic load is still exponentially increasing.

Our main motivation is thus to assist the mobile operators to alleviate the traffic load in wireless networks, while the cellular resource is worthily allocated to every type of mobile users. For the QoS improvement, the main strategy is to persuade the mobile users to connect with WiFi so that the cellular resource is less required but the transmission delay is on the acceptable duration, while the surplus cellular resource is reallocated to hungry-bandwidth users or additional incoming users. Under this concept, the blocking probability of QoS sessions is reduced and the satisfaction of mobile users is accordingly improved. As such, the present study investigates the performance trade-off between the blocking probability, the user payment, and the operator revenue. Based on our results, the operator can further adjust the bandwidth dynamically in order to admit new users and the user satisfaction of a given service is improved as the growing incomes, which represents the delay tolerance that mobile users accept. We are the first to consider the user satisfaction in terms of the blocking probability as well as the user payment for the given service.

This paper proposes a Satisfaction-based Dynamic Bandwidth Reallocation (SDBR) method for multipath mobile data offloading in cellular and WiFi networks. The main goal is to maximize the revenue of the mobile operator and offer the users with incentives such as a reasonable price fitting their received satisfaction. Moreover, our proposed method is able to accommodate new incoming sessions that can generate greater incomes and decrease the blocking probability. The effectiveness of the proposed scheme is first evaluated by the blocking probability under two baseline schemes and then is compared to that of the existing static and DBR schemes [6,13,14] by means of numerical simulations.

The remainder of this paper is organized as follows. Section 2 discusses some related works regarding DBR on multipath data offloading and a pricing-based strategy. Section 3 explains the detailed concept of a pricing-based strategy. Section 4 describes the system model considered in the present study. Section 5 provides the details of the proposed SDBR scheme and Section 6 shows the theoretical analysis on blocking probability. Section 7 presents the numerical results. Section 8 discusses the remarkable results shown in theoretical analysis and numerical examples. Finally, Section 9 provides some brief conclusions and indicates the intended direction of future research.

2. Background and related work

In performing the DBR scheme, the main concept is to adjust the amount of cellular bandwidth allocated to the user dynamically in accordance with changes in their WiFi availability. Fig. 1 illustrates a typical scenario involving four UEs, namely UE1, UE2, UE3, and UE4. We assume that a user can always access the cellular network at all places, so every UE has a cellular connection. As shown, the BS distributes its total available bandwidth evenly among the four UEs such that each UE obtains a cellular bandwidth of $b_1 = 10$ Mbps. However, UE1 and UE2 also have the ability to connect to a WiFi network and receive an additional average bandwidth of $b_2 = 10$ Mbps. Consequently, UE1 and UE2 have a total aggregated bandwidth of $b_e = 20$ Mbps, and can therefore deliver their data more quickly than UE3 and UE4 which have a bandwidth of just 10 Mbps. Offloading traffic to WiFi networks can reduce the traffic load over the cellular network. In the scenario described above, the mobile operator is successful in maintaining the QoS provided to four users by adjusting the allocated bandwidth at the BS dynamically as their WiFi connection condition changes.

Recently, many studies have applied multipath data offloading in performing DBR schemes. In [6], a Dynamic Bandwidth Reallocation (DBR) algorithm was proposed to balance a cellular bandwidth for all users connecting to the BS in accordance with the agreed reallocation rate. Some cellular bandwidth belonging to users with WiFi connection is reclaimed to the sharing pool of the BS and then such reclaimed bandwidth in the sharing pool is equally redistributed to mobile users without the access of a WiFi network. Accordingly, such users have a higher bandwidth to complete their activities before the deadline transmission time. In [14], they improved the proposed DBR scheme by augments of a couple of allocation policies, namely an Earliest Deadline First (QEDF) allocation policy and a Smallest Download Size First (QSSF) allocation policy. By the QEDF policy, the surplus cellular bandwidth of users whose session is sufficient to meet QoS requirement is redistributed to hungry-bandwidth users with the earliest deadline. By contrast, in the QSSF policy, those users with the smallest volume of data remaining to be downloaded are prioritized to obtain such reclaim bandwidth. By performing QSSF and QEDF policies, the cellular bandwidth of the BS is fairly allocated to every users; as a result, the overall QoS level of a provided service is also improved.

However, WiFi offloading can also cause a traffic congestion over a WiFi networks, most mobile users with WiFi do not agree to reduce their obtained cellular bandwidth to help hungry-bandwidth users. Recently, many incentive frameworks has been applied in data offloading in order to mitigate a load of cellular bandwidths. In [15], they proposed a Q-learning method to provide the best WiFi AP to mobile users. Therefore, mobile user with WiFi can obtain the seamless connection during the period of the WiFi offloading process and the cellular residence time is accordingly reduced. The incentive framework is also applied in the Device-to-Device (D2D) offloading. [16] proposed D2D fogging in which the mobile operator motivates its users to share their resources to each other in order to minimize the time-average energy consumption for task executions. By doing so, users will be rewarded the improving capabilities of mobile devices particularly processing powers and network bandwidth. In the Internet of Vehicles (IoV), [17] presented the incentive mechanism with delay and cost constraints. By this method, the satellite operator motivates users to sell their idle resources to hungry-resource users for profits. In the future, the sellers will be rewarded with the additional resources by the satellite operator. In multipath offloading, [12] proposed IP Flow Mobility (IFOM) approach. Their algorithm allows mobile users to offload some data via WiFi path, while the rest of the data are transmitted via the cellular path with a pricing-based rate allocation. This method persuades users to avoid using the cellular network since such an action wastes their money. However, mobile users paying money for offloading via the

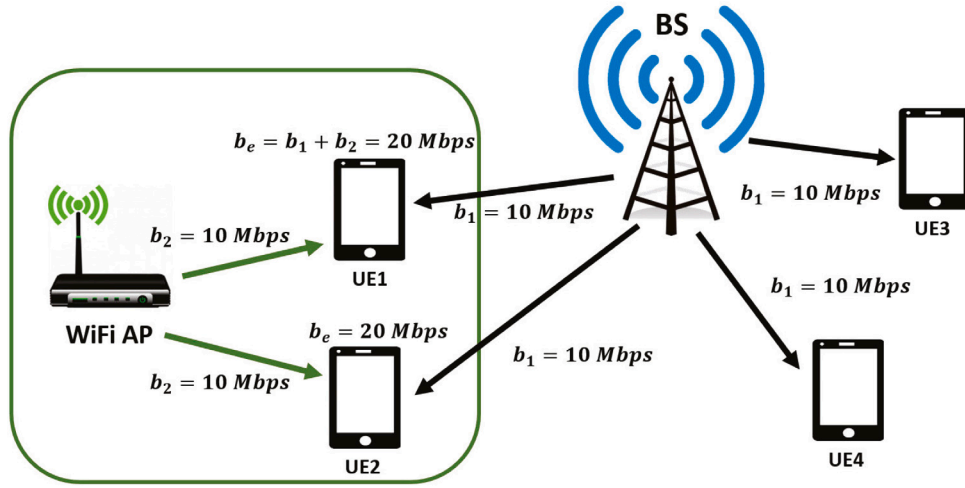


Fig. 1. An illustration of the unbalanced bandwidth by considering the WiFi connection.

cellular path will receive the impressive services throughout the entire periods of the services.

Also, a pricing-based scheme is amongst one of the most effective frameworks in improving the QoS level for WiFi offloading methods. Determining the price and creating acceptable policies can be done in several ways. [9] proposed an incentive framework to persuade mobile users to leverage their delay tolerance for cellular traffic offloading. Such a proposed scheme can be imagined as a reverse auction in which participants send their delay tolerance and desiring offloading volume for the bidding. Mobile users winning the auction are selected to join the offloading process with a designated delay time and a designated offloading volume while data of disqualified bidders are only performed via a cellular network path. In [18,19], they designed analytical models for delayed data offloading in heterogeneous networks to create an optimal offloading policy based on the network environment and QoS constraints; by such methods, the delay time is reduced and hence the user satisfaction is improved accordingly. Meanwhile, [20–22] used the concepts of the queuing model and the Markov chains to optimize policies for performing the delayed WiFi offloading activities such that users can complete their own activities with the suitable deadline time. [10,11,23,24] applied a game-theoretic approach, especially a Stackelberg game, in mobile data offloading to define an offloading policy based on an interaction between the service provider and the service consumers. Due to GBR policy for the cellular environment, the BS needs to restrict a number of UEs accessing its service since it is impossible for the mobile operator to guarantee the bit rate to every subscriber in the case that there are so many subscribers requesting services. In [13], the blocking probability is the output metric indicating the ratio of people that cannot access the BS service. The lower blocking probability can indicate the high level of user satisfaction.

The present study proposes the SDBR method for multipath mobile data offloading in cellular and WiFi networks. For comparison purposes, the effectiveness of the SDBR scheme is compared to those under the static bandwidth allocation algorithm [13] and three dynamic bandwidth reallocation algorithms [6,14]. Both [6] and [14] focused on reducing the deadline miss for QoS sessions, while [13] aimed to reduce the blocking probability. In the nature of mobile users, they are frustrated with the performance degradation (e.g., delay). These existing studies enable the network operator to estimate the resource for allocating to a particular session. However, none of them discuss the incentive strategy for encouraging users to join bandwidth sharing. In this work, an incentive pricing is afforded to motivate users to leverage their delay tolerance for cellular traffic offloading. Users are offered with incentives such as a reasonable price fitting their received satisfaction. Our proposed SDBR scheme takes both deadline

miss for QoS sessions and blocking probability in bandwidth sharing scheme. The bandwidth reallocation algorithms in [6,14] are improved by computing the optimal reallocation rate based on the transmission latency and the satisfaction function. In the remaining of the paper, we will show how to reduce the user payment and to increase the operator revenue as incentive to promote the SDBR scheme.

3. Satisfaction-based pricing

We select the user delay tolerance model to demonstrate the relationship between the degradation of the user satisfaction and the growth of the download delay. If the offloading process of mobile users can be completed before the deadline, they are willing to pay for the data service. If the delay reaches the bound, the user satisfaction becomes zero, indicating that users are not willing to pay for the data services. The user satisfaction in this paper is used to indicate the offloading price per user. There are different kinds of mathematics equations demonstrating the reverse relationship between the download delay and user satisfaction, namely sigmoid function and polynomial function. Sigmoid function in [25] has domain of all real numbers and the user satisfaction monotonically increases from 0 to 1. In the present study, we prefer to adopt a polynomial model in [9] because it is sufficient to see the relationship between the offloading price and delay.

In [9], the authors proposed to use a pricing strategy based on a function of the data transmission delay. Fig. 2 illustrates an example of function $S(t)$ of delay t , where the constant P_{\max} is the maximum downloading price of the satisfaction function and t_d is the maximum delay that a user can accept. Therefore, this figure shows $S(0) = P_{\max}$ and $S(t_d) = 0$. when delay is larger than t_d , the satisfaction becomes a negative value, which means that users do not satisfy the services. In the present study, we consider to assess the user satisfaction based on such a model. According to the user satisfaction of the service, the mobile operator can further announce the price and estimate the overall earning revenue for all serving users.

For the concave function shown in Fig. 2a, the user satisfaction reduces slowly as the time delay increases and drops to 0 as the time delay increases toward a critical value of $t = t_d$. This means that users are willing to wait for a certain period of time to complete their activities. This scenario is used to model the delay-tolerant activity (e.g., downloading files or updating software). By contrast, in Fig. 2b, the user satisfaction reduces rapidly and immediately as the time delay increases. This means that users cannot suffer a longer time delay to complete their activities. This scenario is that of delay-sensitive activity (e.g., live streaming services or Voice over Internet Protocol (VoIP) applications).

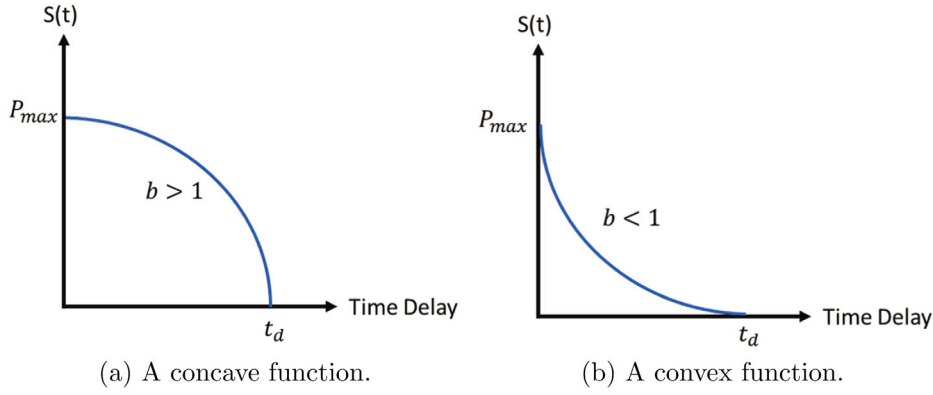


Fig. 2. A satisfaction function.

This paper takes the user satisfaction as the main consideration in determining the price of data downloading. In particular, the bandwidth allocation is induced in accordance with the function shown in Eq. (1), where a and b are scale and shape parameters of the function $S(\cdot)$, respectively.

$$S(t_l) = \begin{cases} P_{\max} - at_l^b, & \text{if } t_l < t_d \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

and

$$p = \mathbb{P}(S(t_l)) \quad (2)$$

In Eq. (1), a smaller scale parameter a results in a higher delay tolerance of the file transmission [9]. Also, the shape parameter b has a direct effect on the revenue accruing to the mobile operator. Referred to Fig. 2, when $b > 1$, $S(t)$ is a concave function. The revenue decreases slowly as the download time further increases (see Fig. 2a). On the other hand, when $b < 1$, a longer download time results in a rapid drop in the revenue (see Fig. 2b). The price p for a particular download session is then determined by the pricing function $\mathbb{P}(S(t_l))$ based on the user satisfaction of the session, which computed by the transmission delay t_l , the scale parameter a and the shape parameter b regarding to that particular type of the service. The pricing function is typically a monotonically increasing function of the input $S(t_l)$. That is, the user pays a higher price for the session of a shorter transmission delay.

In [26,27], they found that the satisfaction with price has a positive impact on the satisfaction with the provided service and [28] presented the complicated pricing model where users pay with high price in the case that they obtain desirable services. Therefore, this paper regards the satisfaction function with the download delay as the offloading price, i.e., $p = S(t_l)$. The sum of prices generated from each user is regarded as the mobile operator overall revenue and the average price computed from all users in the entire services is regarded as the user payment. Both output metrics can indicate the QoS performance in which users are willing to spend money for impressive services.

4. Analytical model

This section describes the system model of file downloading in multipath offloading. Fig. 3 shows the timing diagram for a typical file downloading service. Let random variables t_0 and t_1 be the respective time periods for which the user is *disconnected* from WiFi and *connected* to WiFi. Assuming the WiFi disconnection period t_0 follows an exponential distribution with rate μ_d . Similarly, the WiFi connection period t_1 follows an exponential distribution with rate μ_c . A download session commences at time τ_{start} and continues until the file is completely downloaded at τ_{end} . Thus, the downloading time, i.e., the transmission delay, is equal to $t_l = \tau_{\text{end}} - \tau_{\text{start}}$. In general, each download session designates the deadline requirement t_d . If the downloading time t_l is

Table 1

A notation list.

| Notation | Description |
|-------------------------------------|--------------------------------------------------------|
| W | Total available bandwidth of the BS |
| U_{BS} | Set of UE connecting the BS |
| u_i | UE device i |
| w_g | Guaranteed minimum bandwidth of UE |
| N_{max} | Highest number of UEs that the BS can handle |
| t_0 | WiFi disconnection period |
| t_1 | WiFi connection period |
| t_l | Download time |
| μ_d | The rate of the disconnection period |
| μ_c | The rate of the WiFi connection period |
| F | File size |
| α | Bandwidth reallocation rate |
| b_1 | Cellular bandwidth |
| b_2 | WiFi bandwidth |
| $b_1^i(t)$ | Cellular bandwidth of u_i at time t |
| $b_2^i(t)$ | WiFi bandwidth of u_i at time t |
| $b^i(t)$ | Total bandwidth of u_i at time t |
| S | Total available bandwidth for reallocation at the BS |
| $S(t)$ | User satisfaction function at time t |
| P_{\max} | Maximum downloading price for $S(t)$ |
| a | Scale parameter of $S(t)$ |
| b | Shape parameter of $S(t)$ |
| λ_u | User arrival rate at the BS |
| μ | Service rate for each user |
| $P_{N_{\text{max}}}$ | Blocking probability without WiFi |
| $P_{N_{\text{max}}}^{\text{Bound}}$ | Blocking probability that all of UEs connect with WiFi |
| $P_{N_{\text{max}}}^{\text{SDBR}}$ | Blocking probability under SDBR |
| P_1 | Probability of user starting with WiFi connection |

less than or equal to t_d , the UE pays the mobile operator for an agreed price p , which is calculated in accordance with the satisfaction function $S(t_l)$ given in Eqs. (1) and (2). The notation list is given in Table 1 for the ease of reading.

Let the user download sessions arriving at the BS be a Poisson process with rate λ_u . The allocated cellular bandwidth of UE u_i at time t is denoted by $b_1^i(t)$. The download session holding time is a random variable with exponential distribution of rate μ . Referring to Fig. 1, initially in the static bandwidth allocation, when u_i connects to the BS, the cellular bandwidth $b_1^i(t)$ of u_i is assumed to be:

$$b_1^i(t) = \frac{W}{N_{\text{max}}} = w_g, \tau_{\text{start}} \leq t \leq \tau_{\text{end}} \quad (3)$$

where W is the total available bandwidth of the BS and N_{max} is maximum number of served users in the BS.

In the static bandwidth allocation (i.e., the baseline scheme), the obtained cellular bandwidth is a constant during the whole session. Let $b_2^i(t)$ be the WiFi bandwidth of u_i . The WiFi bandwidth $b_2^i(t)$ changes time by time. If some of the users have WiFi connection, i.e., $b_2^i(t) > 0$, it is possible that the cellular bandwidth assigned to these devices

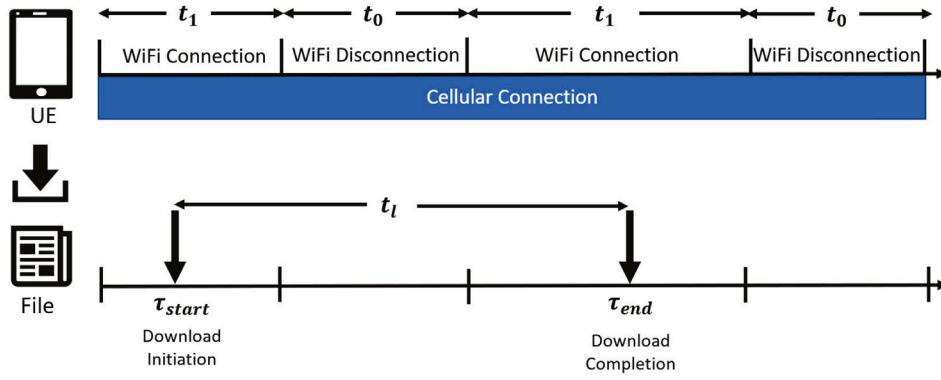


Fig. 3. The timing diagram for the download service with WiFi offloading.

can be partially reclaimed and then temporarily distributed to those devices which do not have WiFi connection. In the present study, we propose the SDBR scheme which assuming the cellular operator cooperates with the WiFi network and is thus able to determine the WiFi connection status of all the users connected to the BS. Let S be the amount of reclaimed cellular bandwidth available for reallocation. Let $b_1^i(\tau_{start}) = b_1^i = w_g$. For $\tau_{start} < t \leq \tau_{end}$, $b_1^i(t)$ changes as the status of WiFi connection changes. The relationship of the bandwidth, the transmission delay t_l , and the download file size F can be expressed as

$$F = \int_{t=\tau_{start}}^{\tau_{start}+t_l} (b_1^i(t) + b_2^i(t)) dt \quad (4)$$

According to the DBR scheme proposed in [6], let α be the cellular bandwidth reallocation rate for $0 \leq \alpha \leq 1$. When the parameter $\alpha = 1$, the cellular bandwidth from the UEs with WiFi is fully reallocated; when $\alpha = 0$, the BS does not perform any bandwidth reallocation (which is the standard setting for a static bandwidth policy). In this paper, we consider to use satisfaction-based pricing as the incentive for DBR. It is worth to note that individual user satisfaction is different, in order to increase the operator revenue and reduce the individual user payment, we assign different reallocation rate α_i to each UE $u_i \in U_{BS}$. For UE u_i of $b_2^i(t) > 0$, i.e., that with WiFi connection, the adjusted cellular bandwidth at time $t+1$ can be computed as:

$$b_1^i(t+1) = w_g - \alpha_i * [\min(w_g, b_2^i(t))] \quad (5)$$

If $b_2^i(t) = 0$ (i.e., no WiFi connection), no cellular bandwidth is reclaimed. If $b_2^i(t)$ is high (i.e., $b_2^i(t) > w_g$), the cellular bandwidth is reduced by $\alpha_i w_g$. In the case that $b_2^i(t)$ is small (i.e., $0 < b_2^i(t) < w_g$), the cellular bandwidth is reclaimed by just $\alpha_i b_2^i(t)$.

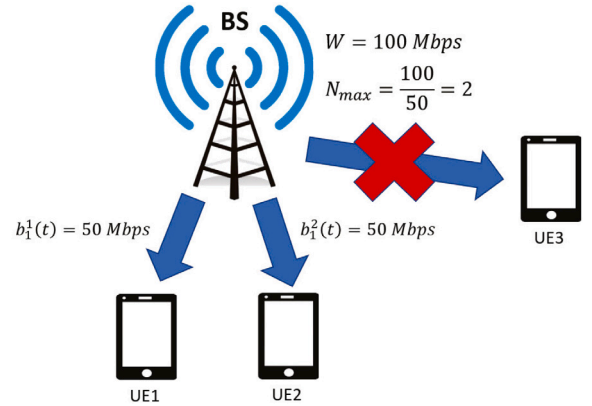
Meanwhile, the amount of cellular bandwidth contributed by UE u_i to S for reallocation is given by:

$$S = S + [b_1^i(t+1) - b_1^i(t)] \quad (6)$$

Eq. (6) simply states that the increment of returned bandwidth at time $t+1$ is allocated to the sharing pool S . Our goal is to effectively allocate S to more users with a higher economic benefits. We refer such benefits to the increment of operator revenue, and the reduction of the user payment and the session blocking rate. In the following, we firstly assess the blocking probability with and without serving users having WiFi connections and then we discuss how the mobile operator can utilize a satisfaction-based pricing scheme to manage reclaimed cellular bandwidth and to accommodate new sessions that can boost the operator overall revenue and reducing user payment.

4.1. Blocking probability

This subsection derives the blocking probability in DBR. In order to maintain the QoS, the BS should perform bandwidth allocation in

Fig. 4. An example of access control on BS, where $W = 100$ Mbps and $N_{max} = 2$.

accordance with a *Guaranteed Bit Rate* (GBR) policy. For example, each UE receives at least the guaranteed minimum bandwidth w_g . Hence, the maximum number N_{max} of UEs which can connect to the BS is obtained simply as $N_{max} = W/w_g$. Usually, the BS first distributes the available bandwidth W equally among the set U_{BS} of connected UEs with the same QoS level. When the total number of UEs connected to the BS reaches N_{max} , new arriving UEs are blocked and are unable to access its service as shown in Fig. 4; such blocking events can affect the user satisfaction of the system [13].

Let $P_{N_{max}}$ be the probability that an incoming download session cannot be accommodated at the BS in the original scheme (the baseline scheme) and let N_{max} , λ_u , and μ be the maximum number of users allowed in the system, the user arrival rate, and the service rate giving the average bandwidth b_1 , respectively. Then, we can compute the blocking probability via the classic Erlang-B formula. That is

$$P_{N_{max}} = B(N_{max}, \lambda_u/\mu) = \frac{(\lambda_u/\mu)^{N_{max}}/N_{max}!}{\sum_{k=0}^{N_{max}} (\lambda_u/\mu)^k/k!} \quad (7)$$

4.2. Revisit of blocking probability

Let $b_1 = E[b_1^i(t)]$ and $b_2 = E[b_2^i(t)]$. By considering the amount of serving users having WiFi connection of bandwidth b_2 , we revisit the blocking probability with two factors: (1) the reduction of the download session holding due to the aggregate cellular bandwidth (b_1) and the WiFi bandwidth (b_2); (2) the potential to reduce b_1 of those users with WiFi in order to increase the operator overall revenue as well as to reduce the average user payment by accommodating more new download sessions. As a result, the blocking probability value will decrease.

Regarding to factor (1), when all the sessions are with WiFi bandwidth $b_2 > 0$, the blocking probability reaches its lower bound, denoting by $P_{N_{\max}}^{\text{Bound}}$, where

$$P_{N_{\max}}^{\text{Bound}} = B \left(N_{\max}, \frac{\lambda_u}{\left(1 + \frac{b_2}{b_1}\right)\mu} \right) = \frac{\left[\frac{\lambda_u}{\left(1 + \frac{b_2}{b_1}\right)\mu} \right]^{N_{\max}} / N_{\max}!}{\sum_{k=0}^{N_{\max}} \left[\frac{\lambda_u}{\left(1 + \frac{b_2}{b_1}\right)\mu} \right]^k / k!} \quad (8)$$

The lower bound value of $P_{N_{\max}}^{\text{Bound}}$ indicates that when users are with WiFi bandwidth, the cellular bandwidth can slightly decrease without increasing the blocking probability. Suppose the session is with WiFi with probability

$$P_1 = \frac{E[t_1]}{E[t_1] + E[t_0]} \quad (9)$$

Let $\alpha = E[\alpha_i]$. The blocking probability under SDBR by reclaiming cellular bandwidth αb_1 from those having WiFi connections can approximate by

$$P_{N_{\max}}^{\text{SDBR}} = B \left(N_{\max}, \frac{\lambda_u}{\left(1 - \alpha P_1 + \frac{P_1 b_2}{b_1}\right)\mu} \right) = \frac{\left[\frac{\lambda_u}{\left(1 - \alpha P_1 + \frac{P_1 b_2}{b_1}\right)\mu} \right]^{N_{\max}} / N_{\max}!}{\sum_{k=0}^{N_{\max}} \left[\frac{\lambda_u}{\left(1 - \alpha P_1 + \frac{P_1 b_2}{b_1}\right)\mu} \right]^k / k!} \quad (10)$$

5. SDBR: Bandwidth reallocation

5.1. Workflow of proposed SDBR method

Fig. 5 presents a flow chart of the proposed SDBR method proposed in the present study. The corresponding pseudocode is shown in Algorithm 1. The main steps in the SDBR scheme are described in the following.

When a new user u_j requests a service from the BS, the operator first check whether the number of existing users n is less than the maximum number of users N_{\max} that the BS can support. If so, the BS distributes its available cellular bandwidth to u_j . In the event that u_j has WiFi connection, the cellular bandwidths of both u_j and the BS are reallocated using a SDBR_BW_Assignment function (See Algorithm 2). However, if N_{\max} has already been reached, the operator checks whether the cellular bandwidth S available for allocation is sufficient to meet the requirement of u_j . If the bandwidth is sufficient, the operator distributes the minimum guaranteed cellular bandwidth w_g to u_j and reduces S accordingly. If u_j has WiFi connection, $b_1^i(t)$ of u_j and S are both reallocated using the SDBR_BW_Assignment function, as aforementioned. The final case is that u_j is blocked to acquire the service if S is insufficient.

After u_j completes its service session, the operator will calculate the satisfaction from Eq. (1) and estimates its revenue accordingly.

5.2. Details of proposed SDBR procedure

Referring to Algorithm 1, the detailed steps in the proposed SDBR procedure are as follows:

Algorithm 1 SDBR Algorithm

```

1: Require:  $b_1, b_2, b_1^i(t)$  and  $b_2^i(t)$  for  $u_i \in U_{\text{BS}}$ 
2: Ensure:  $b_1^i(t+1)$  and  $b_2^i(t+1)$  for  $u_i \in U_{\text{BS}}$ 
3: {Collect bandwidth when initialize}
4: if  $n \leq N_{\max}$  then
5:   for all  $u_i \in U_{\text{BS}}$  do
6:     if  $b_2^i > 0$  then
7:       SDBR_BW_ASSIGNMENT( $u_i$ )
8:     end if
9:   end for
10: end if
11:
12: {Distribute bandwidth & Check Admission}
13: if  $S \geq w_g$  then
14:   Admission = True
15:   Add  $u_j$  to  $U_{\text{BS}}$ 
16:    $b_1^j(t+1) = w_g$ 
17:    $S = S - w_g$ 
18:   if  $u_j$  has WiFi connection then
19:     SDBR_BW_ASSIGNMENT( $u_j$ )
20:   end if
21: else if  $(S > 0)$  and  $(S < w_g)$  then
22:   Admission = True
23:   Add  $u_j$  to  $U_{\text{BS}}$ 
24:    $b_1^j(t+1) = S$ 
25:    $S = 0$ 
26: else if  $S == 0$  then
27:   Admission = False
28:   Block  $u_j$ 
29: end if

```

Algorithm 2 SDBR Bandwidth Assignment function

```

1: function SDBR_BW_ASSIGNMENT( $u_i$ )
2:    $\alpha_i = \text{FindReclaimRatio}()$ 
3:   if  $b_2^i(t) > 0$  then
4:      $r = \min(w_g, b_2^i(t)) * \alpha_i$  //reclaiming a cellular bandwidth
5:      $a = w_g - r$  //newly assigned bandwidth
6:   else
7:      $a = w_g$ 
8:   end if
9:   {Bandwidth Update}
10:  if  $b_1^i(t) \geq a$  then
11:     $S = S + (b_1^i(t) - a)$  //Return bandwidth to Sharing pool
12:     $b_1^i(t) = a$ 
13:  else
14:    if  $S > (a - b_1^i(t))$  then
15:       $S = S - (a - b_1^i(t))$  //Return bandwidth to user i
16:       $b_1^i(t) = a$ 
17:    end if
18:  end if
19: end function

```

- Lines 1–2: Cellular bandwidth of b_1 and $b_1(t)$ and WiFi bandwidth of b_2 and $b_2(t)$ for all of the UEs in U_{BS} are provided as inputs to the algorithm, and the algorithm then outputs cellular bandwidth at $t+1$ and WiFi bandwidth at $t+1$ (respectively denoted as $b_1^i(t+1)$ and $b_2^i(t+1)$) for all the connected UEs.
- Lines 3–10: The UEs in the set U_{BS} are all allocated the cellular bandwidth. If some of them have WiFi connections, their obtained cellular bandwidths are adjusted via Algorithm 2 and any surplus amount of bandwidth is returned to the BS before reallocating to all the UEs having only cellular bandwidths in the set U_{BS} .

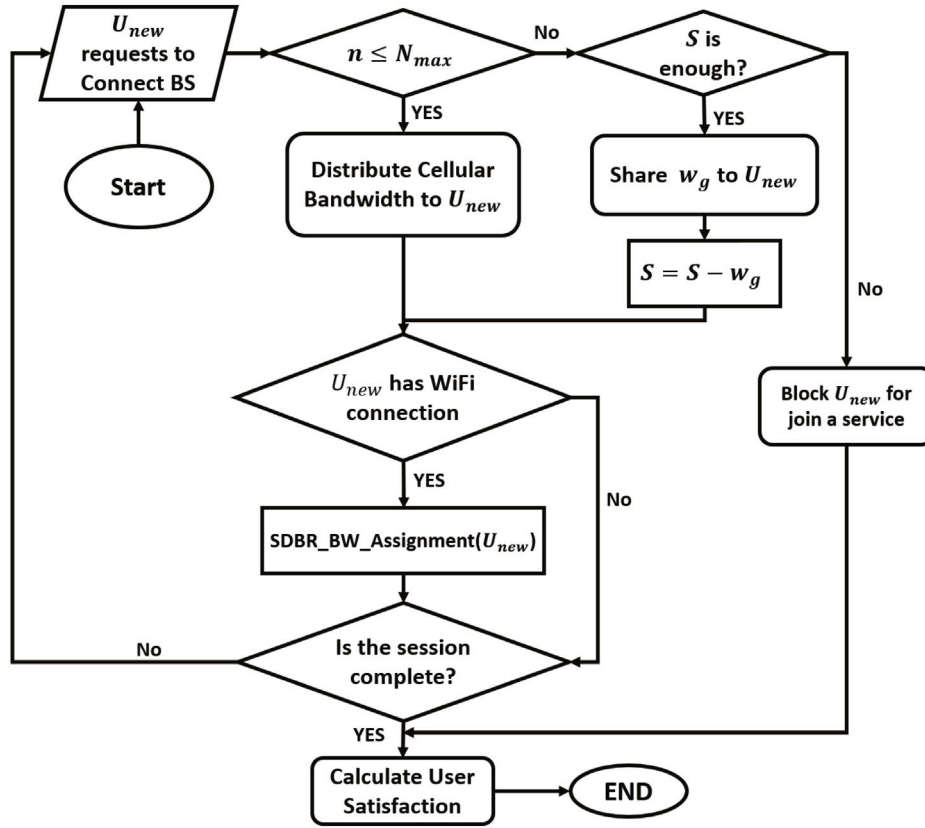


Fig. 5. The access control in the SDBR scheme.

Algorithm 3 Find Reclaim Ratio Function

```

1: function FindReclaimRatio()
2:    $t_l = \frac{F}{w_g}$  //Average transmission latency
3:    $G = S(t_l)$ 
4:    $t_{l,1} = \frac{F}{w_g(1-\alpha)+b_2}$  //Average transmission latency of WiFi
5:    $t_{l,2} = \frac{F}{w_g*\alpha}$  //Average transmission latency of Cellular
6:    $G' = S(t_{l,1}) + S(t_{l,2})$ 
7:   if  $G > G'$  then
8:      $\alpha_i = 0$  //Reclaiming Bandwidth is disabled
9:   else
10:    {Finding optimal  $\alpha_i$ }
11:     $\alpha^* = \operatorname{argmax}_{\alpha} (S(\frac{F}{w_g(1-\alpha)+b_2}) + S(\frac{F}{w_g*\alpha}))$ 
12:     $\alpha_i = \alpha^*$ 
13:   end if
14: end function

```

- Lines 13–29: When the additional UE u_j requests for a service after N_{\max} is reached, it can be happened as follows. First, if the bandwidth of the sharing pool of the BS S is sufficient to distribute w_g to that u_j , the admission value is set as true and then the u_j receives w_g . If that u_j has WiFi, Algorithm 2 is called for the bandwidth reallocation. Secondly, if S exists, but is insufficient to satisfy w_g , the admission value is defined as true and then all of the remaining cellular bandwidth in S is distributed to the u_j . Otherwise, the operator will designate the value of admission as false and then that u_j is denied for the service.

In the case that u_i has WiFi connection, the SDBR_BW_Assignment function described in Algorithm 2 is used to adjust the cellular bandwidth of u_i . It can be described as follows:

- Line 2: Call the FindReclaimRatio function to obtain the reallocation rate for the u_i (α_i) in which the satisfaction of the u_i is maximized.
- Lines 3–8: If $b_2^i(t) > 0$, the cellular bandwidth the u_i is reclaimed in accordance with α_i . After this, the u_i obtains the newly assigned cellular bandwidth a by subtracting the reclaimed bandwidth r from w_g . However, If $b_2^i(t) = 0$, the a value is set as w_g .
- Lines 10–18: The step of updating the cellular bandwidth of the u_i is performed. If $b_1^i(t) \geq a$, the cellular in the sharing pool S is retrieved by $S = S + (b_1^i(t) - a)$ and then $b_1^i(t)$ is updated as the value of a . In the case of If $b_1^i(t) < a$, the mobile operator will check whether the cellular resource of the sharing pool is sufficient to return the cellular bandwidth of $a - b_1^i(t)$ back to the u_i or not. If so, S is decreased by $a - b_1^i(t)$ and $b_1^i(t)$ is then equal to a . Otherwise, the u_i receives the original value of $b_1^i(t)$.

To find the reallocation rate α_i in which the satisfaction of the u_i is maximized, the function of FindReclaimRatio() is called. It can be explained as follows:

- Lines 2–3: Find the lower-bound user satisfaction $G = S(t_l)$ in which t_l is only calculated by F and w_g .
- Lines 4–6: Compute the user satisfaction after the bandwidth allocation G' in which the transmission latency values via the cellular and WiFi paths are considered.
- Lines 7–13: Find the value of α_i to be used in Algorithm 2. If $G > G'$, which means that the u_i satisfies using only w_g rather than using the SDBR schemes, α_i is set as 0, which means that the dynamic bandwidth allocation is disable. Otherwise, α_i is

Table 2

A default values of parameters for blocking probability analysis.

| Parameters | Values |
|------------------------------------------------------------------|---------|
| Arrival rate at the BS (λ_{μ}) | 3 |
| Service rate (μ) | 0.1 |
| Max available UE of BS (N_{\max}) | 10 |
| Cellular bandwidth (b_1) | 5 Mbps |
| WiFi bandwidth (b_2) | 10 Mbps |
| Probability that users will start with Wifi connection (P_1) | 68.5% |
| Reallocation rate (α) | 0.5 |

searched by the arguments of maximum on $(S(t_{i,1}) + S(t_{i,2}))$ in which α is the domain of the function.

5.3. Analysis of computational complexity of proposed SDBR algorithm

The time complexity analysis of Algorithm 1 can be explained as follows. Lines 1–10 is the step of bandwidth collection of each UE in the U_{BS} . For mobile users with WiFi in the U_{BS} , Algorithm 2 stated at Line 7 is called to reallocate their cellular bandwidth in order to maximize the user satisfaction of the entire system. After the step of the collection, Algorithm 1 handles additional users requesting a service given in Lines 13–20. If the additional user have WiFi, that additional users need to be reclaimed the cellular bandwidth to the sharing pool by Algorithm 2 stated at Line 19. Therefore, the computational complexity of Algorithm 2 is required to be discussed first.

Regarding Algorithm 2, at Line 2, in order to maximize the user satisfaction, Algorithm 3 is called to get the value of α_i in which the maximum of user satisfaction is computed. Once α_i is obtained, the cellular bandwidth of the u_i is reclaimed in accordance with α_i . The step of reclaiming cellular bandwidth is given in Lines 2–8. After this, $b'_i(t)$ is updated in accordance with Lines 10–18. For the step of bandwidth update, each line is executed only 1 time, while each line for reclaiming step except Line 2 is also executed only 1 time. Hence, we need to consider the time complexity of Algorithm 3.

Algorithm 3 is the step of calculating the reallocation rate of u_i in which the satisfaction of user i is maximized. To compute the satisfaction, Lines 2–6 are to compute the lower-bound transmission latency given at Line 2 and the transmission latency after the reallocation given in Lines 4–5. After the user satisfaction computation in Lines 7–13, the α_i is obtained by the argument of maximum of the user satisfaction in which α is the domain. According to [29], the time complexity analysis of a search algorithm for unsorted data is $\mathcal{O}(n)$. Because α is unsorted data, the time complexity at Line 11 is $\mathcal{O}(n)$. Therefore, the time complexity analysis of the function of FindReclaimRatio is $\mathcal{O}(n)$.

For Line 2 in Algorithm 2, the time complexity analysis is accordingly $\mathcal{O}(n)$ and the worst case of Line 4 is executed for n times due to unsorted input data. For the other lines are individually executed only 1 time. Therefore, the time complexity analysis of Algorithm 2 is $\mathcal{O}(n)$. Computing the time complexity of Algorithm 1, the time complexity at Line 7 and Line 19 is $\mathcal{O}(n)$. However, a For loop shown in Lines 5–9 is executed for n time, which is the number of users in the U_{BS} , at most. In conclusion, the time complexity of the overall Algorithm 1 is accordingly $\mathcal{O}(n^2)$.

6. Analysis on blocking probability

In this section, the theoretical analysis of the blocking probability is demonstrated. For comparison purposes, it is demonstrated with the other two different schemes, namely baseline and lower-bound schemes. The baseline scheme allows only N_{\max} users to participate the offloading activities and no users that have WiFi, while the lower-bound scheme allows mobile users equipped with both cellular and WiFi connections to join such activities. In the SDBR scheme, we assume that a user has WiFi connection with the probability P_1 and

the reallocation ratio α . Therefore, we use the Erlang-B formula and Eqs. (7)–(10), which are for the blocking probability analysis, to obtain the results. We summarize the setup of relevant parameters in Table 2. In Fig. 6a, there are only cellular users in the baseline scheme, so increasing b_2 does not affect the change of the blocking probability and it accounts for the highest rate of three models roughly at 0.6813 for all higher values of b_2 . Meanwhile, with the increment of b_2 values, the blocking probability values under the SDBR scheme decrease rapidly because the increment of b_2 values can allow the mobile users complete their download data earlier. As a result, the mobile operator can accommodate more new users/sessions and can earn more revenue accordingly. For example, when $b_2 = 10$ Mbps, there is a decrease of 0.1777 in the rate of blocking probability under the SDBR scheme while that under the low-bound scheme drops by 0.1957. The lower-bound scheme achieves a faster drop rate since the SDBR scheme reclaims some cellular bandwidth from WiFi users in order to reallocate to users with no WiFi while the lower-bound scheme does not. Hence, every user using the service of the lower-bound scheme is equipped with the greater aggregated bandwidth than users using that of the SDBR scheme. However, the blocking probability values under both SDBR and lower-bound schemes are almost identical, around 0.0015, when $b_2 \geq 40$ because that amount of b_2 is so sufficient that almost all users under both schemes complete the activities prior to the deadline.

As shown in Fig. 6b, since no users have WiFi in the baseline scheme, there is no change in the rate of blocking probability. It accounts for the highest value of 0.6813 for all values of b_2 . Meanwhile, every user in the service of the lower-bound scheme has WiFi connection with the probability P_1 of 1, which indicates that every user definitely starts the process with WiFi connection, which results in the most smooth connection of the three methods owing to the greater aggregated bandwidth than the others. Hence, the lower-bound scheme achieves the lowest blocking probability of the three methods at around 0.2146 for all values of P_1 . For the SDBR scheme, when increasing P_1 values, the blocking probability values decrease. For example, when $P_1 = 0.1$, the blocking probability value starts nearly at the same value as the baseline scheme begins because the small value of P_1 results in a lack of a WiFi use and the data traffic is congested in a cellular path accordingly. However, with the increment of P_1 values, mobile users connect with WiFi more frequently. Such an action results in the rapid decrease in the blocking probability values. For example, below 50% of the blocking probability values occur when $P_1 \geq 0.45$ and such values will be less than 40% when $P_1 \geq 0.7$.

Referring to Fig. 6c, there is no dynamic bandwidth reallocation mechanism in both baseline and lower-bound schemes, so increasing the reallocation rate α does not result in the change of the blocking probability value. Accordingly, the highest and lowest blocking probability values of the three methods are still achieved by the baseline and lower-bound schemes, respectively. As we mentioned in Section 4.2, the reclaiming cellular bandwidth is directly proportional to the α value. Higher value of α results in the greater reclaiming bandwidth from WiFi users. Hence, when some cellular bandwidth of WiFi users is heavily reclaimed, such users are at risk that their offloading activities do not complete by deadline. Accordingly, increasing α value results in the marked growth in the blocking probability value under the SDBR scheme since so many users are still in the service with lengthy download times. For example, it increases from 0.3365 to 0.4848 in all cases of α values.

7. Numerical examples

The effectiveness of the proposed SDBR scheme was evaluated by means of C++ distribution simulations with three output metrics, namely the mobile operator overall revenue, the average user payment, and the blocking probability. For reference purposes, the performance of our SDBR scheme was compared with that of four existing bandwidth allocation algorithms, namely a static mechanism [13] in which the

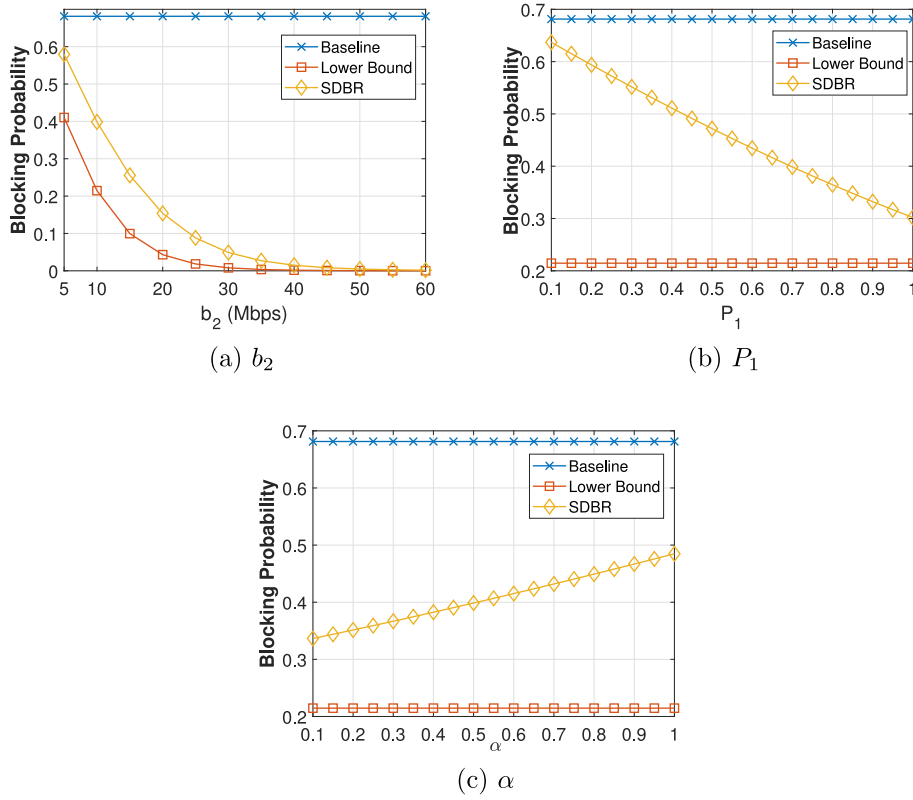


Fig. 6. A comparison of blocking probability.

Table 3

A basic simulation settings for file downloading scenarios.

| Parameter | Time-based | Volume-based |
|------------------------------------------------------------------|---------------|---------------|
| Max available UE of BS (N_{\max}) | 10 | 10 |
| BS total bandwidth (W) | 300 (Mbps) | 200 (Mbps) |
| Guaranteed bandwidth (w_g) | 30 (Mbps) | 20 (Mbps) |
| WiFi bandwidth (b_2) | 60 (Mbps) | 60 (Mbps) |
| Mean connection time of Wifi ($E[t_1]$) | 1679 (s) | 1679 (s) |
| Mean disconnection time of Wifi ($E[t_0]$) | 439 (s) | 439 (s) |
| Probability that users will start with Wifi connection (P_1) | 68.50% | 68.50% |
| Reallocation rate (α) | 0.5 | 0.5 |
| Deadline (t_d) | 500 (s) | 600 (s) |
| Mean file size | 187.5 (MB) | 3 (GB) |
| Pareto shape parameter | 1.2 | 1.8 |
| b of satisfaction function | 1.2 | 1.2 |
| P_{\max} of satisfaction function | 1 | 1 |
| P_s | 0.9 | 0.9 |
| Run time | 3 000 000 (s) | 3 000 000 (s) |

bandwidth was allocated initially and then remained unchanged until deadline, and three dynamic allocation mechanisms, namely DBR [6], QEDF [14], and QSSF [14].

For the sake of realism, the simulations were demonstrated with two offloading scenarios, namely time-based file downloading and volume-based file downloading, and we used the WiFi dataset given in [30] to respectively compute the average WiFi connection and disconnection periods as $E[t_1] = 1679$ s and $E[t_0] = 439$ s, where t_1 and t_0 had exponential distributions with rate μ_c and μ_d , respectively. Table 3 summarizes the basic parameter settings of a couple of scenarios. Note that the unit of the overall revenue and the user payment is P_{\max} and we let $P_{\max} = 1$, so the unit is not shown in our graphs.

7.1. Time-based file downloading

The time-based file downloading scenario is modeled for delay-sensitive applications such as live-streaming video or multimedia teleconferencing. Mobile users in this scenario are not willing to suffer a lengthy downloading time t_l . If t_l is longer than the deadline time of t_d , the mobile operator will lose the revenue in accordance with the user satisfaction function. Referring Table 3, for each user session, the generation of the file size in this scenario is in accordance with $w_g \cdot t_s$ where t_s is the length of session time exponentially distributed with rate λ_f . About the user satisfaction computation (see Eq. (1)), the maximum service charge P_{\max} and the shape parameter b were set as 1 and 1.2, respectively. For QEDF and QSSF methods, we followed [14] to set the parameter P_s , the percentage of sessions which complete transmission before deadline, as 0.9. Finally, the deadline period t_d for each session is designated as 500 s.

7.1.1. Effect of user arrival rate

In Fig. 7, overall, increasing the user arrival rate λ_μ results in a growth in the mobile operator overall revenues and the blocking probability values, while there is a decrease in the rates of the average user payments.

As shown in Fig. 7a, when there are a few mobile users connecting to the BS, the mobile operator overall revenues under the five schemes are similar since there is no load of data traffic over wireless networks. For example, all the mobile operator overall revenues under the five methods increase from roughly 27 700 for λ_μ of 0.01 to roughly 51 000 for λ_μ of 0.02. However, it is clearly seen that when $\lambda_\mu \geq 0.05$, the SDBR scheme results in the continual growth in the rate of the overall revenues and achieves the highest revenues of the five methods, roughly 92 000. This shows that the SDBR scheme uses the reclaim policy to accommodate more additional users in order to generate higher revenues, which is described in Algorithms 2 and 3. Meanwhile, the static and DBR schemes accept only N_{\max} mobile users and hence they

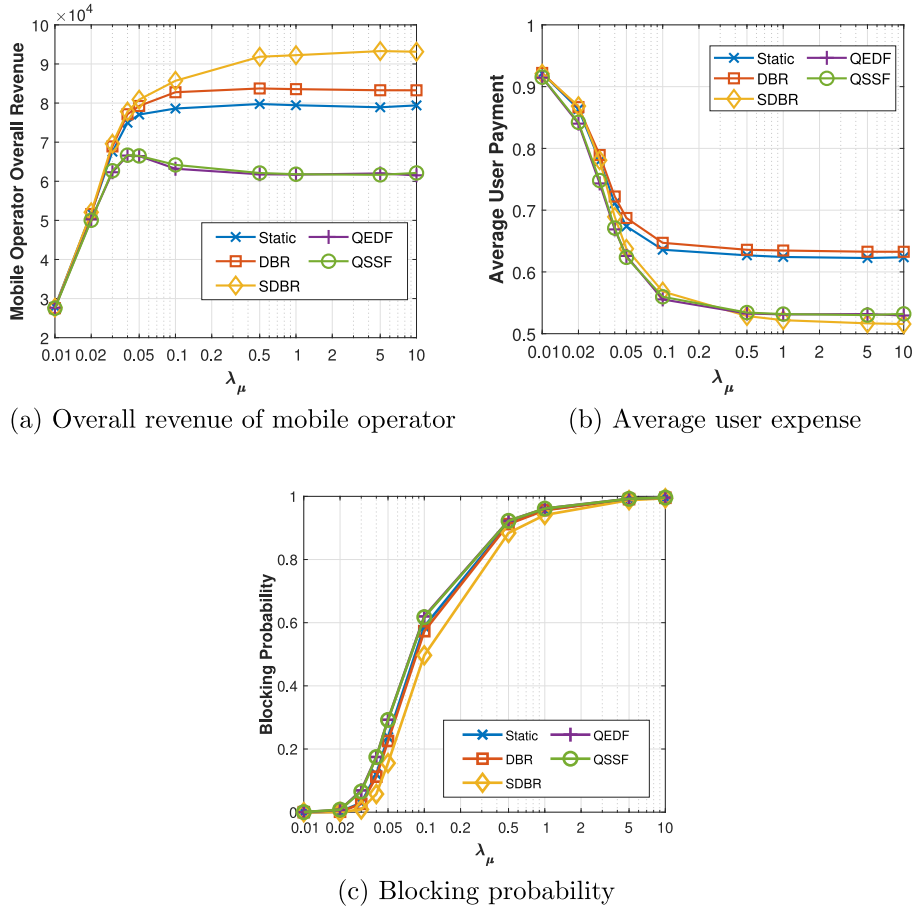


Fig. 7. Impact of user arrival rate in a time-based file downloading scenario.

cannot earn the revenues as high as the SDBR scheme does. However, the QSSF and QEDF schemes result in the lowest revenues since their priorities is to reallocate reclaimed bandwidth to QoS-unsatisfied users first. Hence, users not receiving the additional bandwidth may not satisfy the services.

Referring to Fig. 7b, when the λ_μ value is too small, there is a slight difference among the values of the average user payment. However, with increasing λ_μ , the difference among these values is obviously shown. For example, the average payment values under the SDBR and QSSF-and-QEDF schemes for λ_μ of 0.05 all are at roughly 0.63, while those under the DBR and static schemes all are roughly at 0.68. After this, the SDBR scheme results in the lowest average payments of the five schemes, which is around 0.015 and around 0.106 lower than those under the QSSF-and-QEDF and static-and-DBR schemes for λ_μ of 5, respectively. Although increasing λ_μ makes every scheme lose the user satisfaction requirements, we can see that the average user payments under the QSSF and QEDF schemes are almost identical to those under the SDBR schemes since they earn the lowest revenue and have a limited number of users, while the SDBR scheme accommodates so many users and results in the highest revenue. Hence, it is possible that the average payment values under these three schemes are similar. For the static and DBR schemes, they have the same N_{\max} users as the QSSF and QEDF schemes have, but gain the greater revenues. Therefore, their average payment values are always higher than those under the QSSF and QEDF schemes.

In Fig. 7c, when the λ_μ value is too small, there are a few users denied to acquire services. The blocking probability values under the five methods are all around at 0.0001. However, when the λ_μ value gets higher, the SDBR scheme is able to accept more new sessions than the others. Therefore, it results in the lowest blocking probability

values, from 0.0087 at λ_μ of 0.02 to 0.9411 at λ_μ of 2. However, when $\lambda_\mu \geq 5$, the blocking probability values under the five values are all approximately at 1, because they all cannot accept more users to acquire the services.

7.1.2. Effect of WiFi bandwidth

Referring to Fig. 8, in general, increasing b_2 values results in the growth of the overall revenues and the user payments, while the blocking probability values decrease with the increment of b_2 values.

As shown in Fig. 8a, when $b_2 < 30$ Mbps, the SDBR scheme tends to disable the bandwidth reclaim policy (or let $\alpha_i = 0$) and hence its performance is similar to the static and DBR schemes. For example, the overall revenue under the DBR, static, and SDBR schemes for b_2 of 20 Mbps is 55 190.1, 53 433.4, and 50 035.5, respectively, while those under the QSSF and QEDF schemes are all at approximately 41 000. The QSSF and QEDF schemes results in the lower revenues since most of users do not satisfy services. However, when b_2 is higher than 30 Mbps, the SDBR scheme accommodates more additional sessions than the others do, so the higher number of sessions results in generating higher revenues. This is why the SDBR scheme achieves the highest revenues of the five methods.

In Fig. 8b, not only does increasing b_2 values result in the growth in the mobile operator overall revenues but also that in the average user payments. The DBR scheme results in the highest average payments of the five schemes in all cases of b_2 values since many users satisfy the given services even though the total number of users is not as high as given under the SDBR scheme. The QSSF and QEDF schemes results in the lowest average payment of the five methods, which is only 5.35% lower than those under the SDBR scheme at the final value of b_2 . However, it does not indicate that their performance is better

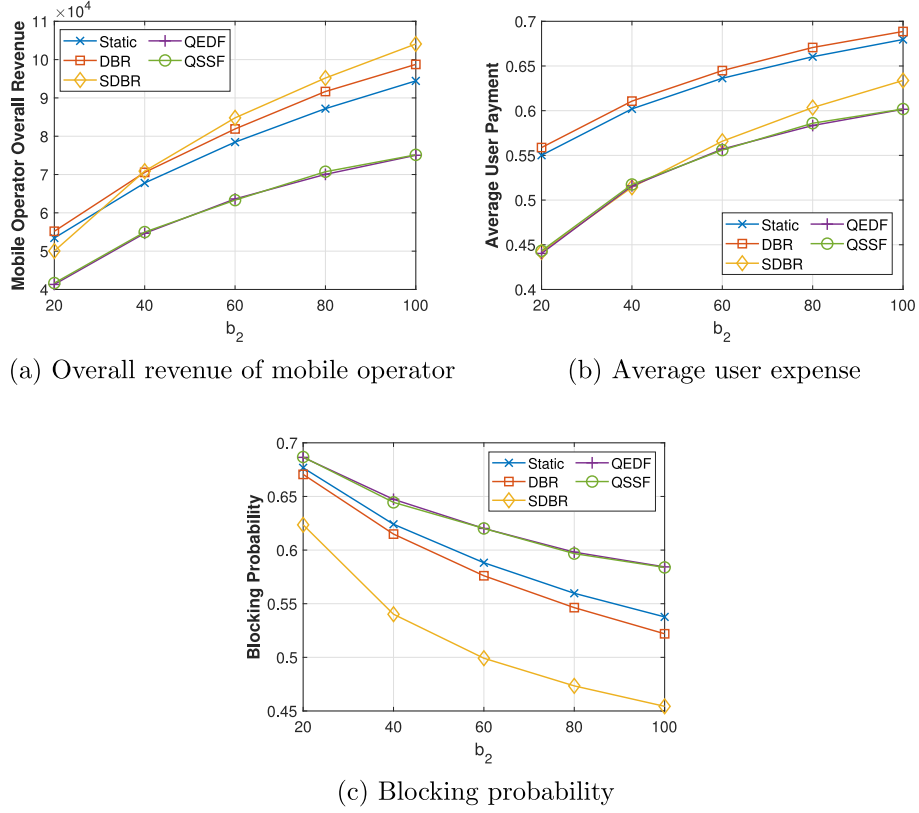


Fig. 8. Impact of WiFi bandwidth in a time-based file downloading scenario.

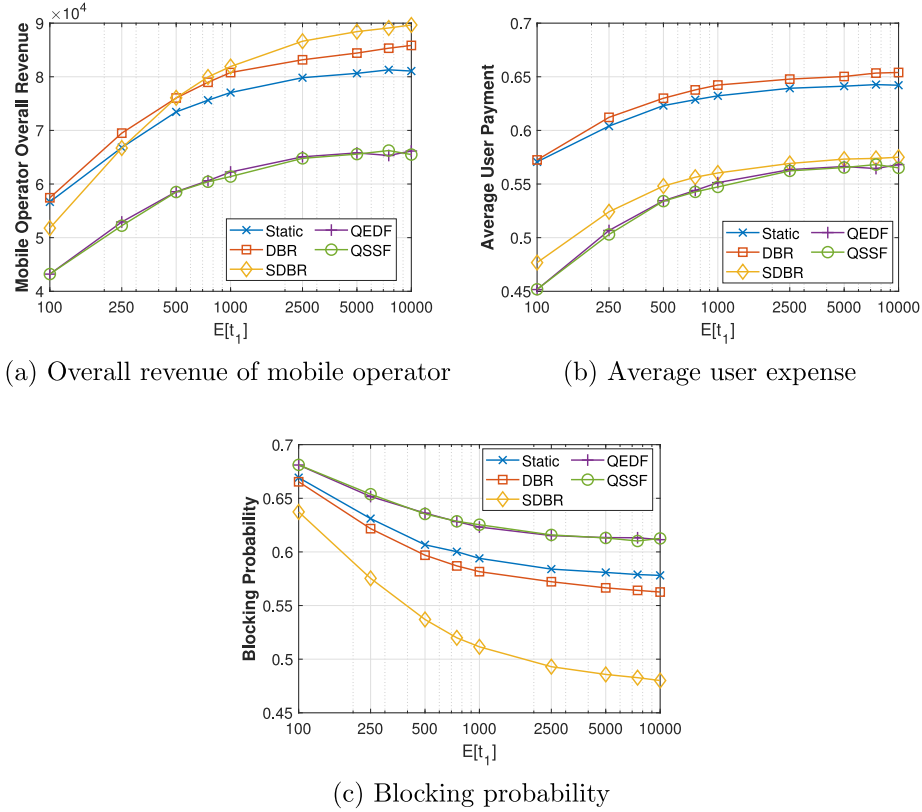


Fig. 9. Impact of average WiFi connection time in a time-based file downloading scenario.

than the SDBR scheme since most of their users do not satisfy the services. Accordingly, they cannot earn the desirable incomes and also no additional users acquire the services. Undoubtedly, their average payments are not as high as the others.

Fig. 8c shows that with the increment of b_2 values, there is a rapid decrease in the rates of the blocking probability under the five methods. Increasing b_2 values results in higher values of the user satisfaction due to the reduced transmission latency. Thus, for all higher values of b_2 , the SDBR scheme accommodates the greater number of users than N_{\max} while the others still can accept only N_{\max} users to acquire services. Hence, the SDBR scheme results in the lowest blocking probability values of the five methods, which markedly drops from 0.6235 to 0.4543, while those under the QSSF-and-QEDF, static, and DBR schemes end at 0.5840, 0.5377, and 0.5219, respectively.

7.1.3. Effect of average WiFi connection time

The results in Fig. 9 show that overall, increasing the average WiFi connection time results in the growths of the overall revenues and the average payments for all five schemes, while the blocking probability values of the five schemes decrease.

In Fig. 9a, when $E[t_1]$ is short, the reclaim policy of every scheme tends to be disabled since most data traffic is performed via the cellular path, so the SDBR scheme does not outperform the others. For example, the overall revenue under the DBR, static, and SDBR schemes for $E[t_1]$ of 100 s is approximately 57 438.5, 56 667.1, and 51 780.1, respectively. However, when $E[t_1]$ gets longer, the data offloading process via a WiFi path is more frequently performed and the reclaim policy of the SDBR scheme accommodates more additional sessions than those of the other four schemes. Therefore, longer WiFi connection time is to assist the SDBR scheme to achieve the highest overall incomes of the five methods. For example, the overall income under the SDBR scheme for $E[t_1]$ of 5000 s is at about 88 416.6, while that under the DBR scheme are 84 420.8. Meanwhile, the QSSF and QEDF schemes are not successful in earning incomes since their services are not satisfied by many users not receiving the additional cellular bandwidths.

As shown in Fig. 9b, increasing $E[t_1]$ values results in the growth of the average payments. Although the average payments under the QSSF-and-QEDF schemes are always smaller than those under the other four schemes, it cannot be claimed that they outperform the others. Their price per person is the cheapest since there are a few users satisfying the given services, while the SDBR scheme results in the highest revenues and simultaneously provides almost the same price as the QSSF-and-QEDF schemes do. This indicates that there are more users satisfying the services provided the SDBR schemes than those provided by the others. Referring to Fig. 9c, when the WiFi network is more frequently connected, the SDBR schemes can use the reclaim policy to accept more additional sessions than the other schemes. Therefore, higher values of $E[t_1]$ encourage the SDBR scheme to accommodate the greatest number of users. By the virtue of accepting more users that the SDBR scheme supports, it achieves the lowest blocking probability values of the five methods for all values of $E[t_1]$, which decrease markedly from 0.6373 to 0.48, while those under the others are just over 55%.

7.1.4. Effect on b

Because increasing λ_μ results in the marked growth in the blocking probability value that lead to a loss of QoS level. To compensate for the revenue detriment under a load of users, the information regarding the change of the values of the mobile overall revenue and the average user expense under the effect of varying b values is provided to discover the method to boosting the revenue.

Referring to Fig. 10, the mobile operator overall revenue under both SDBR and DBR methods increases in accordance with a growth of the shape parameter b , whereas the average user payment under those methods drops for higher values of the user arrival rate. As shown in Fig. 10a, since b is a crucial factor to increase the overall revenue and the proposed SDBR scheme is capable of accommodating more users

than the existing DBR scheme, the SDBR always results in a higher revenue than that under the DBR scheme for all values of b . Owing to the same reason, the results in Fig. 10b show that SDBR always costs the lower average payment to users than that under the DBR scheme in all cases of b values.

7.2. Volume-based file downloading

This volume-based file downloading scenario is modeled for delay-tolerant applications in which users can suffer a lengthy delay of a downloading time t_l such as performing software updates. A decline of the user satisfaction against t_l in this scenario is not as rapid as that in the delay-sensitive scenario. Referring Table 3, almost all basic parameter settings in this scenario are as same as the previous scenario with the following exceptions. Firstly, the total cellular bandwidth of the BS W and the guaranteed cellular bandwidth w_g are respectively reduced to 200 Mbps and 20 Mbps. Secondly, for each user session, the file size was uniformly generated via a Pareto distribution with a shape parameter of 1.8 and a mean file size of 3 GB. To make every user suffer a lengthy delay, a deadline time t_d is adjusted to 600 s.

7.2.1. Effect of user arrival rate

In Fig. 11, overall, higher values of λ_μ result in the growth in the values of the mobile revenues and average payments. However, the blocking probability values under all five methods decrease for all values of λ_μ . Interestingly, the SDBR scheme significantly outperforms the others in all cases of λ_μ values.

In Fig. 11a, since all mobile users under the five methods tolerate lengthy download times, the reclaiming bandwidth is vital. When λ_μ is low, the performance under the SDBR scheme is similar to that under the DBR scheme. For example, all the overall revenues under the SDBR and DBR schemes for λ_μ of 0.02 are approximately 21 800, while the figure under the static scheme is just under 20 300. It is seen that the QSSF and QEDF schemes suffer the delay and so many of their users do not satisfy the services. Hence, they result in the lowest overall revenue. However, with the increment of λ_μ values, the SDBR scheme shows its effectiveness to handle the huge number of users via the reclaim policy that the revenue is generated with more users than the other schemes. Hence, it achieves the highest revenues of the five methods when $\lambda_\mu \geq 0.03$, while those under the other schemes are all just under 22 000.

As shown in Fig. 11b, when $\lambda_\mu \geq 0.02$, there are not so many users in the system, so the performances of maintaining the QoS levels under all five methods are similar and such five methods result in the almost same expenses accordingly. However, owing to the reclaim policy that the SDBR scheme accommodates more users to generate higher incomes than the others do, with increasing λ_μ , the user payments under the SDBR scheme decrease rapidly since there are more users than the others and it results in the cheapest offloading price for $\lambda_\mu > 0.02$, while the offloading prices under the others are just over 0.35.

Fig. 11c proves that under the situation of a load of users requesting and acquiring services, the SDBR scheme is the most effective in accommodating the greater number of users for the services, while the others suffer lengthy delay and they cannot use their reclaim policy to accept additional sessions. Therefore, when $\lambda_\mu \leq 1$, the blocking probability values under the SDBR scheme always around 11.5% less than those under the others before increasing to approximately 0.99 when $\lambda_\mu > 1$. Meanwhile, the blocking probability values under the other four schemes are almost identical for all values of λ_μ .

7.2.2. Effect of WiFi bandwidth

Referring to Fig. 12, with increasing b_2 values, the five schemes result in increasing the overall revenues and the average user payments and reducing the blocking probability values. In addition, the SDBR scheme achieves the most outstanding results of the five methods.

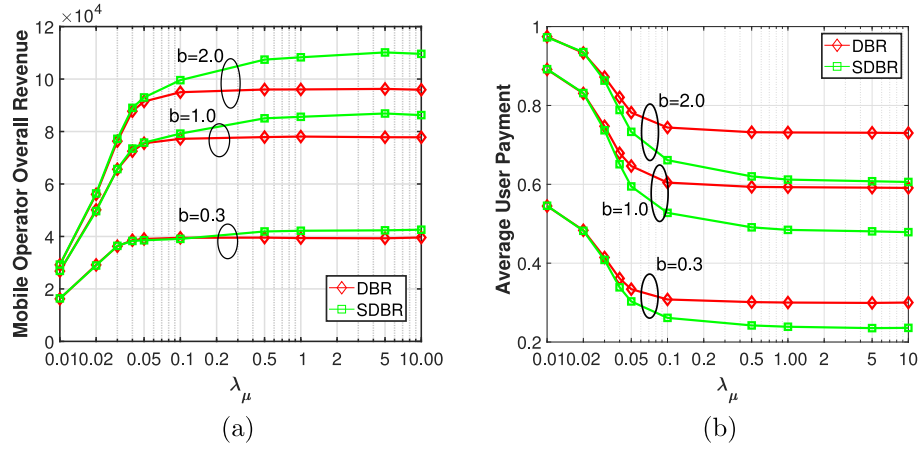
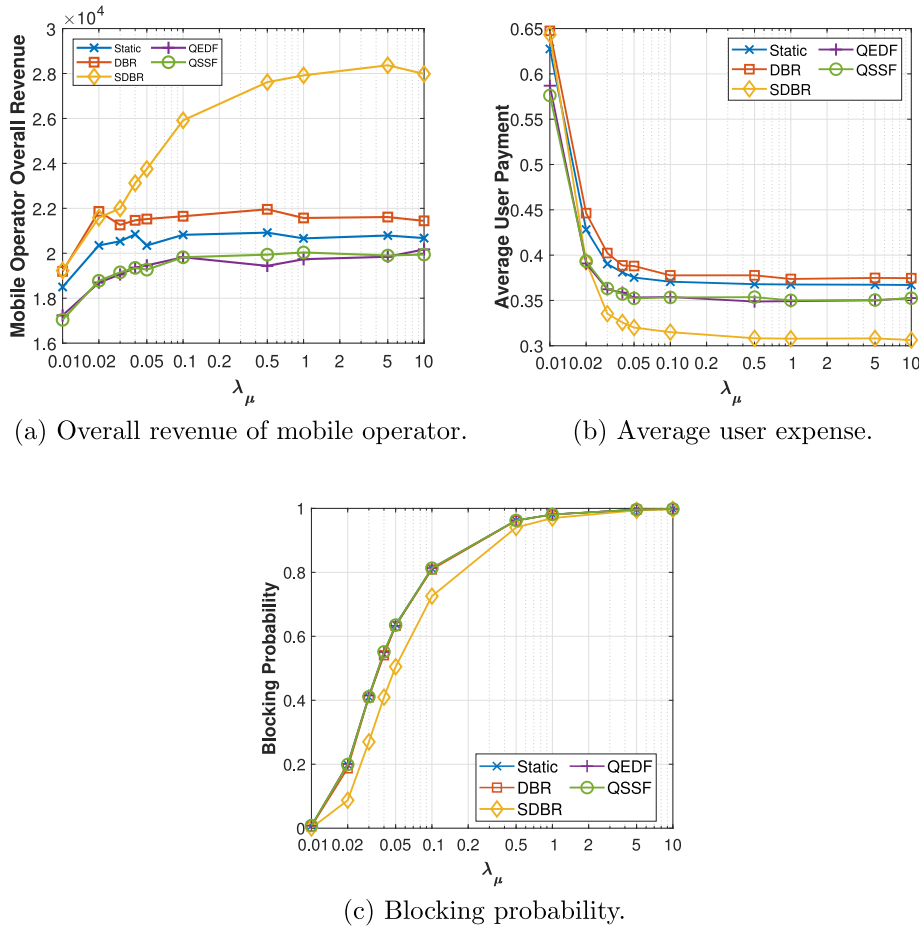
Fig. 10. Different b in a time-based file downloading scenario.

Fig. 11. Impact of user arrival rate in a volume-based file downloading scenario.

In Fig. 12a, since the SDBR scheme tends to disable its reclaim policy when b_2 is low, most users suffer a high transmission latency for downloading a large file size. Therefore, it cannot outperform the others in term of earning revenues. For example, when $b_2 \leq 30$ Mbps, the overall revenue under the SDBR scheme starts at roughly 1476.15 while those under the others are all at roughly 3600. However, with increasing b_2 values, the SDBR scheme can take advantage in its policy to accept more incoming sessions that can generate higher revenues than the other schemes do. Hence, the overall incomes under the SDBR scheme for b_2 of 80 and 100 Mbps are 25% and 28% greater than those under the others, respectively.

As shown in Fig. 12b, increasing b_2 values is to assist mobile users to complete their processes early, so the SDBR scheme accommodates more additional users and simultaneously generates higher incomes than the others owing to the reclaim policy. Hence, it can charge users with the lowest prices of the five methods. For example, the average payment values under the SDBR scheme is always around 18% greater than those under the others in all cases of b_2 values.

Referring to 12c, when users complete the offloading processes early, the mobile operators can accommodate more new arriving sessions. Hence, the lowest blocking probability value of the five methods is achieved by the SDBR scheme since it accommodates the significantly

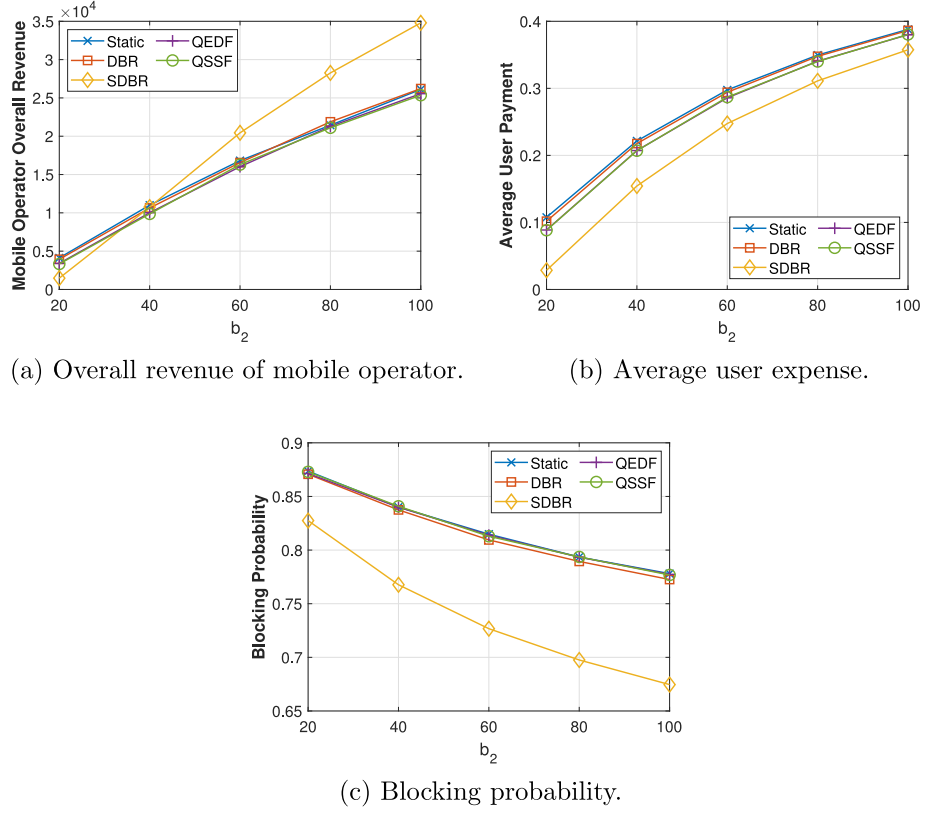


Fig. 12. Impact of WiFi bandwidth in a volume-based file downloading scenario.

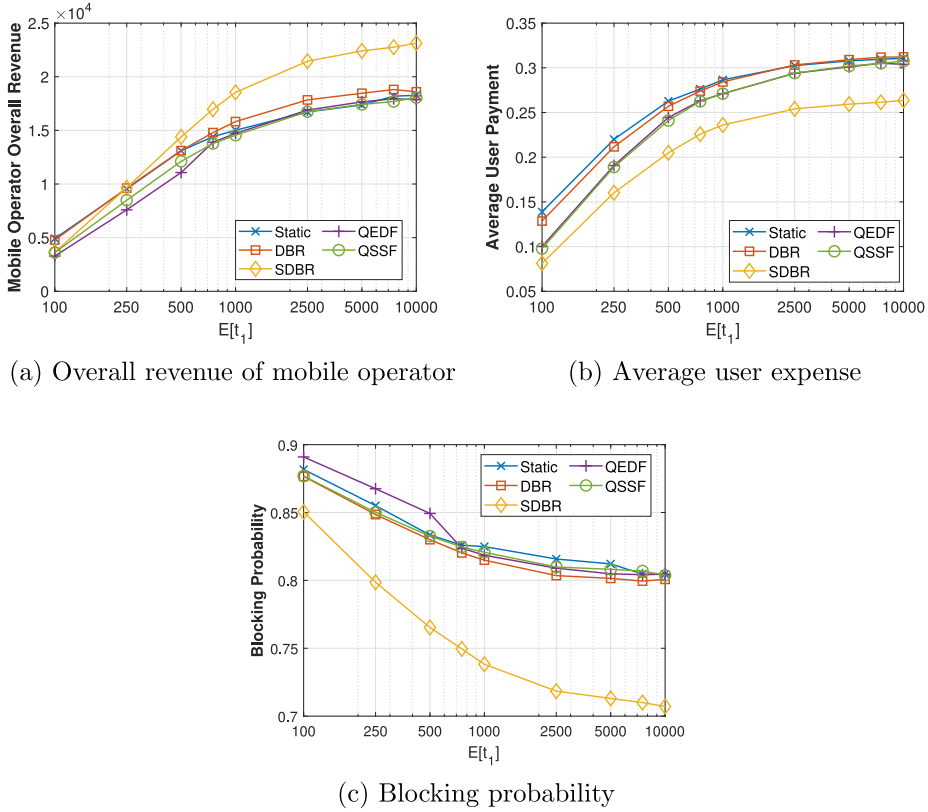


Fig. 13. Impact of average WiFi connection time in a volume-based file downloading scenario.

greater number of users than the other four schemes do. For example, the blocking probability values under the SDBR scheme are always around 13% lower than those under the other four schemes.

7.2.3. Effect of average WiFi connection time

The results in Fig. 13 show that with expanding the WiFi connection time, the five schemes results in a growth in the overall revenues and the average payments. By contrast, an increase of $E[t_1]$ results in a decline in the blocking probability values since the WiFi network is more frequently connected in accordance with the increment of $E[t_1]$ values.

As shown in Fig. 13a, when $E[t_1] < 250$, mobile users under the five schemes suffer a huge delay tolerance since most data traffic is performed via the cellular path. Therefore, the SDBR scheme tends to disable its reclaim policy and hence its effectiveness is similar to the other schemes. For example, all the overall revenues under the five methods for $E[t_1]$ of 100 s are just under 5000 and the revenue under the SDBR scheme starts at around 3600. However, when the data traffic is mostly performed via the WiFi path, the reclaim policy is more frequently used to assist the SDBR scheme to earn higher revenues and to accommodate more users. As a result, the SDBR scheme results in the highest overall revenues when $E[t_1] \geq 500$ s, which increase from 4393.8 for $E[t_1]$ of 500 s to 23 136.7 for $E[t_1]$ of 10 000 s.

Referring to Fig. 13b, increasing $E[t_1]$ values encourages mobile subscribers to use WiFi bandwidths more frequently. Therefore, the SDBR scheme takes advantage in its policy to accommodate more users by allocating cellular resource to hungry-bandwidth users, while the other schemes need to deal with the lengthy transmission latency. As a result, the SDBR scheme results in the lowest average user payments for all values of $E[t_1]$, which is around 15% lower than those under the others.

In Fig. 13c, as increasing $E[t_1]$ values, the WiFi resource is increasingly maximized. Thus, a higher number of users are encouraged to join the service early because users in the services finish the activities before the deadline. By virtue of supporting a larger number of users than the theoretical limit N_{\max} , the SDBR scheme is successful in minimizing the blocking probability and it achieves the lowest blocking probability values of the five methods, which are around 10% lower than those under the other four schemes.

7.2.4. Effect of shape parameter b

Fig. 14 shows that the results under the SDBR scheme are still always by far effective than those under the DBR scheme. Higher values of b result in the performance enhancement of both SDBR and DBR schemes. Referring to Fig. 14a, increasing b results in the overall revenue growth and hence the overall revenue of the mobile operator under the SDBR scheme is always higher than that under the DBR scheme for all values of b . At $b = 2.0$, the overall revenue under the SDBR scheme is markedly higher than that under the DBR scheme. As shown in Fig. 14b, since every UE in this scenario suffers a lengthy delay throughout the offloading session, the average user payment under the SDBR scheme is almost as same as that under the DBR scheme at $b = 0.3$. However, the average payment increases with increasing b and the SDBR scheme always costs a lower expense than the DBR scheme. This evaluation based on the user satisfaction explicitly proves that for experimenting with delay-tolerant applications, the value of b should be at least 1.0 so that the mobile operator obtains the worth revenue.

7.3. Time-complexity comparison

The numerical examples was conducted by a 3.2 GHz Intel core-i7 with 16 GB of memory. Table 4 compares the time complexity among different algorithms. Its details can be described as follows.

According to [6,13], the static and SDBR schemes execute the bandwidth distribution step for n times, which is the number of users in

Table 4

Time complexity comparison.

| Algorithm | Time complexity |
|---------------|--------------------|
| Static | $\mathcal{O}(n)$ |
| DBR | $\mathcal{O}(n)$ |
| QSSF | $\mathcal{O}(n^2)$ |
| QEDF | $\mathcal{O}(n^2)$ |
| Proposed SDBR | $\mathcal{O}(n^2)$ |

the U_{BS} . Hence the time complexity of both DBR and static algorithms is $\mathcal{O}(n)$. Regarding the QSSF and QEDF algorithms [14], their time complexity is $\mathcal{O}(n^2)$ since they need to distribute bandwidth to users in the U_{BS} and need to find users with the earliest deadline and users with the smallest download volume. As we mentioned in Section 5.3, the time complexity of the SDBR algorithm is $\mathcal{O}(n^2)$.

It seems that the DBR and static algorithms are equipped with the better run time than the others. However, when considering improving the user satisfaction of the provided services, the SDBR scheme creates a win-win situation in which the mobile operators earn the desirable revenues commensurate with their QoS requirements, while their mobile users pay the offloading fees with an inexpensive price. Compared to the others, although the DBR and static schemes have the best run time of the five methods and also gain desirable incomes from their users, they cannot make users satisfy services as much as the SDBR scheme does. This is why such methods result in the higher rate of blocking probability under various factors. About QSSF and QEDF algorithms, most of users do not satisfy given services even their run time is identical to that of the SDBR scheme since such methods are effective for only users classified to their reallocation policies. Therefore, the operator cannot attain its goals in terms of earning incomes and reducing blocking probability.

In conclusion, the SDBR scheme is much worthier than the others in the aspect of the user satisfaction improvement even though it cannot result in the best run time.

8. Discussion

Regarding the theoretical analysis, the SDBR scheme notably outperforms the baseline scheme for all considering factors. Compared to the lower-bound scheme, the WiFi resource is crucial for the dynamic bandwidth reallocation of the SDBR scheme. We can see that when increasing b_2 and P_1 , its blocking probability values tend to be similar to those under the lower-bound schemes. However, with the increment of α , the cellular bandwidth from WiFi users is heavily reclaimed, so the blocking probability value under the SDBR scheme increases as the α value increases.

About the simulation results, since there is a huge transmission latency in the scenario of volume-based file downloading, it is clearly seen that the overall revenues and user payments in the time-based file downloading scenario are always higher than those in the volume-based file downloading scenario. Most importantly, the WiFi resource is really important for the SDBR schemes. It can outperform the other four schemes since higher values of b_2 and $E[t_1]$ assist the SDBR scheme to enable the reclaim policy to accommodate more additional users that can generate higher revenue than the others. Hence, although there are loads of users, the SDBR scheme still can outperform the others for all output metrics as long as the reclaim policy is enabled.

9. Conclusion

This paper proposes the SDBR scheme to maximize the user satisfaction by reducing the service blocking probability and the individual payment for the service, together with increasing the overall revenue on the mobile operator, which is a significant indicator of the service performance. The proposed SDBR scheme encourages the base

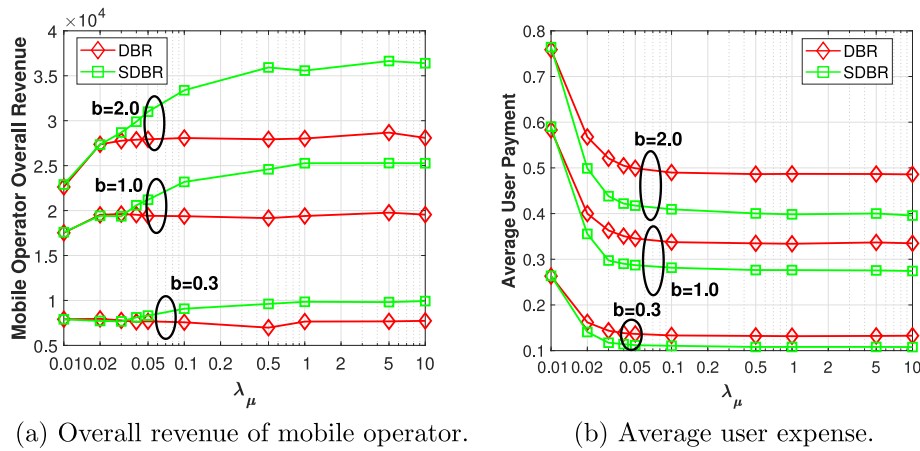


Fig. 14. Impact of shape parameter b in a volume-based file downloading scenario.

station to accommodate a greater number of download sessions by considering the WiFi connection status of the existing connected users in accordance with a price-based satisfaction function to balance the download delay and the serving cellular bandwidth. By doing so, the mobile operator simultaneously maximizes its overall revenue while still satisfying the QoS requirements of the users.

The simulation results have shown that compared to the existing DBR schemes in the time-based file downloading scenario, the SDBR scheme results in the growth of 11.36% and 17.02% in the overall revenues and the average user expenses, respectively, and also results in a drop of 30% in the rates of the blocking probability. In the volume-based file downloading scenario, the SDBR scheme increases the overall revenue and the average user expense, and reduces blocking probability by 47.5%, 13.33%, and 12.5%, respectively. In other words, all five schemes are successful in both maximizing the satisfaction function and minimizing the blocking probability. However, the main reason why the SDBR scheme significantly outperforms the other four schemes even though the BS suffers a massive load of clients is because of the virtue of supporting the greater number of users than the theoretical N_{\max} , which results in the higher revenues than the other four schemes. Thus, the mobile operator can apply the SDBR scheme in providing a reliable services for its subscribers.

In future works, we plan to solve the closed form of the optimization problem of the SDBR. Also, we will consider more kinds of services for different levels of QoS requirements.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of S.-I. Sou was sponsored in part by Ministry of Science and Technology (MOST), Taiwan, under the contract number MOST 105-2628-E-006-007-MY3 and MOST 108-2628-E-006 -006 -MY3.

References

- [1] Cisco, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update (2018–2023), White paper(2020), San Jose, CA, USA, 2018, <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html> (accessed 29 April 2020).
- [2] F. Rebecchi, M.D.D. Amorim, V. Conan, A. Passarella, R. Bruno, M. Conti, Data offloading techniques in cellular networks: a survey, *IEEE Commun. Surv. Tutor.* 17 (2) (2015) 580–603.
- [3] R. Maallawi, N. Agoulmine, B. Radier, T.B. Meriem, A comprehensive survey on offload techniques and management in wireless access and core networks, *IEEE Commun. Surv. Tutor.* 17 (3) (2015) 1582–1604.
- [4] S.-I. Sou, Y.-T. Peng, Performance modeling for multipath mobile data offloading in cellular/wifi networks, *IEEE Trans. Commun.* 65 (9) (2017) 3863–3875.
- [5] Y.-T. Peng, S.-I. Sou, M.-H. Tsai, C.-S. Lin, Multipath mobile data offloading of deadline assurance with policy and charging control in cellular/wifi networks, *Comput. Netw.* 129 (2017) 17–27.
- [6] C.-S. Lin, S.-I. Sou, Deadline assurance through dynamic bandwidth reallocation in multipath mobile data offloading, in: *IEEE Wireless Communications and Networking Conference (WCNC)*, April, 2018, IEEE, 2018.
- [7] A. Rahmati, C. Shepard, C.-C. Tossell, L. Zhong, P. Kortum, A. Nicoara, J. Singh, Seamless TCP migration on smartphones without network support, *IEEE Trans. Mob. Comput.* 13 (3) (2014) 678–692.
- [8] G. Gao, M. Xiao, J. Wu, K. Han, L. Huang, Z. Zhao, Opportunistic mobile data offloading with deadline constraints, *IEEE Trans. Parallel Distrib. Syst.* 28 (12) (2017) 3584–3599.
- [9] X. Zhou, W. Gao, G. Cao, S. Hua, An incentive framework for cellular traffic offloading, *IEEE Trans. Mob. Comput.* 13 (3) (2014) 541–555.
- [10] J. Xu, S. Wu, L. Xu, N. Zhang, Q. Zhang, User satisfaction-aware WiFi offloading in heterogeneous networks, in: *IEEE VTC-Fall*, Toronto, Canada, September 24–27 (2017).
- [11] K. Wang, F.C.M. Lau, L. Chen, R. Schober, Pricing mobile data offloading: a distributed market framework, *IEEE Trans. Wireless Commun.* 15 (2) (2016) 913–927.
- [12] V. Miliotis, L. Alonso, C. Verikoukis, Weighted proportional and pricing based resource allocation for uplink offloading using IP flow mobility, *Ad Hoc Netw.* 49 (2016) 17–28.
- [13] Y.B. Lin, Y.J. Shih, P.W. Chao, Design and implementation of LTE rrm with switched LWA policies, *IEEE Trans. Veh. Technol.* 67 (2) (2018) 1053–1062.
- [14] C.-S. Lin, S.-I. Sou, Qos-aware dynamic bandwidth reallocation with deadline assurance for multipath data offloading, *Comput. Netw.* 153 (2019) 103–112.
- [15] E. Fakhfakh, S. Hamouda, Incentive reward for efficient WiFi offloading using Q-learning approach, in: *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Valencia, 2017, pp. 1114–1119.
- [16] L. Pu, X. Chen, J. Xu, X. Fu, D2d fogging: An energy-efficient and incentive-aware task offloading framework via network-assisted d2d collaboration, *IEEE J. Sel. Areas Commun.* 34 (12) (2016) 3887–3901.
- [17] M. LiWang, S. Dai, Z. Gao, X. Du, M. Guizani, H. Dai, A computation offloading incentive mechanism with delay and cost constraints under 5g satellite-ground iov architecture, *IEEE Wirel. Commun.* 26 (4) (2019) 124–132.
- [18] H. Ko, J. Lee, S. Pack, Performance optimization of delayed wifi offloading in heterogeneous networks, *IEEE Trans. Veh. Technol.* 66 (10) (2017) 9436–9447.
- [19] M.H. Cheung, J. Huang, DAWN: delay-aware wi-fi offloading and network selection, *IEEE J. Sel. Areas Commun.* 33 (6) (2015) 1214–1223.
- [20] H. Xu, X. Wen, Z. Lu, Z. Hu, W. Jing, K. Chen, Performance analysis of delayed mobile data offloading with multi-level priority, in: *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, (2016) pp. 1–6.
- [21] Z. Hu, Z. Lu, X. Wen, Q. Li, Stochastic-geometry-based performance analysis of delayed mobile data offloading with mobility prediction in dense IEEE 802.11 networks, *IEEE Access* 5 (2017) 23060–23068.
- [22] F. Mehmeti, T. Spyropoulos, Performance. modeling, analysis, Performance modeling analysis and optimization of delayed mobile data offloading for mobile users, *IEEE/ACM Trans. Netw.* 25 (1) (2017) 550–564.

- [23] H. Shah-Mansouri, V.W.S. Wong, J. Huang, An incentive framework for mobile data offloading market under price competition, *IEEE Trans. Mob. Comput.* 16 (11) (2017) 2983–2999.
- [24] X. Song, L. Qin, H. Qi, S. Li, H. Qian, L. Dong, et al., Incentive framework for mobile data offloading market under qoe-aware users, *IET Commun.* 14 (13) (2020) 2151–2161.
- [25] S. Pal, S.K. Das, M. Chatterjee, User-satisfaction based differentiated services for wireless data networks, in: *IEEE International Conference on Communications*, 2005. ICC 2005. Vol. 2, Seoul (2005) pp. 1174–1178.
- [26] R.N. Bolton, K.N. Lemon, A dynamic model of customers' usage of services: usage as an antecedent and consequence, *J. Mark. Res.* 36 (1999) 171–186.
- [27] R. Leszinski, M.V. Marn, Setting value, not price, *McKinsey Quart.* 1 (1997) 99–115.
- [28] Y. Cao, T.S. Gruca, B.R. Klemz, Internet pricing, price satisfaction, and customer satisfaction, *Int. J. Electron. Commer.* 8 (2) (2003) 31–50.
- [29] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, *Introduction to algorithms*, in: *Community Resource for Archiving Wireless Data At Dartmouth (CRAWDAD)*, MIT press, 2009.
- [30] J. Shi, C. Qiao, D. Koutsonikolas, G. Challen, CRAWDAD Dataset buffalo/phonelab-wifi (v. 2016-03-09), 2016, [Online]. Available: <https://crawdad.org/buffalo/phonelab-wifi/20160309> (accessed 29 April 2020).



Nuntanut Bhooanus received the BEng degree of Computer Engineering from King Mongkut's University of Technology Thonburi, Thailand, in 2011 and the MS degree of Computer Science from Thammasat University, Thailand, in 2015. Currently, He is a Ph.D. candidate in the Institute of Computer and Communication Engineering at National Chen Kung University, Taiwan.



Sok-Ian Sou is a professor in the Institute of Computer and Communication Engineering and the Department of Electrical Engineering, National Cheng Kung University (NCKU), Tainan, Taiwan. She received the BS, the MS and the PhD degrees in Computer Science and Information Engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1997, 2004, and 2008, respectively. She was a visiting scholar at Carnegie Mellon University (CMU) during July–August 2009. She was a recipient of the Investigative Research Award from the Pan Wen Yuan Foundation in 2009 and the Young Researcher in Service Science Award from the Sayling Wen Cultural & Educational Foundation in 2012. She also received Outstanding Teacher Award in 2017. Her current research interests include Internet of things, mobile networks, and wireless signal analysis. She is the co-author of the book entitled "Charging for Mobile All-IP Telecommunications" (with Yi-Bing Lin; published by Wiley, 2008).



Kai-Chun Cheng received the BS degree from the Department of Communications Engineering, Feng Chia University, Taiwan in 2017, the MS degree in the Institute of Computer and Communication Engineering, National Cheng Kung University, Taiwan, in 2019.