A User Centered Multi-Objective Handoff Scheme for Hybrid 5G Environments

Li Qiang, Jie Li, Corinne Touati

To cite this version:

Li Qiang, Jie Li, Corinne Touati. A User Centered Multi-Objective Handoff Scheme for Hybrid 5G Environments. IEEE Transactions on Emerging Topics in Computing, Institute of Electrical and Electronics Engineers, 2016. hal-01384756

HAL Id: hal-01384756
https://hal.inria.fr/hal-01384756

Submitted on 20 Oct 2016

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2016.2551042, IEEE Transactions on Emerging Topics in Computing

A User Centered Multi-Objective Handoff Scheme for Hybrid 5G Environments

Li Qiang, *Member, IEEE*, Jie Li, *Senior Member, IEEE* and Corinne Touati, *Member, IEEE*

Abstract—In this paper, we propose a user centered handoff scheme for hybrid 5G environments. The handoff problem is formulated as a multi-objective optimization problem which maximizes the achievable data receiving rate and minimizes the block probability simultaneously. When a user needs to select a new Base Station (BS) in handoff, the user will calculate the achievable data receiving rate and estimate the block probability for each available BS based on limited local information. By taking the throughput metric into consideration, the formulated multi-objective optimization problem is then transformed into a maximization problem. We solve the transformed maximization problem to calculate the network selection result in a distributed method. The calculated network selection result is proved to be a Pareto Optimal solution of the original multi-objective optimization problem. The proposed scheme guarantees that based on limited local information, each user can select a new BS with high achievable data receiving rate and low block probability in handoff. Comprehensive experiment has been conducted. It is shown that the proposed scheme promotes the total throughput and ratio of users served significantly.

Index Terms—Hybrid 5G environments, multi-objective optimization, vertical handoff, network selection, distributed algorithm.

◆

1 INTRODUCTION

The emergence of 5G will not replace the existing technologies[1] but be more integrative and hybrid: combining with existing technologies to provide ubiquitous high-rate and seamless communication service [1]. As we move toward 5G era, environment becomes so complex that the handoff problem faces with new challenges. The data rate in 5G is expected to be roughly 1000× compared with current 4G technology [2], hence the handoff problem requires a faster processing [3]. Furthermore, as the number of Base Stations (BSs) and mobile devices dramatically increases, the centralized control may not be efficient. On the contrary, more intelligent mobile devices can play important roles in handoff. Moreover, increasingly serious data security problem reminds users[2] do not share their private information with others. Thus, it is glad to see a fast, distributed, privacy-preservation and user centered handoff scheme in hybrid 5G environments. Motivated by this, we will study the handoff problem for hybrid 5G environments in this paper.

Consider a scenario as shown in Fig.1 where 3G [4], LTE, WiMAX and 5G BSs construct a hybrid 5G environment. Users in the hybrid 5G environment do not share their private information with others. Moving in this scenario, users may need to transfer their network connections from one BS to another. This kind of transferring operation is called handoff [5]. The handoff problem refers that when a user has several available BSs in a handoff, the user needs to decide to which BS the network connection should be transferred [6]. Take a user for instance. As the user moves far away from 3G BS, the signal strength received from 3G BS gets so weak that the user has to transfer his (or her) network connection to a new BS. This user has three possible choices: LTE, 5G and WiMAX BSs [7]. He (or she) has to decide which BS should be selected. It seems that the handoff problem is very simple, the user only needs to select the best performance one. However, the user has difficulties to know the network selection behaviors of other users. If there are too many other users making the same selection, this user is possible to be blocked [8], [9]. As a result, the objectives of network selection are to select a high performance BS and avoid being blocked.

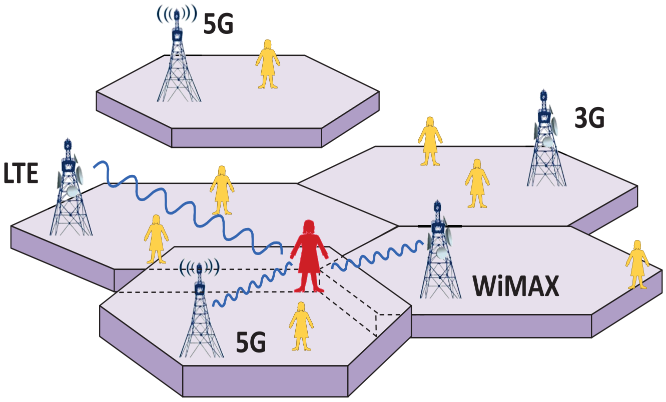


Fig. 1. Illustrative example for handoff problem in Hybrid 5G environment.

There are two kinds of approaches to solve the handoff problem in general: the network centered approach and the user centered approach. In the network centered approach,

---

- *Li Qiang is with Graduate School of Systems and Information Engineering, University of Tsukuba, Japan. E-mail: qiangli@osdp.cs.tsukuba.ac.jp*
- *Jie Li (Corresponding Author) is with Faculty of Engineering Information and Systems, University of Tsukuba, Japan. E-mail:lijie@cs.tsukuba.ac.jp*
- *Corinne Touati is with Inria and CNRS, LIG, University Grenoble Alpes, France. E-mail:corinne.touati@inria.fr*

Manuscript received Oct. 11, 2015. Major revision Dec. 8, 2015. Accepted Feb. 28, 2016.

1. The existing technologies include 3G, LTE, and so on.
2. The terms *user* and *mobile device* are interchangeable in the paper.

networks are responsible for computing and making the decisions. In the user centered approach, users will be in charge of the network selection. Considering the requirement of privacy-preservation in hybrid 5G environment, users are not suggested to send out their private information (*e.g.*, number of available networks, basic bandwidth requirement and so on) [10]. Under this limitation, networks are unable to obtain adequate information from users for the network selection. As a result, the user centered approach is more suitable for the hybrid 5G environment than the network centered approach.

In this paper, we propose a user centered multi-objective handoff scheme for hybrid 5G environment. In our proposed scheme, users are divided into two classes: *non-handoff users* and *handoff users*. Non-handoff users will stay in the connections with their current BSs. While handoff users will transfer their network connections to new BSs based on limited local information. Local information refers to the private information of the user itself, the parameters of BSs and two pieces of public information (*i.e.*, the total numbers of handoff and non-handoff users inside each available BS). When a user needs to select a new BS in a handoff, it will calculate the achievable data receiving rates of all its available BSs. Furthermore, the user also has to **infer** the network selection behaviors of other users in order to **estimate** its block probability for each available BS. By jointly considering the achievable data receiving rate and block probability, the user can select the most appropriate BS in a handoff. The main contributions of this paper are summarized as follows:

- We study the relations between two users, and define the *correlation degree*. The correlation degree could efficiently distinguish the categories of relations, and sufficiently reflect the association strength.
- We formulate the handoff problem as a multi-objective optimization problem which maximizes the achievable data receiving rate and minimizes the block probability. Then, we transform the formulated multi-objective optimization problem into an equivalent maximization problem.
- We solve the transformed maximization problem by a distributed method in polynomial time and linear space. We further prove that the solution of the transformed maximization problem is a Pareto Optimal [11] result of the original multi-objective optimization problem.

The rest of this paper is organized as follows. Section 2 presents the related works. The system description and problem formulation are given in Section 3. In Section 4, we study the estimation method of block probability. Section 5 presents our proposed handoff scheme for hybrid 5G environments. Section 6 is the performance evaluation. Section 7 concludes this paper.

2 RELATED WORK

Although a lot of works have been conducted in addressing the handoff problem, most of the these existing works could not be directly used for hybrid 5G environments. As we have explained in Introduction, the high data rate, numerous mobile devices and BSs, and security awareness of hybrid 5G environments appeal for a fast, distributed and privacy-preservation handoff scheme. In this section, we will introduce some interesting handoff schemes and discuss the experiences which should be concerned in our study.

A Quality of Service (QoS) aware handoff scheme is proposed by Yang *et al.* [12]. In the proposed scheme, the QoS metric is the received Signal to Interference and Noise Ratio (SINR) which can be used to evaluate the achievable bandwidths of networks. The proposed scheme can be operated under active mode or passive mode. In the active mode, a user will select the network which provides the maximum achievable bandwidth by itself. In the passive mode, users will periodically send their received SINRs to a Radio Network Controller (RNC) which centrally makes the network selections for users. Authors pointed out that the passive mode will result in higher latency than the active mode. This experience also suggests us to consider the user centered rather than the network centered handoff scheme from another aspect.

In our previous work [13], we proposed a Software-Defined Network (SDN) based vertical handoff scheme. In the proposed scheme, users append their private information to the handoff request frames and send these frames to a SDN controller. The SDN controller formulates the handoff problem as a 0-1 integer programming problem and calculates the network selection results. As the number of mobile devices and BSs dramatically increases in hybrid 5G environment, the centralized control of a SDN controller is no longer feasible. If many SDN controllers are deployed, the cooperation between SDN controllers will cause a lot of overhead. Furthermore, sending the private information to the SDN controller is not conducive to privacy-preservation.

In order to improve the Quality of Experience (QoE), a Multiplicative Utility based Automatic Handoff scheme is proposed by Nguyen-Vuong Q.t. *et al.* [14]. In the proposed multiplicative scheme, the handoff problem is formulated as a Multiple Attribute Decision Making (MADM) problem. After calculating the multi-criteria utility function value for each available network, a user selects the highest scoring network as the new network. Since the network selection behaviors of users have influence on each other, a user should not only consider the network attributes but also needs to consider the network selection behaviors of other users during the handoff.

Chao *et al.* [15] proposed a two-step handoff scheme. The first step is pre-decision progress, in which a filtering function is used to evaluate the performance of networks. If no network can pass the pre-decision, a user will stay in the connection with its current network. If there is only one network passing the pre-decision, a user will handoff to the sole network. If there are several networks passing the pre-decision and a user has insufficient power, the user will randomly handoff to a network. If there are several networks passing the pre-decision and a user has sufficient power, the user will execute the second step. In the second step, the handoff scheme is formulated as an MADM problem and the highest scoring network device will be selected. The complex procedure of two-step handoff scheme dissatisfies the fast decision requirement of handoff in hybrid 5G environment.

3 SYSTEM DESCRIPTION AND PROBLEM FORMULATION

In this section, we formulate the handoff problem for hybrid 5G environments. Specifically, we consider a hybrid 5G environment which consists of n BSs. Let $\mathcal{B}$ be the set of BSs, $\mathcal{B} = \{b_1, b_2, \cdots, b_n\}$. These BSs support different wireless technologies. With the support of the Media-Independent Handover (MIH) standard [16], we can focus on the handoff problem from the perspective of algorithm without caring about the differences between communication technologies. Denote the frequency band of BS b_i ($b_i \in \mathcal{B}, i = 1, 2, \cdots, n$) as ω_i in MHz. b_i equally allocates its frequency band among serving users. In order to guarantee the quality of service of each user, b_i will serve at most η_i users at the same time.

Consider that there are m users. Let $\mathcal{U}$ be the set of users, $\mathcal{U} = \{u_1, u_2, \cdots, u_m\}$. If user u_j ($u_j \in \mathcal{U}, j = 1, 2, \cdots, m$) is inside the coverage area of a BS, this BS is called an **available BS** of u_j. An **adjacency matrix** $\delta(t)$ is used to reflect the available relationship between BSs and users at time t as follows. By introducing the time-slotted idea [17], a continuous period of time is divided into discrete time samples. In the rest of paper, time t is referred to the t th time slot. The system status in a time slot is assumed to be stable.

$$\delta(t) = \begin{array}{c} \\ \boldsymbol{b_1} \\ \boldsymbol{b_2} \\ \vdots \\ \boldsymbol{b_n} \end{array} \begin{array}{c} \begin{array}{cccc} \boldsymbol{u_1} & \boldsymbol{u_2} & \cdots & \boldsymbol{u_m} \end{array} \\ \begin{bmatrix} \delta_{11}(t) & \delta_{12}(t) & \dots & \delta_{1m}(t) \\ \delta_{21}(t) & \delta_{22}(t) & \dots & \delta_{2m}(t) \\ \vdots & & \ddots & \vdots \\ \delta_{n1}(t) & \delta_{n2}(t) & \dots & \delta_{nm}(t) \end{bmatrix} \end{array},$$

where

$$\delta_{ij}(t) = \begin{cases} 1, & \text{BS } b_i \text{ is available to user } u_j \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

For BS b_i, the number of users inside its coverage area is $\sum_{j=1}^{m} \delta_{ij}(t)$ which should satisfy the following constraint.

$$0 \leq \sum_{j=1}^{m} \delta_{ij}(t) \leq m. \tag{2}$$

For user u_j, the number of available BSs is $\sum_{i=1}^{n} \delta_{ij}(t)$. In hybrid 5G environment, u_j may have several available BSs. That is the value of $\sum_{i=1}^{n} \delta_{ij}(t)$ should satisfy the following constraint.

$$0 \leq \sum_{i=1}^{n} \delta_{ij}(t) \leq n. \tag{3}$$

Although user u_j has several available BSs, it can connect to at most one of its available BSs at any time. The connected available BS is called the **current BS** of user u_j.

TABLE 1
Notation Summary

n	Number of base stations
$\mathcal{B}$	Set of base stations $\{b_i\}$, $i = 1, 2, \cdots, n$
ω_i	The frequency band of base station b_i in MHz
ω_i'	The bandwidth that a user can get from base station b_i in MHz
η_i	The maximum users that base station b_i can serve simultaneously
m	Number of users
$\mathcal{U}$	Set of users $\{u_j\}$, $j = 1, 2, \cdots, m$
$\delta(t)$	Adjacency matrix of $\mathcal{B}$ and $\mathcal{U}$ at time t, i.e., $[\delta_{ij}(t)]$
$\theta(t)$	Conjunction matrix of $\mathcal{B}$ and $\mathcal{U}$ at time t, i.e., $[\theta_{ij}(t)]$
γ_{ij}	Basic bandwidth requirement of user u_j for base station b_i in Mbps
$s_{ij}(t)$	Received signal power of user u_j from base station b_i at time t in watts
$d_{ij}(t)$	Euclidean distance between base station b_i and user u_j at time t
ρ_i	Transmission power of base station b_i in watts
h_{ij}	Channel fading gain of channel (b_i, u_j)
λ	Pass loss exponent
ζ^2	Background additive white Gaussian noise in watts
$g_{ij}(t)$	The interference caused by base station b_i to user u_j in watts at time t
$q_{ij}(t)$	The achievable data receiving rate of user u_j from base station b_i at time t in Mbps
$v_j(t)$	Identifier that if user u_j at time t is a handoff user or a non-handoff user
$\mathcal{V}(t)$	$(v_j(t))$, $j = 1, 2, \cdots, m$
$\mathcal{B}_j(t)$	Set of available base stations for user u_j at time t, i.e., $\{b_{j_i}\}$, $\mathcal{B}_j(t) \subseteq \mathcal{B}$
$\mathcal{F}_j(t)$	Network selection result of handoff user u_j at time t, i.e., $(f_{j_i j})$, $\lvert\mathcal{F}_j(t)\rvert = \lvert\mathcal{B}_j(t)\rvert$
$\mathcal{P}_j(t)$	Block probabilities of user u_j for available base stations at time t, i.e., $(p_{j_i j})$, $\lvert\mathcal{P}_j(t)\rvert = \lvert\mathcal{B}_j(t)\rvert$
$Q_j(t)$	Achievable data receiving rates provided by available base stations for user u_j at time t, i.e., $(q_{j_i j})$, $\lvert Q_j(t)\rvert = \lvert\mathcal{B}_j(t)\rvert$
$\Theta_i(t)$	Number of non-handoff users which are connecting to base station b_i at time t
$\Delta_i(t)$	Number of hand-off users inside the coverage area of base station b_i at time t
$\alpha_{ij}(t)$	Probability that base station b_i will be selected by handoff user u_j at time t
$\beta_{ij}(t)$	Probability inferred by u_j that base station b_i will be selected by another handoff user at time t
$\mathbb{P}_j(r_i(t))$	Probability inferred by u_j that there are $r_i(t)$ other handoff users who have selected b_i as their new base station at time t
$\tau_{ij}(t)$	Throughput of channel (b_i, u_j) at time t in Mbps
ε	The maximal moving velocity of user in m/s

A **conjunction matrix** $\theta(t)$ is used to reflect the connected relationship between BSs and users at time t as follows.

$$\theta(t) = \begin{array}{c} \\ \boldsymbol{b_1} \\ \boldsymbol{b_2} \\ \vdots \\ \boldsymbol{b_n} \end{array} \begin{array}{c} \begin{array}{cccc} \boldsymbol{u_1} & \boldsymbol{u_2} & \cdots & \boldsymbol{u_m} \end{array} \\ \begin{bmatrix} \theta_{11}(t) & \theta_{12}(t) & \dots & \theta_{1m}(t) \\ \theta_{21}(t) & \theta_{22}(t) & \dots & \theta_{2m}(t) \\ \vdots & & \ddots & \vdots \\ \theta_{n1}(t) & \theta_{n2}(t) & \dots & \theta_{nm}(t) \end{bmatrix} \end{array},$$

where

$$\theta_{ij}(t) = \begin{cases} 1, & \text{current BS } b_i \text{ is connected by user } u_j \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

For BS b_i, the number of serving users is $\sum_{j=1}^{m}\theta_{ij}(t)$ which should satisfy the following constraint.

$$0 \le \sum_{j=1}^{m}\theta_{ij}(t) \le min\left(\eta_i, \sum_{j=1}^{m}\delta_{ij}(t)\right). \tag{5}$$

Each serving user can get $\omega_i'(t)$ MHz bandwidth from BS b_i at time t. The value of $\omega_i'(t)$ is calculated as follows.

$$\omega_i'(t) = \frac{\omega_i}{\sum_{j=1}^{m}\theta_{ij}(t)}, \forall i, \forall j, 0 \le \theta_{ij}(t) \le \delta_{ij}(t). \tag{6}$$

Since user u_j can connect to at most one current BS, the number of current BS $\sum_{i=1}^{n}\theta_{ij}(t)$ should satisfy the following constraint.

$$0 \le \sum_{i=1}^{n}\theta_{ij}(t) \le min\left(1, \sum_{i=1}^{n}\delta_{ij}(t)\right). \tag{7}$$

For each BS-user pair (b_i, u_j), assume that the received signal power of user u_j from available BS b_i at time t is $s_{ij}(t)$ in watts. Let $d_{ij}(t)$ denote the Euclidean distance between BS b_i and user u_j at time t. When b_i transmits a signal for each channel with power ρ_i in watts, $s_{ij}(t)$ is then calculated as follows.

$$s_{ij}(t) = \delta_{ij}(t)\cdot\rho_i\cdot h_{ij}\cdot d_{ij}(t)^{-\lambda}, \tag{8}$$

where the channel fading gain h_{ij} follows an exponential distribution with rate μ ($h_{ij} \sim exp(\mu)$), and the pass loss exponent $\lambda \ge 2$ (varies depending on channel conditions).

Since different BSs are assumed to use different frequency bands, there is no interference among BSs. For 5G supported BS which utilizes the Orthogonal Frequency Division Multiple Access (OFDMA, also commonly applied to LTE, WiMAX and IEEE 802.11 b/g supported devices) to avoid the interference among users. For those BSs which do not utilize the OFDMA, some techniques such as Code Division Multiple Access (CDMA, commly applied to 3G devices) and orthogonal codes are assumed to be used in order to waken the interference among users. Let $g_{xj}(t)$ in watts be the interference caused by BS b_x ($b_x \in \mathcal{B}, x \ne i$) to user u_j at time t, where b_x transmits signal by using the same frequency as user u_j. The value of g_{xj} can be calculated as follows.

$$g_{xj}(t) = \rho_x\cdot h_{xj}\cdot d_{xj}(t)^{-\lambda}, \tag{9}$$

where $d_{xj}(t)$ is the Euclidean distance between BS b_x and user u_j at time t. According to the Shannon theorem, the achievable data receiving rate of user u_j from BS b_i at time t denoted by $q_{ij}(t)$ in Mbps is calculated as follows.

$$q_{ij}(t) = \omega_i'(t)\cdot\log\left[1 + \frac{s_{ij}(t)}{\sum_{b_x\in\mathcal{B},x\ne i} g_{xj}(t) + \zeta^2}\right], \tag{10}$$

where ζ^2 is the background additive white Gaussian noise (AWGN).

In hybrid 5G environment, different kinds of BSs are assumed with different basic bandwidth requirements. Let γ_{ij} denote the basic bandwidth requirement of user u_j for BS b_i in Mbps. Suppose that the current BS of user u_j is b_c. If the achievable data receiving rate from b_c cannot meet the basic bandwidth requirement ($q_{cj}(t) < \gamma_{cj}$), user u_j will perform handoff. We call these users who need to perform handoff **handoff users**. If the achievable data receiving rate can satisfy the basic bandwidth requirement ($q_{cj}(t) \ge \gamma_{cj}$), user u_j will stay in the connection with its current BS b_c. We call these users who do not need handoff **non-handoff users**. A vector $\mathcal{V}(t) = (v_1(t), v_2(t), \cdots v_m(t))$ is used to identify the kinds of users. The value of $v_j(t)$ is given as follows, where $j = 1, 2, \cdots, m$.

$$v_j(t) = \begin{cases} 0, & \text{user } u_j \text{ is a handoff user at time } t, \\ 1, & \text{user } u_j \text{ is a non-handoff user at time } t. \end{cases} \tag{11}$$

Let $\Delta_i(t)$ be the number of **handoff users** which are **inside** the coverage area of BS b_i at time t. The value of $\Delta_i(t)$ is then calculated as follows.

$$\Delta_i(t) = \sum_{j=1}^{m}\{\delta_{ij}(t)\cdot[1 - v_j(t)]\}. \tag{12}$$

Let $\Theta_i(t)$ be the number of **non-handoff users** which are **connecting** to BS b_i at time t. The value of $\Theta_i(t)$ is then calculated as follows.

$$\Theta_i(t) = \sum_{j=1}^{m}[\theta_{ij}(t)\cdot v_j(t)]. \tag{13}$$

Note that there is no centralized control entity. Users perform network selection in a distributed way. Furthermore, users are assumed do not share their private information (such as the number of available BSs, channel capacities, and so on) for privacy preservation. Therefore, each user has to make its own network selection based on local information. Local information is acquired by a user including the private information of itself, parameters of BSs and two pieces of public information (*i.e.*, $\Delta_i(t)$ and $\Theta_i(t)$).

Users have a lot of ways to obtain the public information, such as BSs periodically broadcast, device-to-device communication and standard location update. At the beginning of each time slot, users can send Hello messages to their available BSs to announce their presences. After collecting these Hello messages, BSs count the number of handoff and non-handoff users, then broadcast the values. This procedure can be enhanced through the device-to-device communications in some special scenarios [18]: those devices which have already known the public information can notify their neighbors about the public information. In order to further reduce the overhead and information refresh time, BSs can make use of the location update processes provided by the communication standards (*e.g.*, GSM 03.12 [19], 3GPP TS 23.012 [20], Mobile IP [21], [22]). By embedding the Hello message and public information into the Channel Request, Immediate Assignment and other control frames, the overhead and refresh time will be reduced to a very low level even can be neglected [23].

Let $\mathcal{B}_j(t) = \{b_{j_1}, b_{j_2}, \cdots, b_{j_k}\}$ be the set of available BSs for user u_j at time t, where $\mathcal{B}_j(t) \subseteq \mathcal{B}$, $k = |\mathcal{B}_j(t)| = \sum_{i=1}^{n}\delta_{ij}(t)$. Since u_j is a handoff user at time t, it has to select **a new** BS from $\mathcal{B}_j(t)$. Let $\mathcal{F}_j(t) = (f_{j_1 j}(t), f_{j_2 j}(t), \cdots, f_{j_k j}(t))$ be the network selection result of user u_j at time t, where $|\mathcal{F}_j(t)| = |\mathcal{B}_j(t)|$. The value of $f_{j_i j}(t)$ is given as follows, where $i = 1, 2, \cdots, k$.

$$f_{j_i j}(t) = \begin{cases} 1, & \text{new BS } b_{j_i} \text{ is selected by } u_j \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

For handoff user u_j, let $Q_j(t) = (q_{j_1 j}(t), q_{j_2 j}(t), \cdots, q_{j_k j}(t))$ be the achievable data receiving rates provided by its available BSs, where $|Q_j(t)| = |\mathcal{B}_j(t)|$. Hence, the achievable data receiving rate that u_j can obtain from its new BS is $\mathcal{F}_j(t) \cdot [Q_j(t)]^T$ in Mbps, where $[Q_j(t)]^T$ is the transposition of $Q_j(t)$.

$$\mathcal{F}_j(t) \cdot [Q_j(t)]^T = \sum_{i=1}^{k} [f_{j_i j}(t) \cdot q_{j_i j}(t)]. \tag{15}$$

The achievable data receiving rate provided by new BS should satisfy the basic bandwidth requirement of user u_j for the new BS. That is, the value of $\mathcal{F}_j(t) \cdot [Q_j(t)]^T$ should be subject to the following constraint.

$$\mathcal{F}_j(t) \cdot [Q_j(t)]^T \geq \sum_{i=1}^{k} [f_{j_i j}(t) \cdot \gamma_{j_i j}(t)]. \tag{16}$$

If no available BS can satisfy the above constraint, the handoff of u_j will fail. In order to guarantee the quality of experience of other users, u_j will be discarded by its **current** BS. For handoff user u_j, its available BS b_{j_i} ($b_{j_i} \in \mathcal{B}_j(t)$) can serve at most η_{j_i} users simultaneously. Since there are $\Theta_{j_i}(t)$ non-handoff users connecting to BS b_{j_i} at time t, b_{j_i} can serve at most $\eta_{j_i} - \Theta_{j_i}(t)$ handoff users. Note that time t refers to the t th time slot. Handoff requests will come to a BS successively during a time slot. If there are more than $\eta_{j_i} - \Theta_{j_i}(t)$ handoff users that have chosen b_{j_i} as their new BS at time t, the after coming handoff requests will be blocked. These blocked handoff users will wait in a First-In-First-Out (FIFO) queue.

Let $p_{j_i j}(t)$ be the probability that handoff user u_j is blocked, when it tries to handoff to the BS b_{j_i} at time t. The calculation method of $p_{j_i j}(t)$ will be given in Section 4. Let $\mathcal{P}_j(t) = (p_{j_1 j}(t), p_{j_2 j}(t), \cdots, p_{j_k j}(t))$. Then, the block probability of u_j for its new BS is $\mathcal{F}_j(t) \cdot [\mathcal{P}_j(t)]^T$, where $[\mathcal{P}_j(t)]^T$ is the transposition of $\mathcal{P}_j(t)$.

$$\mathcal{F}_j(t) \cdot [\mathcal{P}_j(t)]^T = \sum_{i=1}^{k} [f_{j_i j}(t) \cdot p_{j_i j}(t)]. \tag{17}$$

For a single handoff user, the objectives of its network selection are to maximize the achievable data receiving rate provided by the new BS, and to minimize the block probability. We theoretically formulate the handoff problem as a multi-objective optimization problem as follows.

$$\begin{aligned} \mathcal{O}_1 &= \text{Maximize } \ \mathcal{F}_j(t) \cdot [Q_j(t)]^T \\ \mathcal{O}_2 &= \text{Minimize } \ \mathcal{F}_j(t) \cdot [\mathcal{P}_j(t)]^T \end{aligned} \tag{18}$$

subject to

$$\sum_{i=1}^{k} f_{j_i j}(t) + v_j(t) \leq 1, \ j \in \mathbb{Z}^+, \ 1 \leq j \leq m, \tag{19a}$$

$$\sum_{i=1}^{k} [f_{j_i j}(t) \cdot \gamma_{j_i j}(t)] \leq \sum_{i=1}^{k} [f_{j_i j}(t) \cdot q_{j_i j}(t)], \tag{19b}$$

$$b_{j_i} \in \mathcal{B}_j(t), \ \ \mathcal{B}_j(t) \subseteq \mathcal{B}, \ \ k = |\mathcal{B}_j(t)|. \tag{19c}$$

The first constraint Eqn. (19a) indicates that a non-handoff user ($v_j(t) = 1$) does not have any new BS and a handoff user ($v_j(t) = 0$) has at most one new BS. The second constraint Eqn. (19b) guarantees that the achievable data receiving rate provided by the new BS can satisfy the basic bandwidth requirement of a handoff user. The last constrain Eqn. (19c) reveals that the network selection of a handoff user should be implemented within its available BS set.

4 BLOCK PROBABILITY ESTIMATION

Based on limited local information, each handoff user tries to select a new BS which can provide the maximal achievable data receiving rate and minimal block probability. The calculation method of achievable data receiving rate has been given in Section 3. In this section, we will explain the estimation method of block probability.

4.1 Relations Between Users

The block probability of a handoff user relates to the network selection behaviors of other handoff users. However, under the premise of privacy preservation, a user has no idea of other handoff users. The calculation of block probability relies on the **inferences** made by a handoff user to other handoff users. In order to assist a handoff user in inferring the network selection behaviors of other handoff users, we study the relations between handoff users in this subsection.

In hybrid 5G environments, each handoff user has several available BSs. We investigate the relations between any two handoff users based on their available BS sets. In general, the relations of a pair of handoff users can be divided into two categories: independent relation and correlated relation.

Definition 4.1 (*The independent relation of a pair of handoff users*). Let (u_i, u_j) denote any pair of handoff users. Their available network sets at time t are $\mathcal{B}_i(t)$ and $\mathcal{B}_j(t)$ respectively. If $|\mathcal{B}_i(t) \cap \mathcal{B}_j(t)| = 0$, u_i and u_j have the independent relation. ☐

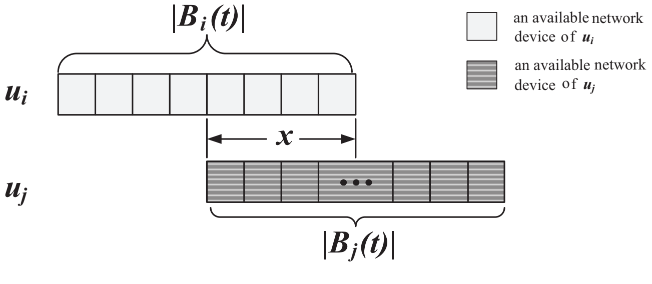


Fig. 2. The correlation degree of (u_i, u_j).

When u_i and u_j are independent, the network selection behavior of u_i has no direct impact on u_j, and vice versa. Hence in the handoff process, a user only needs to consider those users who are in the correlated relations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2016.2551042, IEEE Transactions on Emerging Topics in Computing

Definition 4.2 (*The correlated relation of a pair of handoff users*). For any pair of handoff users (u_i, u_j), if there is at least one BS which is available to both of them, then they have the correlated relation. Consequently, u_i and u_j are mutually neighbors. □

In order to reflect the strength of correlated relation, we define the *correlation degree* as follows.

Definition 4.3 (*The correlation degree of a pair of handoff users*). (u_i, u_j) is any pair of handoff users, their available BS sets are $\mathcal{B}_i(t)$ and $\mathcal{B}_j(t)$ respectively. The correlation degree of (u_i, u_j) is the probability that when selecting a BS from $\mathcal{B}_i(t) \cup \mathcal{B}_j(t)$, the selected BS is available to both u_i and u_j.□

Let $\mathcal{L}(u_i, u_j)$ denote the correlation degree of (u_i, u_j). Suppose that $\mid \mathcal{B}_i(t) \cap \mathcal{B}_j(t) \mid = x$ as shown in Fig.2, then $\mathcal{L}(u_i, u_j)$ can be calculated by the following equation,

$$\mathcal{L}(u_i, u_j) = \frac{x}{|\mathcal{B}_i(t)| + |\mathcal{B}_j(t)| - x},$$
$$\text{where } x \in \mathbb{Z}^+,\ 0 < x \le min(|\mathcal{B}_i(t)|, |\mathcal{B}_j(t)|). \quad (20)$$

Note that if u_i and u_j are independent, the value of x is 0, and the correlation degree $\mathcal{L}(u_i, u_j) = 0$. If u_i and u_j are correlated, $\mathcal{L}(u_i, u_j) \in (0, 1]$. As a result, we can extend the defining field of x in Eqn. (20) to $[0, min(|\mathcal{B}_i(t)|, |\mathcal{B}_j(t)|)]$ and use just one metric *correlation degree* to distinguish the categories of relations, and reflect the strength of association. The *correlation degree* metric also has the following attribute.

Theorem 4.1: For any pair of handoff users (u_i, u_j), their available BS sets are $\mathcal{B}_i(t)$ and $\mathcal{B}_j(t)$ respectively. If the correlation degree $\mathcal{L}(u_i, u_j) = 1$, $\mathcal{B}_i(t)$ is equal to $\mathcal{B}_j(t)$.

Proof. Based on Eqn. (20), if $\mathcal{L}(u_i, u_j) = 1$, then $2x = |\mathcal{B}_i(t)| + |\mathcal{B}_j(t)|$. For the correlation degree $\mathcal{L}(u_i, u_j) = 1$, we have the following cases.

Case 1: $|\mathcal{B}_i(t)| < |\mathcal{B}_j(t)|$, *i.e.*, $min(|\mathcal{B}_i(t)|, |\mathcal{B}_j(t)|) = |\mathcal{B}_i(t)|$. Substitute this equation into the constraint $x \le min(|\mathcal{B}_i(t)|, |\mathcal{B}_j(t)|)$ of Eqn. (20), we can get that $2x \le |\mathcal{B}_i(t)| + |\mathcal{B}_i(t)|$. Since $2x = |\mathcal{B}_i(t)| + |\mathcal{B}_j(t)|$, $|\mathcal{B}_j(t)|$ should be not bigger than $|\mathcal{B}_i(t)|$ which contradicts with the premise of Case 1. That is Case 1 will not happen when the correlation degree $\mathcal{L}(u_i, u_j) = 1$.

Case 2: $|\mathcal{B}_i(t)| > |\mathcal{B}_j(t)|$, *i.e.*, $min(|\mathcal{B}_i(t)|, |\mathcal{B}_j(t)|) = |\mathcal{B}_j(t)|$. Similar to the previous case, we can get that $2x \le |\mathcal{B}_j(t)| + |\mathcal{B}_j(t)|$. Since we already know that $2x = |\mathcal{B}_i(t)| + |\mathcal{B}_j(t)|$, then $|\mathcal{B}_i(t)|$ should be smaller than or equal to $|\mathcal{B}_j(t)|$ which contradicts with the premise of Case 2. That is Case 2 will not happen when the correlation degree $\mathcal{L}(u_i, u_j) = 1$.

For the relationship between $\mathcal{B}_i(t)$ and $\mathcal{B}_j(t)$, we have excluded $|\mathcal{B}_i(t)| < |\mathcal{B}_j(t)|$ and $|\mathcal{B}_i(t)| > |\mathcal{B}_j(t)|$ through the above discussions. Therefore, $|\mathcal{B}_i(t)| = |\mathcal{B}_j(t)|$. Furthermore, from the Definition 4.3 we observed that $x = \mid \mathcal{B}_i(t) \cap \mathcal{B}_j(t) \mid = |\mathcal{B}_i(t)| = |\mathcal{B}_j(t)|$ when the correlation degree $\mathcal{L}(u_i, u_j) = 1$. As a result, the available BS sets $\mathcal{B}_i(t)$ and $\mathcal{B}_j(t)$ are completely overlapping when their correlation degree is equal to 1. Theorem 4.1 is proved. □

Here, we want to explain the reason that why we specially proposed and studied the *correlation degree* metric in this subsection. Remember that, we investigate handoff user relations for the purpose of assisting a handoff user to infer the network selection behaviors of other handoff users. It requires a metric which can reflect the relation between handoff users. Thus, we proposed the *correlation degree* metric in Definition 4.3. During the behavior inference, since a handoff user does not know any private information of other handoff users, the handoff user will consider the worst case (*i.e.*, the correlation degree is 1) to be on the safe side. Through Theorem 4.1 we observed that the available BS sets of two handoff users will be completely overlapping in the worst case. This conclusion is meaningful since a handoff user can infer the network selection behaviors of other handoff users based on its own available BS set.

4.2 Behaviors Inference

For a handoff user u_j ($u_j \in \mathcal{U}$), since u_j has no idea of other handoff users, these handoff users are indistinguishable for u_j. We use u to represent an arbitrary one of them. In order to estimate its block probability for each available BS, u_j has to infer the network selection behavior of u [24].

Suppose that BS b_i is available to handoff user u_j at time t (*i.e.*, $\delta_{ij}(t) = 1$). There are two conditions needed to be satisfied simultaneously, if u_j is blocked when it tries to handoff to b_i [25]. These two conditions are: 1) u_j selects b_i as the new BS in a handoff; 2) before u_j tries to handoff to b_i, b_i is already full load.

For the first condition, we assume that u_j selects BSs based on their achievable data receiving rates. The larger achievable data receiving rate, the higher probability to be selected. As a result, the BS b_i will be selected as the new BS by u_j at time t with the probability $\alpha_{ij}(t)$ as follows.

$$\alpha_{ij}(t) = \frac{q_{ij}(t)}{\sum_{k=1}^{n} [\delta_{kj}(t) \cdot q_{kj}(t)]}. \quad (21)$$

During the network selection behavior inference, u_j always considers the worst case with the other handoff user u (*i.e.*, $\mathcal{L}(u, u_j) = 1$). According to Theorem 4.1 we can get that the available BS sets of u_j and u are completely overlapping in the worst case. Since u_j does not know the private information of u (such as how much achievable data receiving rate that u can obtain from each available BS, the specific location of u, and so on), u_j has no choice but to assume that u selects BS based on the remaining bandwidth. Note that the bandwidth of BS b_i is ω_i MHz. Moreover, there are $\Theta_i(t)$ non-handoff users are connecting to BS b_i at time t. Each non-handoff users will occupy $\omega'_i(t)$ MHz bandwidth of b_i. As a result, the remaining bandwidth of b_i is $\omega_i - \omega'_i(t) \cdot \Theta_i(t)$ MHz. Hence, u_j infers that b_i will be selected as the new BS by u at time t with the probability $\beta_{ij}(t)$ as follows.

$$\beta_{ij}(t) = \frac{\omega_i - \omega'_i(t) \cdot \Theta_i(t)}{\sum_{k=1}^{n} \delta_{kj}(t) \cdot [\omega_k - \omega'_k(t) \cdot \Theta_k(t)]}. \quad (22)$$

Note that there are $\Delta_i(t)$ handoff users (including u_j) inside the coverage area of BS b_i at time t. Let $\mathbb{P}_j(r_i(t))$ denote the probability that before u_j, there are $r_i(t)$ handoff users that have chosen b_i as their new BS at time t. The value of $\mathbb{P}_j(r_i(t))$ is calculated as follows.

$$\mathbb{P}_j(r_i(t)) = \binom{\Delta_i(t) - 1}{r_i(t)} \cdot \beta_{ij}(t)^{r_i(t)} \cdot (1 - \beta_{ij}(t))^{\Delta_i(t) - 1 - r_i(t)},$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2016.2551042, IEEE Transactions on Emerging Topics in Computing

$$\text{where } \binom{x}{y} = \frac{x!}{y! \cdot (x-y)!}. \tag{23}$$

If u_j is blocked when it tries to handoff to the new BS b_i, that means b_i has been full load. As a result, the value of $r_i(t)$ should satisfy the following constraint.

$$\eta_i - \Theta_i(t) \le r_i(t) \le \Delta_i(t) - 1. \tag{24}$$

Based on its private information, parameters of BSs and two pieces of public information (*i.e.*, the number of non-handoff users $\Theta_i(t)$ and the number of handoff users $\Delta_i(t)$), handoff user u_j estimates its block probability for BS b_i at time t denoted by $p_{ij}(t)$ as follows.

$$p_{ij}(t) = \alpha_{ij}(t) \cdot \sum_{r_i(t)=\eta_i-\Theta_i(t)}^{\Delta_i(t)-1} \mathbb{P}_j(r_i(t)). \tag{25}$$

5 PROPOSED HANDOFF SCHEME

The handoff problem has been formulated as a multi-objective optimization problem [11]. Unfortunately, for most of multi-objective optimization problems, there does not exist a solution which simultaneously optimizes each objective. In our scheme, a handoff user is unable to find a BS which exactly provides maximal achievable data receiving rate and minimal block probability simultaneously either. However, our proposed scheme is able to find a Pareto Optimal [26] network selection for the formulated multi-objective optimization problem. A network selection is Pareto Optimal if and only if there does not exist another network selection which promotes at least one objective without demoting any one objective [27]. In this section, we will explain how to solve the formulated multi-objective optimization problem and find a Pareto Optimal network selection.

By taking the *throughput* metric into consideration, we firstly transform the original multi-objective optimization problem into a maximization problem. As an available BS of handoff user u_j, BS b_{j_i} is tagged with two attributes: achievable data receiving rate denoted by $q_{j_ij}(t)$ and block probability denoted by $p_{j_ij}(t)$. Let $\tau_{j_ij}(t)$ be the throughput of channel (b_{j_i}, u_j) at time t in Mbps. If b_{j_i} is not selected as the new BS by handoff user u_j at time t (*i.e.*, $f_{j_ij}(t) = 0$), $\tau_{j_ij}(t) = 0$. If b_{j_i} is selected as the new BS (*i.e.*, $f_{j_ij}(t) = 1$) but handoff user u_j is blocked in b_{j_i}, $\tau_{j_ij}(t) = 0$. If b_{j_i} is selected as the new BS (*i.e.*, $f_{j_ij}(t) = 1$) and u_j successfully gets the network service, $\tau_{j_ij}(t) = q_{j_ij}(t)$. In summary, the value of throughput $\tau_{j_ij}(t)$ is calculated as follows.

$$\tau_{j_ij}(t) = f_{j_ij}(t) \cdot q_{j_ij}(t) \cdot [1 - p_{j_ij}(t)]. \tag{26}$$

Note that, the *throughput* metric involves both of two attributes (*i.e.*, achievable data receiving rate and block probability) what a handoff user is concerned. Furthermore, the throughput metric is proportional to the achievable data receiving rate attribute and inversely proportional to the block probability attribute. Thus, it is reasonable to substitute the following objective $\mathcal{O}_3$ for the original multiple objectives $\mathcal{O}_1$ and $\mathcal{O}_2$ under the same constraints listed in Eqn. (19).

$$\mathcal{O}_3 = \text{Maximize } \sum_{i=1}^{k} \{f_{j_ij}(t) \cdot q_{j_ij}(t) \cdot [1 - p_{j_ij}(t)]\}. \tag{27}$$

By solving the maximization problem (*i.e.*, $\mathcal{O}_3$), a handoff user can select a new BS. We will prove that this selected new BS is a Pareto Optimal solution of the original multi-objective optimization problem (*i.e.*, $\mathcal{O}_1$ and $\mathcal{O}_2$) through the following theorem.

Theorem 5.1: The solution of the transformed maximization problem $\mathcal{O}_3$ is a Pareto Optimal result of the original multi-objective optimization problem $\mathcal{O}_1 \& \mathcal{O}_2$.

Proof. The objective $\mathcal{O}_1$ can be equivalently transformed into $\mathcal{O}_1'$ as follows:

$$\mathcal{O}_1' = \text{Maximize} \sum_{i=1}^{k} [f_{j_ij}(t) \cdot q_{j_ij}(t)] + \sum_{i=1}^{k} f_{j_ij}(t)$$

$$= \text{Maximize} \sum_{i=1}^{k} [f_{j_ij}(t) \cdot q_{j_ij}(t)] + \sum_{i=1}^{k} \{f_{j_ij}(t) \cdot [1 - p_{j_ij}(t)]\}$$

$$+ \sum_{i=1}^{k} [f_{j_ij}(t) \cdot p_{j_ij}(t)].$$

The $\mathcal{O}_2$ can be equivalently transformed into $\mathcal{O}_2'$ as follows:

$$\mathcal{O}_2' = \text{Maximize} \sum_{i=1}^{k} [-f_{j_ij}(t) \cdot p_{j_ij}(t)].$$

The most common approach to multi-optimization problem is the weighted sum method [28], in which multiple objectives are weighted summed and merged into a single objective. Let the weights of $\mathcal{O}_1'$ and $\mathcal{O}_2'$ are 1, then these two objectives can be merged into a single objective $\mathcal{O}_{1\&2}'$ as follows:

$$\mathcal{O}_{1\&2}' = \mathcal{O}_1' + \mathcal{O}_2'$$

$$= \text{Maximize } \sum_{i=1}^{k} \{f_{j_ij}(t) \cdot [q_{j_ij}(t) + 1 - p_{j_ij}(t)]\}.$$

Zadeh *et al.* [29] proved that if all of the weights are positive, the newly merged objective is Pareto Optimal. That is, the solution of $\mathcal{O}_{1\&2}'$ is a Pareto Optimal solution of the multi-objective (i.e., $\mathcal{O}_1'$ and $\mathcal{O}_2'$) optimization problem. Furthermore, since log(x) is a monotone increasing function, the objective $\mathcal{O}_{1\&2}'$ can be equivalently transformed into $\mathcal{O}_{1\&2}''$ as follows:

$$\mathcal{O}_{1\&2}'' = \text{Maximize} \sum_{i=1}^{k} \{f_{j_ij}(t) \cdot [\log q_{j_ij}(t) + \log(1 - p_{j_ij}(t))]\}.$$

Note that, the objective $\mathcal{O}_3$ can be equivalently transformed into $\mathcal{O}_3'$ as follows,

$$\mathcal{O}_3' = \text{Maximize} \sum_{i=1}^{k} \{f_{j_ij}(t) \cdot \log [q_{j_ij}(t) \cdot (1 - p_{j_ij}(t))]\}$$

$$= \mathcal{O}_{1\&2}''$$

Through several times equivalent transformations and once weighted sum, the multiple objectives $\mathcal{O}_1$ and $\mathcal{O}_2$ are transformed into $\mathcal{O}_3$. Therefore, the solution of $\mathcal{O}_3$ must be a Pareto Optimal solution of the original multi-objective (i.e., $\mathcal{O}_1$ and $\mathcal{O}_2$) optimization problem. Theorem 5.1 is proved. □

At the beginning of each time slot, users compare the achievable data receiving rates of their current BSs with the basic bandwidth requirements, and decide whether to implement the handoffs or not. If a user does not need handoff (non-handoff user), it will stay in the connection with its current BS. Otherwise, the user (handoff user) will select a new BS through Algorithm 1. In a real system, Algorithm 1 will be executed on mobile terminals. It is necessary to consider the general limitations of mobile terminals, such as small storage space and limited processing capacity. Hence, we will analyze the computation and memory complexities of Algorithm 1 through Theorem 5.2.

Algorithm 1: Steps of the Proposed Handoff Scheme for Handoff User u_j

```
Input: available BS set at time $t$ $\mathcal{B}_j(t)$; for $\forall b_{j_i} \in \mathcal{B}_j(t)$:
       maximum number of users can be served $\eta_i$,
       frequency band $\omega_{j_i}$, bandwidth that a user can
       get $\omega'_{j_i}$, number of handoff and non-handoff
       users $\Delta_{j_i}(t)$ and $\Theta_{j_i}(t)$, received signal power
       $s_{j_ij}(t)$, noise power $n_{j_ij}(t)$, interference caused
       by other BS $g_{xj}(t)$ and basic bandwidth
       requirement $\gamma_{j_ij}$.
Output: network selection result $\mathcal{F}_j(t)$.
1  $max = 0, index = 0$;
2  for $\forall b_{j_i} \in \mathcal{B}_j(t)$ do
3      Calculate the achievable data receiving rate $q_{j_ij}(t)$
       by Eqn. (10);
4      Estimate the block probability $p_{j_ij}(t)$ by Eqn. (25);
5      if $q_{j_ij}(t) \geq \gamma_{j_ij}$ then
6          Calculate the throughput $\tau_{j_ij}(t)$ by Eqn. (26);
7          if $\tau_{j_ij}(t) \geq max$ then
8              $\tau_{j_ij}(t) \rightarrow max$;
9              the index of the selected BS $index = i$;
10 for $i = 1$; $i \leq |\mathcal{B}_j(t)|$; $i$++ do
11     if $i == index$ then
12         the selected BS is $b_{j_i}$, $f_{j_ij}(t) = 1$;
13     else
14         $f_{j_ij}(t) = 0$;
15     return $\mathcal{F}_j(t)$;
```

Theorem 5.2: The computation complexity of the proposed scheme is $O(mn)$, the memory complexity of the proposed scheme is $O(n)$.

Proof. The major computational work of Algorithm 1 consists of three parts: calculate the achievable data receiving rates of available BSs (Line 3); estimate the block probabilities for available BSs (Line 4); scan the available BSs and calculate their throughput (Line 6), then determine the new BS (from Line 5 to 15).

Consider a scenario which has m users and n BSs. The first part is just a numerical calculation, its computation complexity is $O(n)$. For the second part, the computation complexity of $\mathbb{P}_j(r_i(t))$ is $O(1)$ (Eqn. (23)). In order to estimate the block probability for an available BS, a handoff user has to perform at most $m-1$ times calculations of $\mathbb{P}_j(r_i(t))$ (Eqn. (25)). Therefore, the computation complexity of the second part is $O(mn)$. Since the calculation of throughput is also a simple numerical calculation, the computation complexity of the third part is $O(n)$. As a result, the computation complexity of our proposed scheme is $O(mn)$.

Since the first part is just a numerical calculation, the memory complexity of this part is $O(n)$. For the second part, we can make use of the recurrence method during the calculation process of Eqn. (25). Therefore, the memory complexity of Eqn. (25) is $O(1)$. Consequently, the memory complexity of the second part is $O(n)$. For the third part, since we only need to store the information of the current optimal BS, the memory complexity of the third part is $O(1)$. As a result, the memory complexity of our proposed scheme is $O(n)$. ☐

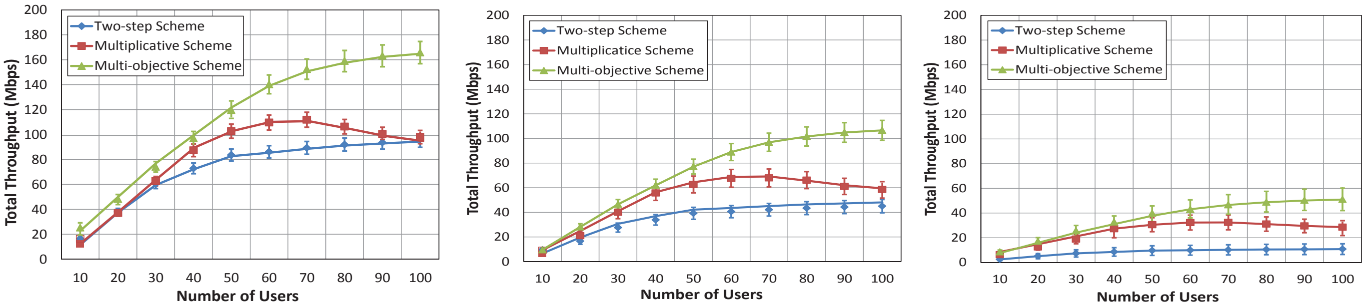
Above discussions illustrate that our scheme has polynomial time and linear space complexities which are suitable for ordinary mobile terminals.

6 PERFORMANCE EVALUATION

We compare the proposed scheme with two recent typical distributed handoff schemes: the multiplicative scheme [14] and the two-step scheme [15] under various network conditions. Over a 500m × 500m rectangular flat space, we randomly place 3 BSs and several users. A BS is available to a user when the distance between them is smaller than the coverage radius of this BS. In order to simulate a small hybrid 5G environment, we set the parameters of these 3 BSs refer to 3G, 4G and 5G techniques respectively. According to the 3G (W-CDMA/HSDPA) standard [30], we set the coverage radius of 3G BS to be 7 km, set the bandwidths and transmission power to be 5 MHz and 10 watts. According to the 4G (802.16a) standard [31], we set the coverage radius, bandwidth and transmission power to be 50 km, 20 MHz and 20 watts respectively. So far the 5G standard is still being figured out. However, Andrews *et al.* [1] pointed out the 5G BS will have higher bandwidth, higher transmission power, smaller cell size and ever-smaller serving users compared with 4G BS. Thus, we set the coverage radius, bandwidth and transmission power of 5G BS to be 25 km, 40 MHz and 40 watts accordingly. Users are moving around inside the hybrid 5G environment. If the current location of a user is denoted by a two-dimensional coordinate (x, y), this user will be inside $(x \pm \triangle t \cdot \varepsilon, y \pm \triangle t \cdot \varepsilon)$ after a period of time $\triangle t$, where ε is the maximal moving velocity of the user [32], [33]. For convenience we assume that users have the same basic bandwidth requirements for a single BS. We set the basic bandwidth requirements of users to be equal to or greater than 2 Mbps, which corresponds to the video conference demanding. Some important experimental parameters are presented in Table 1 [31]. The concerned performance metrics are total throughput and ratio of users served. Simulation experiments are repeated one thousand times and the results are presented with 95% confidence interval.

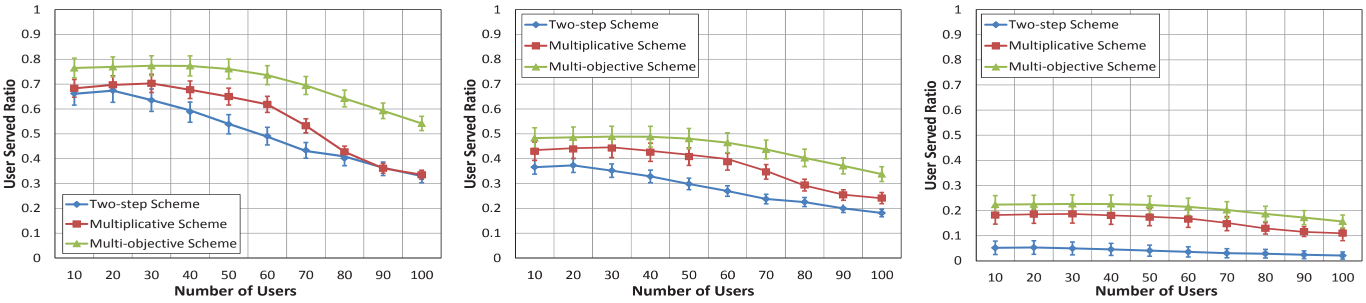
6.1 Total Throughput

Total throughput is defined as the sum of throughput that handoff and non-handoff users can obtain. According to the analysis and discussion in Section 5, the throughput metric can reflect two performance attributes that users care



(a) pass loss exponent $\lambda = 2$ (b) pass loss exponent $\lambda = 3$ (c) pass loss exponent $\lambda = 4$

Fig. 3. Number of users vs. throughput.



(a) pass loss exponent $\lambda = 2$ (b) pass loss exponent $\lambda = 3$ (c) pass loss exponent $\lambda = 4$

Fig. 4. Number of users vs. ratio of users served.

TABLE 2
Experimental Parameters

Parameter	Value
Number of BSs	3
Coverage radii of BSs	7km, 50km, 25km
Maximum number of serving users in BSs	10, 20, 15
Bandwidths of BSs	5MHz, 20MHz, 40MHz
Transmission powers of BSs	10 watts, 20 watts, 40 watts
Basic bandwidth requirements of users for BSs	2 Mbps, 4 Mbps, 6 Mbps
Time slot	1 second
Channel fading gain h	$h \sim exp(1)$
Additive white Gaussian noise power ζ^2	$\zeta^2 \sim N(0,1)$ watts
Moving velocities of users	$0 \sim 5$ m/s

about: achievable data receiving rate and block probability. We study the total throughput when the number of users varies under free space propagation ($\lambda = 2$), flat-earth reflection ($\lambda = 3$) and diffraction losses ($\lambda = 4$) environment conditions in Fig. 3.

The general trend is that the total throughput will be higher as more users join in. For the same scenario, the proposed multi-objective scheme always has the highest total throughput. From the crosswise comparison we observe that the total throughput in three schemes declines in tougher environments. Another interesting observation is that the total throughput in multiplicative scheme slightly reduces when the number of users is bigger than around 50. After careful deliberation, we consider that the reason behind this phenomenon is network congestion.

6.2 Ratio of Users Served

Ratio of users served refers to the ratio of users who have the network service. Following the notations made in problem formulation, the ratio of users served is equal to $\frac{\sum_{i=1}^{n}\Theta_i(t)+\sum_{i=1}^{n}\sum_{j=1}^{m}\{f_{ij}(t)\cdot[1-p_{ij}(t)]\}}{m}$, where m is the number of uses, $\sum_{i=1}^{n}\Theta_i(t)$ is the number of non-handoff users and $\sum_{i=1}^{n}\sum_{j=1}^{m}\{f_{ij}(t)\cdot[1-p_{ij}(t)]\}$ is the number of handoff successful users. The ratio of users served metric is used to reflect the fairness in three handoff schemes.

The experiment results shown in Fig. 4 revel that there are more users can get service in our proposed scheme. Furthermore, the ratio of users served in our scheme will maintain stable then decline as the number of users increases. Comparatively, the ratios of users served in two contrast schemes will slightly increase then decrease. Moreover, there is an obvious downtrend in multiplicative scheme when the number of users is around 50. It will not be difficult to find that the inflection point of multiplicative scheme in Fig. 4 is very close to that in Fig. 3. This is another proof of network congestion.

7 CONCLUSION

We proposed a user centered handoff scheme fulling multiple objectives for hybrid 5G environments. We consider the general limitations in hybrid 5G environments that users are unwilling to share their private information and centralized control usually is inefficient in large scare scenario. Based on limited local information, a user has to make the network selection by itself. We exploited two performance attributes to evaluate BSs: achievable data receiving rate and blocking probability. When a handoff user needs to select a new BS, it will calculate the achievable data receiving rates of available BSs. Then the user has to infer the network selection behaviors of other users in order to estimate its blocking probability for each available BS. By jointly considering these two attributes, a user can select the most appropriate one as its new BS for a handoff.

REFERENCE

[1] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Aurelie Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang. What will 5g be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, 2014.

[2] Cheng-Xiang Wang, Fourat Haider, Xiqi Gao, Xiao-Hu You, Yang Yang, Dongfeng Yuan, Hadi Aggoune, Harald Haas, Sam Fletcher, and Erol Hepsaydir. Cellular architecture and key technologies for 5g wireless communication networks. *IEEE Communications Magazine*, 52(2):122–130, 2014.

[3] Ruidong Li, Jie Li, Kui Wu, Yang Xiao, and Jiang Xie. An enhanced fast handover with low latency for mobile ipv6. *IEEE Transactions on Wireless Communications*, 7(1):334–342, 2008.

[4] Yang Xiao, Yi Pan, and Jie Li. Design and analysis of location management for 3g cellular networks. *IEEE Transactions on Parallel and Distributed Systems*, 15(4):339–349, 2004.

[5] Jie Li, Hisao Kameda, and Keqin Li. Optimal dynamic mobility management for pcs networks. *IEEE/ACM Transactions on Networking (TON)*, 8(3):319–327, 2000.

[6] Khoriba Ghada, Jie Li, and Yusheng Ji. Energy and mobility-aware topology control in heterogeneous mobile ad hoc networks. *International Journal of Computational Science and Engineering*, 5(2):147–153, 2010.

[7] Jie Li, Yi Pan, and Yang Xiao. Performance study of multiple route dynamic source routing protocols for mobile ad hoc networks. *Journal of parallel and distributed computing*, 65(2):169–177, 2005.

[8] Toshikazu Sakano, Zubair Md Fadlullah, Thuan Ngo, Hiroki Nishiyama, Masataka Nakazawa, Fumiyuki Adachi, Nei Kato, Atsushi Takahara, Tomoaki Kumagai, Hironori Kasahara, et al. Disaster-resilient networking: a new vision based on movable and deployable resource units. *IEEE Network*, 27(4):40–46, 2013.

[9] Yang Xiao, Haizhon Li, Yi Pan, Kui Wu, and Jie Li. On optimizing energy consumption for mobile handsets. *IEEE Transactions on Vehicular Technology*, 53(6):1927–1941, 2004.

[10] Pyo ChangWoo, Li Jie, and Hisao Kameda. An efficient caching scheme for personal communication service networks. *IEICE transactions on communications*, 87(12):3603–3610, 2004.

[11] Kalyanmoy Deb. Multi-objective optimization. In *Search methodologies*, pages 403–449. Springer, 2014.

[12] Kemeng Yang, Iqbal Gondal, and et al. Combined sinr based vertical handoff algorithm for next generation heterogeneous wireless networks. In *Proceedings of The 2007 Global Telecommunications Conference (IEEE GLOBECOM 2007)*, pages 4483–4487. IEEE, 2007.

[13] Li Qiang, Jie Li, and Changcheng Huang. A software-defined network based vertical handoff scheme for heterogeneous wireless networks. In *Proceedings of The 2014 IEEE Global Communications Conference (IEEE GLOBECOM 2014)*, pages 4671–4676. IEEE, 2014.

[14] Quoc-Thinh Nguyen-Vuong, Nazim Agoulmine, El Hadi Cherkaoui, and Laura Toni. Multi-criteria optimization of access selection to improve the quality of experience in heterogeneous wireless access networks. *IEEE transactions on vehicular technology*, 62(EPFL-ARTICLE-182868):1785–1800, 2013.

[15] Chao Liu, Yong Sun, Peng Yang, Zhen Liu, Haijun Zhang, and Xiangming Wen. A two-step vertical handoff decision algorithm based on dynamic weight compensation. In *Proceedings of The 2013 IEEE International Conference on Communications Workshops (IEEE ICC 2013)*, pages 1031–1035. IEEE, 2013.

[16] Kenichi Taniuchi, Yoshihiro Ohba, Victor Fajardo, Subir Das, Miriam Tauil, Yuu-Heng Cheng, Ashutosh Dutta, Donald Baker, Maya Yajnik, and David Famolari. Ieee 802.21: Media independent handover: Features, applicability, and realization. *IEEE Communications Magazine*, 47(1):112–120, 2009.

[17] Meng Li, Hiroki Nishiyama, Nei Kato, Yasunori Owada, and Kiyoshi Hamaguchi. On the energy-efficient of throughput-based scheme using renewable energy for wireless mesh networks in disaster area. *IEEE Transactions on Emerging Topics in Computing*, 3(3):420–431, 2015.

[18] Jiajia Liu, Shangwei Zhang, Nei Kato, Hirotaka Ujikawa, and Kenichi Suzuki. Device-to-device communications for enhancing quality of experience in software defined multi-tier lte-a networks. *IEEE Network*, 29(4):46–52, 2015.

[19] ETSI TC-SMG. Digital cellular telecommunications system - location registration procedures. [Online] http://www.etsi.org/deliver/etsi_gts/03/0312/05.00.00_60/gsmts_0312v050000p.pdf, 1996.

[20] 3GPP Support Office. 3rd generation partnership project - technical specification group core network - location management procedures. [Online] http://www.qtc.jp/3GPP/Specs/23012-520.pdf, 2003.

[21] Jie Li and Hsiao-Hwa Chen. Mobility support for ip-based networks. *IEEE Communications Magazine*, 43(10):127–132, 2005.

[22] Adrian Vasilache, Jie Li, and Hisao Kameda. Threshold-based load balancing for multiple home agents in mobile ip networks. *Telecommunication Systems*, 22(1-4):11–31, 2003.

[23] Jie Li, Yi Pan, and Xiaohua Jia. Analysis of dynamic location management for pcs networks. *IEEE Transactions on Vehicular Technology*, 51(5):1109–1119, 2002.

[24] Jie Li and Yongping Pan. A dynamic hlr location management scheme for pcs networks. In *Proceedings of The 2014 IEEE Conference on Computer Communications (IEEE INFOCOM 2004)*, volume 1. IEEE, 2004.

[25] Woongsup Lee and Dong-Ho Cho. Enhanced group handover scheme in multiaccess networks. *IEEE Transactions on Vehicular Technology*, 60(5):2389–2395, 2011.

[26] Kalyanmoy Deb and Himanshu Gupta. Searching for robust pareto-optimal solutions in multi-objective optimization. *Lecture Notes in Computer Science*, 3410:150–164, 2005.

[27] Ralph E Steuer. An overview in graphs of multiple objective programming. In *Evolutionary Multi-Criterion Optimization*, pages 41–51. Springer, 2001.

[28] Kalyanmoy Deb. *Multi-objective optimization using evolutionary algorithms*, volume 16. John Wiley & Sons, 2001.

[29] Lofti Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, 8(1):59–60, 1963.

[30] Chris W Liu. Modeling 3g/wcdma/hsdpa handset transmit system. *Microwave Product Digest*, 2007.

[31] Trichy Venkataraman Krishnamurthy and Rajaneesh Shetty. *4G: Deployment Strategies and Operational Implications*. Apress, 2014.

[32] Jie Li, Atsushi Kubota, and Hisao Kameda. Location management for pcs networks with consideration of mobility patterns. In *Proceedings of The 2005 IEEE Conference on Computer Communications (IEEE INFOCOM 2005)*, volume 2, pages 1207–1216. IEEE, 2005.

[33] A Kubota, Jie Li, and Hisao Kameda. Optimal location management of mobile terminals with consideration of moving patterns. *(in Japanese), IPSJ (Information Processing Society of Japan) Journal*, 43(12):3922–3930, 2002.

Li Qiang is a Ph.D. candidate at the Graduate School of Systems and Information Engineering, University of Tsukuba, Japan. She received the B.E. degree from Anhui University, China, in 2009, and M.E. Degree from University of Science and Technology of China, China, in 2012. Her main research interests include mobility management and heterogeneous wireless networks.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2016.2551042, IEEE Transactions on Emerging Topics in Computing

Jie Li (M94-SM04) received the B.E. degree in computer science from Zhejiang University, Hangzhou, China, the M.E. degree in electronic engineering and communication systems from China Academy of Posts and Telecommunications, Beijing, China. He received the Dr. Eng. degree from the University of Electro-Communications, Tokyo, Japan. He is with Faculty of Engineering, Information and Systems, University of Tsukuba, Japan, where he is a Professor. He has been a visiting professor in Yale University, USA, and in Inria, France. His current research interests are in mobile distributed computing and networking, big data and cloud computing, IoT, OS, modeling and performance evaluation of information systems. He is a senior member of IEEE and ACM and a member of IPSJ (Information Processing Society of Japan). He is the Chair of Technical Sub-Committee on Big Data (TSCBD), IEEE Communications Society. He has served as a secretary for Study Group on System Evaluation of IPSJ and on several editorial boards for the international Journals, and on Steering Committees of the SIG of System EVAluation (EVA) of IPSJ, the SIG of DataBase System (DBS) of IPSJ, and the SIG of MoBiLe computing and ubiquitous communications of IPSJ. He has also served on the program committees for several international conferences such as IEEE INFOCOM, IEEE GLOBECOM, and IEEE MASS.

Corinne Touati is a research scientist at the Inria, France. She received M.S from Telecom SudParis, France, and her Ph.D. from University of Nice Sophia-Antipolis, France. Her research interests include performance evaluation, continuous and stochastic optimization, game theory and learning theory in communication networks and distributed systems.