

Research Article

Data-Driven Handover Optimization in Next Generation Mobile Communication Networks

Po-Chiang Lin, Lionel F. Gonzalez Casanova, and Bakary K. S. Fatty

Department of Communications Engineering, Yuan Ze University, Taoyuan 320, Taiwan

Correspondence should be addressed to Po-Chiang Lin; pclin@saturn.yzu.edu.tw

Received 1 April 2016; Revised 3 July 2016; Accepted 5 July 2016

Academic Editor: Young-June Choi

Copyright © 2016 Po-Chiang Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Network densification is regarded as one of the important ingredients to increase capacity for next generation mobile communication networks. However, it also leads to mobility problems since users are more likely to hand over to another cell in dense or even ultradense mobile communication networks. Therefore, supporting seamless and robust connectivity through such networks becomes a very important issue. In this paper, we investigate handover (HO) optimization in next generation mobile communication networks. We propose a data-driven handover optimization (DHO) approach, which aims to mitigate mobility problems including too-late HO, too-early HO, HO to wrong cell, ping-pong HO, and unnecessary HO. The key performance indicator (KPI) is defined as the weighted average of the ratios of these mobility problems. The DHO approach collects data from the mobile communication measurement results and provides a model to estimate the relationship between the KPI and features from the collected dataset. Based on the model, the handover parameters, including the handover margin and time-to-trigger, are optimized to minimize the KPI. Simulation results show that the proposed DHO approach could effectively mitigate mobility problems.

1. Introduction

The first generation (1G) mobile communication systems enabled the release from traditional wireline place-to-place communications to wireless person-to-person communications. As technologies and demands evolve, now the fifth generation (5G) mobile communication systems aim to connect anything to anything all over the world.

One of the engineering requirements for 5G is to achieve 1000x data rate [1]. There are three categories of technologies that would be combined together to achieve the 1000x capacity gain:

- (1) Network densification and traffic offloading to improve the area spectral efficiency.
- (2) Millimeter wave (mmWave) technologies and access to unlicensed spectrum to increase bandwidth.
- (3) Massive Multiple-Input and Multiple-Output (massive MIMO), Multiuser MIMO (MU-MIMO), and Coordinated Multipoint (CoMP) to increase spectral efficiency.

Among these categories of technologies, network densification and traffic offloading are expected to provide the majority (40x to 50x) of the required capacity gain. Network densification is a straightforward but effective method to increase the network capacity by making cell size smaller. The advantages of cell shrinking include frequency reuse and reduction of resource competition among users within each cell.

However, network densification leads to mobility problems since users are more likely to hand over to another cell in dense or even ultradense mobile communication networks. Therefore, supporting seamless and robust connectivity through such networks becomes a very important issue. Moreover, as the number of base stations increases, the installation, configuration, and maintenance efforts also increase. Mitigating these efforts to decrease the capital expenditure (CAPEX) and operational expenditure (OPEX) is also a critical issue [2].

In next generation mobile communication networks, optimizing the handover (HO) parameters to improve the system performance is critical. The objective of handover

optimization, an important part of the Self-Organizing Network (SON), is to provide fast and seamless handover from one cell to another while simultaneously keeping network management simple [3, 4]. The main goals of handover optimization include minimizing call drops, minimizing radio link failures (RLF), minimizing unnecessary handovers, and minimizing idle mode problems.

In this paper, we investigate the handover optimization problem in next generation mobile communication networks. We propose a data-driven handover optimization (DHO) approach, which aims to mitigate mobility problems. The DHO approach collects data from the mobile communication measurement results and provides a model to estimate the relationship between the key performance indicator (KPI) and features from the collected dataset. Based on the model, the handover parameters, which include the handover margin (HOM) and time-to-trigger (TTT), are optimized to minimize the KPI, which is the weighted average of various mobility problem ratios. Simulation results show that the proposed DHO approach could effectively mitigate mobility problems.

The major contributions of this paper are threefold:

- (1) We consider several features, including the location, moving speed, and moving direction of mobile stations, HOM, and TTT. The proposed approach is thus context-aware and adaptive to mobile communication environments.
- (2) We classify the mobility problems into five different types and design the mobility problem identification mechanism based on the combination of three conditional tests.
- (3) We adopt the neural network model to estimate the KPI function and to optimize the handover parameters HOM and TTT.

The rest of this paper is organized as follows. In Section 2 we present a review of related work. The HO problem formulation is described in Section 3. In Section 4, we provide the detailed description of the proposed DHO approach. Performance evaluation and discussions are given in Section 5. Finally, Section 6 concludes this paper.

2. Related Work

There exist several works for the handover optimization in the long-term evolution (LTE) mobile communication networks. Jansen et al. proposed a self-optimizing algorithm that tunes the handover parameters of an LTE base station in order to diminish negative effects, such as call dropping and handover failures [5]. Their algorithm picks the optimal hysteresis and time-to-trigger combination for the current network status. The authors also proposed a weighted performance based handover optimization algorithm (WPHPO) that tunes the hysteresis and time-to-trigger in iterative steps [6]. Li et al. proposed a dynamic hysteresis-adjusting (DHA) method, which uses the network-allowed maximum RLF ratio as a key indicator [7]. Lobinger et al. investigated the coordination of handover parameter optimization and load balance in LTE

SON [8]. Capdevielle et al. investigated the joint interference management and handover optimization in LTE small cells networks [9]. They focus on two challenges for LTE small cell deployment, intercell interference management, and mobility management.

Several works focus on the heterogeneous/inter-RAT handover. Ali and Saquib investigated cellular/wireless LAN handover analysis for different mobility models in order to exploit the heterogeneity of the wireless environment [10]. They provided a medium selection method for vertical handover. Awada et al. investigated the cell-pair specific optimization of the inter-RAT handover parameters in SON [11]. They provided the configuration paradigms for the inter-RAT handover thresholds. Giacomini and Agarwal use QoS reputation and GM(1, 1) prediction to make decision on vertical handover [12]. They built on a novel reputation based vertical handover decision rating system in the handover decision making progress. López-Pérez et al. characterized the relation between handover failure and ping-pong rates in a heterogeneous network scenario [13]. Rath and Panwar proposed a new prefetch-based fast handover procedure that is designed to overcome the higher latency caused by the use of the public internet to connect the femtocell base station with the mobile core network [14]. Xenakis et al. proposed a novel handover decision policy for the two-tier LTE network with the goal to reduce the power consumption at mobile stations [15]. The proposed method can adapt to the handover hysteresis margin with respect to the prescribed SINR target and measurement results.

Some works introduced new concepts and architectures to enhance user mobility in mobile communication networks. Taleb et al. investigated the support of highly mobile users [16]. They introduced a data anchor gateway relocation method based on the information, including user mobility, history information, and user activity patterns, and proposed a handover management policy to select a target base station. Imran et al. investigated how to empower SON with big data for enabling 5G in [17]. They first characterized big data in next generation mobile communication networks and then provided challenges in SON for enabling 5G, including underutilized intelligence, need for self-coordination, need for more transparent SON, scalability, energy efficiency, and need for a paradigm shift from reactive to proactive SON. The authors also proposed a framework for big data empowered SON (BSON) which takes on the challenges mentioned above. In this paper we leverage from their basic idea and investigate the HO problem which is one of the most important issues in SON.

3. Problem Formulation

We formulate the HO problem as follows. Suppose that $\mathbf{X} = [\mathbf{X}_D, \mathbf{X}_M]$ is a feature vector which includes two types of variables, \mathbf{X}_D and \mathbf{X}_M . To be specific, \mathbf{X}_M denotes the measurement result, which includes the location of MS, moving speed of MS, moving direction of MS, and other useful counters and flags described in Section 4.1. \mathbf{X}_D denotes the decision variables of the HO problem, which is the handover parameter vector with two elements, HOM and TTT.

The details of these two handover parameters are described as follows.

(i) *Handover Margin (HOM) (Also Known as Hysteresis)*. HOM is defined as the threshold of the difference in Signal to Interference plus Noise Ratio (SINR) between the serving and the target cells. When the SINR from the target cell is HOM better than that from the serving cell, the HO condition is satisfied.

(ii) *Time-to-Trigger (TTT)*. An HO request would not be sent immediately when the corresponding HO condition mentioned above is satisfied. Instead, the HO condition has to be fulfilled for a certain period denoted as TTT. When the HO condition keeps satisfied for TTT, an HO request is triggered.

The proper values of HOM and TTT are critical to seamless connectivity. Large HOM and TTT values lead to more stable behavior, but they may delay the HO decisions unnecessarily which may cause problems. On the other hand, small HOM and TTT values avoid long delay to trigger HO request, but they may cause ping-pong HO and unnecessary HO.

The HO problem is formulated as

$$\mathbf{X}_{D,opt} = \arg \min_{\mathbf{X}_D} Y, \quad (1)$$

where Y is the corresponding key performance indicator (KPI), which depends on the feature vector $\mathbf{X} = [\mathbf{X}_D, \mathbf{X}_M]$. The KPI should be designed to measure the mitigation of mobility problems.

There are four main challenges in the HO problem:

- (1) There are various kinds of mobility problems in next generation mobile communication networks [18]. Identifying the mobility problems would be the first priority in the HO problem.
- (2) The design of the KPI needs comprehensive and careful considerations for these mobility problems.
- (3) The KPI Y can be formulated as a function of \mathbf{X} ; that is, $Y = f(\mathbf{X})$. However, the exact form of $f(\mathbf{X})$ is unknown. Therefore, one of the challenges we take on in this work is to provide a good estimation $\hat{f}(\mathbf{X})$ to the function $f(\mathbf{X})$. With a good estimate $\hat{f}(\mathbf{X})$ we can make predictions of the KPI Y at some points $\mathbf{X} = \mathbf{x}$, where \mathbf{x} is a vector of some specific values of \mathbf{X} .
- (4) The decision variables \mathbf{X}_D consist of two elements, HOM and TTT. Mitigating the mobility problems by joint-optimizing HOM and TTT would be another challenge to the HO problem.

In Section 4 we describe the proposed DHO approach to deal with these main challenges.

4. Proposed DHO Approach

In this section, we provide the detailed description of the proposed data-driven handover optimization (DHO) approach. Figure 1 shows the system architecture of the proposed DHO

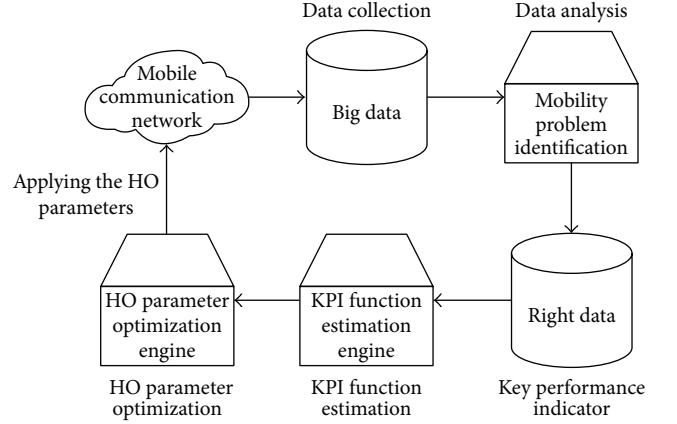


FIGURE 1: System architecture of the proposed DHO approach.

approach. First, measurement results are collected from the mobile communication network. The dataset collected in the database includes the feature vector described in Section 3 and various counters and flags that are used to identify the mobility problems. The details of these counters and flags are provided in Section 4.1. The DHO approach then analyzes and processes the data to calculate the mobility problem ratio (MPR) of each mobility problem and obtain the KPI, which is saved in the database denoted as “right data” in Figure 1. The KPI function estimation engine uses the processed “right data” to provide the KPI function estimate. After that, the HO parameter optimization engine uses the KPI function estimate to optimize the HO parameters including HOM and TTT. Finally, the optimal values of HOM and TTT are applied to corresponding base stations.

In the following paragraphs, we provide the detailed description of each building block in Figure 1.

4.1. Mobility Problem Identification. There would be five different mobility problems (MPs) in next generation mobile communication networks.

(1) *Too-Late HO*. Too-late HO trigger timing leads to low SINR level of the serving cell. A radio link failure (RLF) between the mobile station (MS) and the serving base station (BS) occurs, and then the MS try to reestablish a connection to a nearby BS.

(2) *Too-Early HO*. Too-early HO trigger timing leads to a low SINR level of the target cell. An RLF between the MS and the target BS occurs, and then the MS try to reestablish a connection to the original serving BS.

(3) *HO to Wrong Cell*. HO from a source cell A to a target cell B which provides unstable SINR leads to an RLF. At the same time, if there is another suitable cell C for the MS to reconnect to after the RLF, the MS reestablishes connection with cell C that is neither the serving cell A nor the target cell B. In this situation, instead of the failed HO from cell A to cell B, HO from cell A directly to cell C would be a better decision.

TABLE 1: Mobility problem identification.

Mobility problem	Close to last HO?	RLF status?	New target = previous serving?
Too-late HO	N	Y	Don't care
Too-early HO	Y	Y	Y
HO to wrong cell	Y	Y	N
Ping-pong HO	Y	N	Y
Unnecessary HO	Y	N	N

(4) *Ping-Pong HO*. Cell A hands over a MS to cell B, and cell B hands over the same MS back to cell A shortly after. Since data transmission is temporarily blocked during the connection transferring to the target cell, these two HOs, despite both being successful, should be avoided.

(5) *Unnecessary HO*. Cell A hands over a MS to cell B and cell B hands over the same MS to another cell C shortly after. Those two HOs, despite both being successful, can be combined into one HO from cell A directly to cell C in order to avoid unnecessary blocking of data transmission during the connection transferring time.

In order to identify the mobility problems mentioned above, we design the mobility problem identification criteria, as Table 1 shows. The mobility problem identification criteria depend on three conditional tests:

(1) Close to last HO?

Whether the occurrence time of the current HO event is close to that of the previous HO event?

(2) RLF status?

Is there an RLF for the current connection?

(3) New target = previous serving?

Whether the new target cell is the same as the previous serving one?

Table 1 shows that each of the five mobility problems would be identified based on the test results of the three conditional tests mentioned above.

In the proposed DHO approach, we design the following counters and flags to identify the mobility problems:

- (i) lastHoCounter: to indicate the time interval since last HO.
- (ii) lastHoThreshold: to define a threshold value, combined with lastHoCounter to identify whether the occurrence time of the current HO event is close to that of the previous HO event.
- (iii) rlfStatus: to indicate whether an RLF occurs for the current connection.
- (iv) newTagertCellId: to indicate the new target cell ID.
- (v) previousServingCellId: to indicate the previous serving cell ID, combined with the newTagertCellId to indicate whether the new target cell is the same as the previous serving one.

4.2. *Design of KPI*. In the HO problem, the KPI Y is defined as the weighted average of the various mobility problem ratios (MPRs). There are five different kinds of mobility problems. The KPI Y is calculated as

$$Y = \frac{\sum_{k=1}^5 w_k R_k(\mathbf{X})}{\sum_{k=1}^5 w_k}, \quad (2)$$

where w_k is the weight of the mobility problem k , which is determined by mobile network operators, and $R_k(\mathbf{X})$ is the MPR of the mobility problem k , which is a function of \mathbf{X} . MPR R_k is defined as the probability of the event where the mobility problem (MP) k happens. Consider

$$R_k(\mathbf{X}) = \Pr(\text{MP}_k) = \frac{N_k}{N_{\text{total}}}, \quad (3)$$

where N_k denotes the number of MP_k , and N_{total} denotes the total number of HO. Each BS uses a sliding time window to collect the statistics of N_k and N_{total} in order to obtain each MPR and the KPI Y .

4.3. *KPI Function Estimation*. In order to solve the HO problem, as (1) shows, some previous solutions focus on analyzing and obtaining the KPI function $Y = f(\mathbf{X})$. When the analytical formula $f(\mathbf{X})$ is known, solving the equation $\partial f / \partial \mathbf{X}_D = 0$ would get the optimal handover parameters HOM and TTT. However, note that the function f consists of complex joint effects of the feature vector $\mathbf{X} = [\mathbf{X}_D, \mathbf{X}_M]$. Under these complex effects, it is difficult to analyze and obtain an accurate KPI function.

In this paper we use machine learning techniques to estimate the KPI function. To be specific, we use the multilayer perception (MLP), which is one of the most popular neural networks, to perform the KPI function estimation [19].

The MLP consists of an input layer, one or more hidden layers, and an output layer. In this paper we adopt the MLP with one hidden layer consisting of four hidden neurons. The feature vector \mathbf{X} is applied to the input layer and propagates through the network layer by layer from left to right. Each arrow in this figure represents an adjustable synaptic weight. From our comprehensive experiments, this architecture is empirically sufficient to model the relationship between KPI and the feature vector, whereas the other experiments with more hidden neurons provide similar results. A simulation result about the mean-square-error (MSE) of the KPI function estimation with various numbers of hidden neurons is shown in Figure 2. The epoch limitation is set as 1000. This figure shows that four hidden neurons are sufficient to improve the performance of modeling to the degree of $1E-10$, whereas the performance of two neurons only reaches the degree of $1E-6$. Applying more hidden neurons does not significantly improve the MSE.

In MLP, the backpropagation (BP) algorithm is used to adjust the neural network parameters in order to minimize the error between the target (measured) KPI and the response from the function estimate \hat{f} . The detailed operation of the BP algorithm includes the forward pass and the backward pass. We provide the detailed descriptions of the forward pass and the backward pass in the following paragraphs.

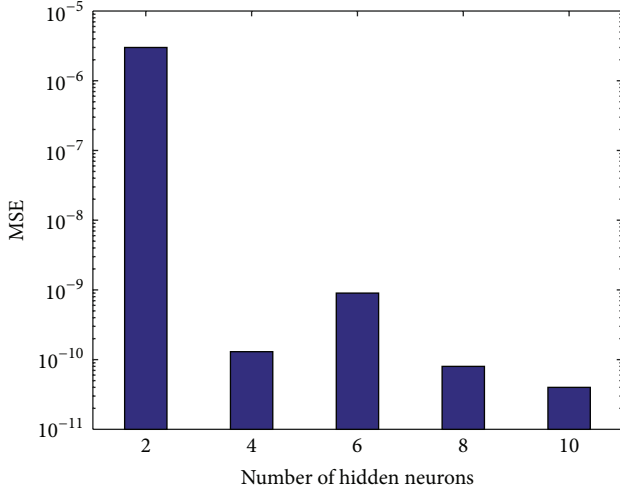


FIGURE 2: MSE of the KPI function estimation with various number of hidden neurons.

4.3.1. Forward Pass. The purpose of the forward pass is to get the actual response of the MLP in accordance with the specified input signal. In the forward pass the feature vector $\mathbf{X}(n)$ at iteration n is applied to the input layer and propagates through the network layer by layer from left to right. The top part of Figure 3 shows the signal flow of the forward pass. $X_1(n)$ denotes the HOM, and $X_2(n)$ denotes the TTT. $X_3(n)$ to $X_m(n)$ denote the other features. When the signal propagates through a synapse, its value is multiplied by the synaptic weight. The input of a neuron is the synaptic weighted sum of the output of its previous layer. For example, the input of the j th neuron in the l th layer is

$$v_j^{(l)}(n) = \sum_{i=0}^{m_{l-1}} w_{ji}^{(l)}(n) y_i^{(l-1)}(n), \quad (4)$$

where $i = 0$ represents the bias term. The input of this neuron is applied to the activation function $\varphi(\cdot)$. The activation function should be differentiable everywhere. We adopt the commonly used sigmoid function as

$$y_j^{(l)}(n) = \varphi(v_j^{(l)}(n)) = \frac{1}{1 + \exp(-v_j^{(l)}(n))}. \quad (5)$$

The output of this neuron propagates through the network to the next layer, and the same operation is performed again. Finally the signal aggregates in the output layer, and the estimated KPI function $\tilde{Y}(n)$ is obtained. During the forward pass the synaptic weights are all fixed.

4.3.2. Backward Pass. The purpose of the backward pass is to adjust the synaptic weights in order to minimize the error between the actual response $\tilde{Y}(n)$ and the desired response $Y(n)$. The error signal $e(n)$ is defined by

$$e(n) = Y(n) - \tilde{Y}(n). \quad (6)$$

The instantaneous square error $E(n)$ is defined as

$$E(n) = \frac{1}{2} e^2(n) = \frac{1}{2} (Y(n) - \tilde{Y}(n))^2. \quad (7)$$

The goal of the back propagation algorithm is to minimize $E(n)$. To achieve this goal, the back propagation algorithm applies a respective correction $\Delta\omega_{ji}^{(l)}(n)$ to every synaptic weight $\omega_{ji}^{(l)}(n)$. The correction $\Delta\omega_{ji}^{(l)}(n)$ is designed to be proportional to the gradient $\partial E(n)/\partial\omega_{ji}^{(l)}(n)$; that is,

$$\Delta\omega_{ji}^{(l)}(n) = -\eta_a \frac{\partial E(n)}{\partial\omega_{ji}^{(l)}(n)}. \quad (8)$$

The minus sign comes from the fact that lower $E(n)$ is preferred. η_a is the adaptive learning rate proposed by Behera et al. in [20], which is expressed as

$$\eta_a = \mu \frac{\|\tilde{\mathbf{y}}\|^2}{\|\mathbf{J}^T \tilde{\mathbf{y}}\|^2}, \quad (9)$$

where $\tilde{\mathbf{y}} = Y(n) - \tilde{Y}(n)$ and $\mathbf{J} = \partial\tilde{Y}/\partial\mathbf{W}_l$. While considering the weight correction for $l = 2$, that is, the output layer, according to the chain rule of calculus, the gradient $\partial E(n)/\partial\omega_{1i}^{(2)}(n)$ can be written as

$$\frac{\partial E(n)}{\partial\omega_{1i}^{(2)}(n)} = \frac{\partial E(n)}{\partial e(n)} \frac{\partial e(n)}{\partial\tilde{Y}(n)} \frac{\partial\tilde{Y}(n)}{\partial v_1^{(2)}(n)} \frac{\partial v_1^{(2)}(n)}{\partial\omega_{1i}^{(2)}(n)}. \quad (10)$$

Applying (4)–(7) to the four partial derivatives at the right of the equal sign in (10), it can be rewritten as

$$\begin{aligned} \frac{\partial E(n)}{\partial\omega_{1i}^{(2)}(n)} &= -e(n) \varphi'(v_1^{(2)}(n)) y_i^{(1)}(n) \\ &= -\delta_1^{(2)}(n) y_i^{(1)}(n), \end{aligned} \quad (11)$$

where $\delta_j^{(l)}(n)$ is the local gradient which is defined as

$$\delta_j^{(l)}(n) = -\frac{\partial E(n)}{\partial v_j^{(l)}(n)}. \quad (12)$$

While considering the weight correction for $l = 1$, that is, the hidden layer, by iterative backpropagation and chain rule manners, the local gradient $\delta_j^{(1)}(n)$ is

$$\delta_j^{(1)}(n) = \varphi'(v_j^{(1)}(n)) \delta_1^{(2)}(n) \omega_{1j}^{(2)}(n). \quad (13)$$

The derivation of the local gradients at each layer is shown in the bottom part of Figure 3. Finally, every synaptic weight is corrected following this formula:

$$\omega_{ji}^{(l)}(n+1) = \omega_{ji}^{(l)}(n) + \Delta\omega_{ji}^{(l)}(n). \quad (14)$$

The forward pass and the backward pass are executed iteratively until the square error $E(n)$ reaches some satisfactory level or when the number of epoch exceeds some threshold. After that, the handover parameter optimization is executed.

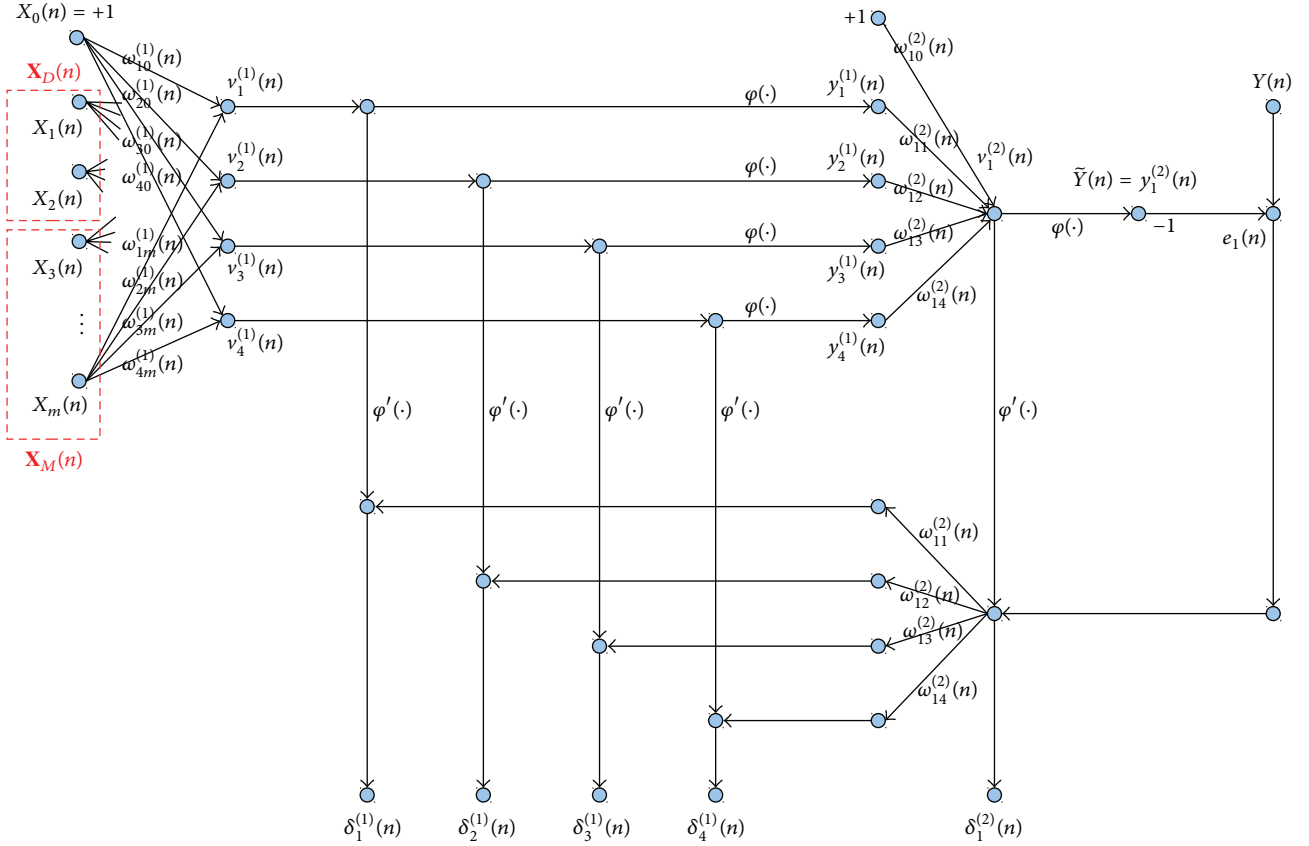


FIGURE 3: The forward pass and the backward pass of the BP algorithm.

4.4. Handover Parameter Optimization. In this paper, the handover optimization problem is solved by the gradient descent algorithm [21]. Here the handover parameters \mathbf{X}_D , including HOM and TTT, are adjusted adaptively. To find the local minimum of the KPI function $Y \approx \hat{f}(\mathbf{X})$ using gradient descent, one takes steps proportional to the gradient. Suppose that at the n th time of adjustment, the handover parameters are $\mathbf{X}_D(n)$ and the KPI is $Y(n)$. At the next time of adjustment, the handover parameters \mathbf{X}_D are set as

$$\mathbf{X}_D(n+1) = \mathbf{X}_D(n) + \Delta\mathbf{X}_D(n). \quad (15)$$

$\Delta\mathbf{X}_D(n)$ depends on the gradient of estimated KPI $Y(n)$ with respect to $\mathbf{X}_D(n)$; that is,

$$\Delta\mathbf{X}_D(n) = \mu \frac{\partial Y(n)}{\partial \mathbf{X}_D(n)}, \quad (16)$$

where μ is the constant adjustment rate.

How to obtain the gradient $\partial Y / \partial \mathbf{X}_D$ is a problem left. The original BP algorithm described in Section 4.3 only consists of two passes: the forward pass and the backward pass. In order to obtain the gradient $\partial Y / \partial \mathbf{X}_D$, the proposed DHO approach modifies the original BP algorithm to add the third pass: the adjustment pass. We provide the detailed description of the adjustment pass in the following paragraphs.

The purpose of the adjustment pass is to obtain the estimated gradient $\partial \tilde{Y}(n) / \partial \mathbf{X}_D(n)$ which is the key to adjust the handover parameter in order to minimize the KPI. The synaptic weights that have been adjusted well in the backward pass in accordance with (14) are set as fixed in the adjustment pass. In the following we show how to obtain the estimated gradient $\partial \tilde{Y}(n) / \partial \mathbf{X}_D(n)$. The signal flow of the adjustment pass is depicted in Figure 4.

First of all, we define the local gradient $\lambda_j^{(l)}(n)$ of the adjustment pass as

$$\lambda_j^{(l)}(n) = \frac{\partial \tilde{Y}(n)}{\partial v_j^{(l)}(n)}. \quad (17)$$

This local gradient $\lambda_j^{(l)}(n)$ is similar to the local gradient $\delta_j^{(l)}(n)$ of the backward pass which is defined in (12), except that the target of the partial derivative is changed to $\tilde{Y}(n)$. While considering the output layer, the local gradient $\lambda_1^{(2)}(n)$ is

$$\lambda_1^{(2)}(n) = \frac{\partial \tilde{Y}(n)}{\partial v_1^{(2)}(n)} = \varphi'(v_1^{(2)}(n)). \quad (18)$$

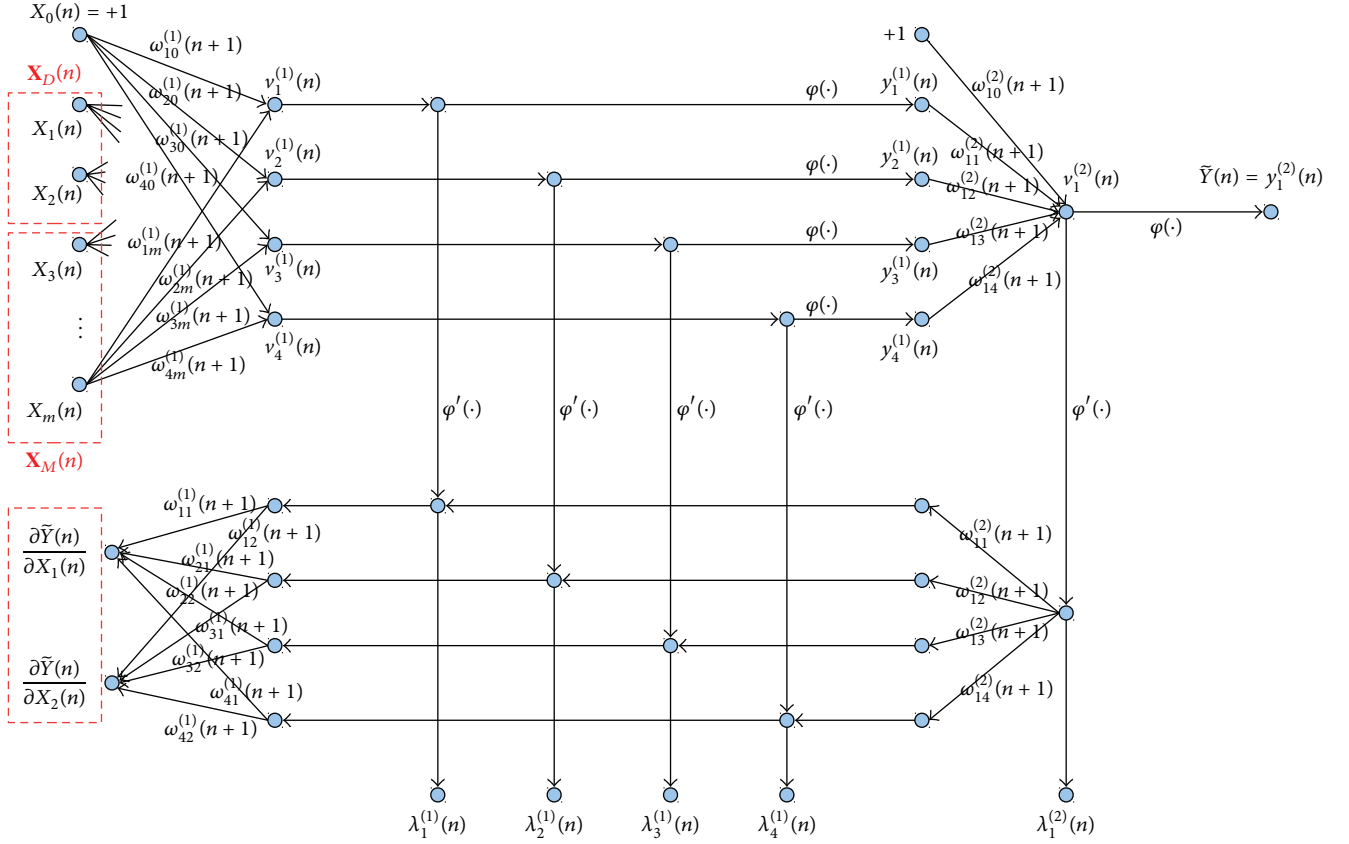


FIGURE 4: Adjustment pass of the proposed DHO approach.

While considering the hidden layer, the local gradient $\lambda_j^{(1)}(n)$ is

$$\begin{aligned} \lambda_j^{(1)}(n) &= \frac{\partial \tilde{Y}(n)}{\partial v_j^{(1)}(n)} = \frac{\partial \tilde{Y}(n)}{\partial v_1^{(2)}(n)} \frac{\partial v_1^{(2)}(n)}{\partial y_j^{(1)}(n)} \frac{\partial y_j^{(1)}(n)}{\partial v_j^{(1)}(n)} \\ &= \lambda_1^{(2)}(n) \omega_{1j}^{(2)}(n+1) \varphi'(v_j^{(1)}(n)). \end{aligned} \quad (19)$$

Using the results from (18) and (19), the gradient can be written as follows:

$$\frac{\partial \tilde{Y}(n)}{\partial X_i(n)} = \frac{\partial \tilde{Y}(n)}{\partial v_1^{(2)}(n)} \frac{\partial v_1^{(2)}(n)}{\partial X_i(n)} = \lambda_1^{(2)}(n) \frac{\partial v_1^{(2)}(n)}{\partial X_i(n)}, \quad (20)$$

where the last equality comes from (18). Since

$$v_1^{(2)}(n) = \sum_{j=0}^4 \omega_{1j}^{(2)}(n+1) y_j^{(1)}(n), \quad (21)$$

the second term at the most right of (20) can be written as

$$\begin{aligned} \frac{\partial v_1^{(2)}(n)}{\partial X_i(n)} &= \sum_{j=0}^4 \omega_{1j}^{(2)}(n+1) \frac{\partial y_j^{(1)}(n)}{\partial X_i(n)} \\ &= \sum_{j=0}^4 \omega_{1j}^{(2)}(n+1) \frac{\partial y_j^{(1)}(n)}{\partial v_j^{(1)}(n)} \frac{\partial v_j^{(1)}(n)}{\partial X_i(n)} \\ &= \sum_{j=1}^4 \omega_{1j}^{(2)}(n+1) \frac{\partial y_j^{(1)}(n)}{\partial v_j^{(1)}(n)} \frac{\partial v_j^{(1)}(n)}{\partial X_i(n)} \\ &= \sum_{j=1}^4 \omega_{1j}^{(2)}(n+1) \varphi'(v_j^{(1)}(n)) \omega_{j1}^{(1)}(n+1), \end{aligned} \quad (22)$$

where the third equality comes from the fact that $y_0^{(1)}(n)$ is always equal to 1 and its partial derivative can be ignored from the summation. Combining (19) and (22) into (20), the gradient $\partial \tilde{Y}(n)/\partial X_i(n)$ is obtained as

$$\frac{\partial \tilde{Y}(n)}{\partial X_i(n)} = \sum_{j=1}^4 \lambda_j^{(1)}(n) \omega_{j1}^{(1)}(n+1). \quad (23)$$

The derivation of the local gradients at each layer and the gradient $\partial \tilde{Y}(n)/\partial X_i(n)$ is depicted in Figure 4. Based on the

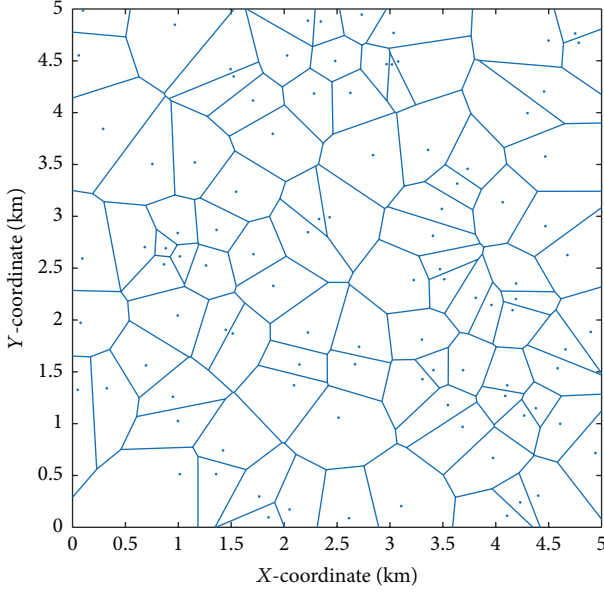


FIGURE 5: Topology of the simulated mobile communication network.

result from (23), the handover parameter X_i is adjusted as (15) and (16) show.

In order to cope with the time-varying channel conditions, we use online training and operate the three passes iteratively to model the instantaneous relationship $Y = f(\mathbf{X})$ and optimize the handover parameters, HOM and TTT.

5. Performance Evaluation and Discussion

In this section we describe the simulation environment, simulation results, and discussions.

In our simulation, a mobile communication network which consists of 100 small cells is built in a $5\text{ km} \times 5\text{ km}$ area. Figure 5 shows the coordinates of these small cell base stations. Each cell has 5 MHz bandwidth with 25 resource blocks and 2 GHz carrier frequency. Each resource block consists of 12 subcarriers of size 15 kHz each. A time slot is 0.5 ms in duration and the transmission time interval (TTI) is 1 ms.

In the simulation, 100 MSs are uniformly distributed over the area with random initialized positions. The mobility trace of these 100 MSs is generated by using the Manhattan mobility model, which is widely used for urban areas and suitable to implement realistic simulations [22–24]. In the Manhattan mobility model, MSs move in either horizontal or vertical directions on an urban map. At each intersection, an MS employs a probabilistic approach to choose the direction of its movements. The probabilities of going straight, turning left, and turning right are 0.5, 0.25, and 0.25, respectively. The blue line in Figure 5 shows an illustrative moving path of an MS in the mobile communication network. While considering the KPI, the weights of mobility problems are set to be equal.

An RLF will occur if the SINR level of the serving cell drops below -6.5 dB before handover execution is

completed [18]. We set this as our SINR threshold value such that if the SINR from the source BS is lower than the threshold, an RLF will occur. The simulation time is set as 5000 seconds. The operating point with 0 dB HOM and 0 s TTT is the starting operating point for all cells in the network.

In this paper we use LTE-Sim, an open source framework to simulate LTE networks, as our basic simulation tool [25]. The design and development of LTE-Sim support the following aspects:

- (i) The Evolved Universal Terrestrial Radio Access (E-UTRAN) and the Evolved Packet System (EPS).
- (ii) Single and heterogeneous multicell environments.
- (iii) Multiusers environment.
- (iv) User mobility.
- (v) Handover procedures.
- (vi) Frequency reuse techniques.
- (vii) Quality-of-service (QoS) management.
- (viii) Scheduling strategies.

However, LTE-Sim still lacks some important functionalities and features that are required in this work. Therefore, we modified LTE-Sim from several aspects. Our modification to LTE-Sim includes the following:

- (1) Add:
 - (i) radio link failure,
 - (ii) handover failure,
 - (iii) HOM and TTT,
 - (iv) proposed DHO approach.
- (2) Modify:
 - (i) mobility pattern: using the seed to generate the same sequence of random numbers each time we run the simulation.

The simulation parameter values are shown in Table 2 unless otherwise specified.

Figure 6 shows the simulation result when the TX power is set as 20 dBm. The horizontal axis denotes the mobility problem index. Indexes 1 to 5 mean the mobility problem of too-late HO, too-early HO, HO to wrong cell, ping-pong HO, and unnecessary HO, respectively. Index 6 means the KPI, which is the weighted average of the five MPRs. The vertical axis means the corresponding values of MPR or KPI. The blue bars show the performance for a static HO parameter setting with 0 dB HOM and 0 s TTT and the green ones for another static setting with 10 dB HOM and 5.12 s TTT. The yellow bars show the performance of the proposed DHO approach. The results shown in Figure 6 validate that the proposed DHO approach can effectively mitigate the mobility problems and achieve better MPRs and KPI.

Figure 7 shows the simulation result when the TX power is set as 43 dBm. Since the TX power is higher than the one in the previous scenario in Figure 6, it is less likely that an

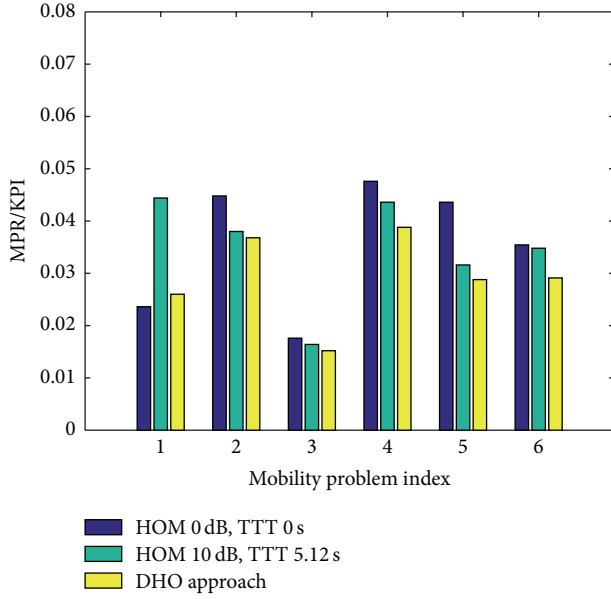


FIGURE 6: Simulation result when the TX power is set as 20 dBm.

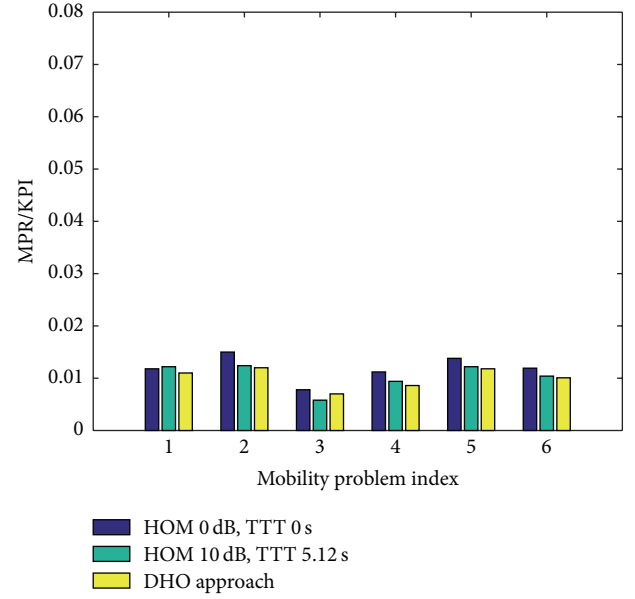


FIGURE 7: Simulation result when the TX power is set as 43 dBm.

TABLE 2: Simulation parameters.

Parameter	Values
Number of small cell base stations	100
Number of mobile stations	100
eNB Tx power	15 dBm/20 dBm/43 dBm
Carrier frequency	2 GHz
MS speed	30 km/h
Mobility model	Manhattan
TTT	1 ms
Subcarrier spacing	15 KHz
Resource block	180 KHz
Super frame time	10 ms
Noise figure	2.5
Noise spectral density	-174 dBm
SINR threshold	-6.5 dB
Simulation time	5000 s
Handover delay	30 ms
System bandwidth	5 MHz 25 RBs/TTI

RLF occurs. Therefore, the MPRs and KPI in Figure 7 are lower than those in Figure 6. Moreover, the proposed DHO approach can still effectively mitigate the mobility problems and provide better connectivity to MSs.

Figure 8 shows the simulation result when the TX power is set as 15 dBm. Since the TX power is lower than the one in the previous scenario in Figure 6, it is more likely that an RLF occurs. Therefore, the MPRs and KPI in Figure 8 are higher than those in Figure 6. Moreover, the proposed DHO approach can still effectively mitigate the mobility problems and provide better connectivity to MSs.

We compare the KPI improvement under the various values of TX power. The KPI improvement is calculated as the difference between the initial-case KPI (HOM 0 dB,

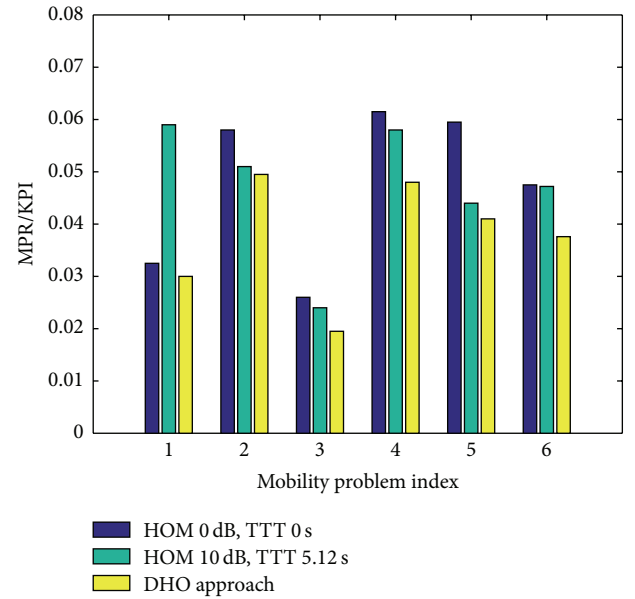


FIGURE 8: Simulation result when the TX power is set as 15 dBm.

TTT 0 s) and the KPI of the proposed DHO method, divided by the initial-case KPI. Figure 9 shows that the KPI improvement ranges from 15% to 20%. It also shows that the KPI improvement is more significant when the TX power is lower; that is, the transmission range of the base station is smaller.

6. Conclusion

In this work, we investigate the mobility problems in next generation mobile communication networks. There are totally five different mobility problems, each with different

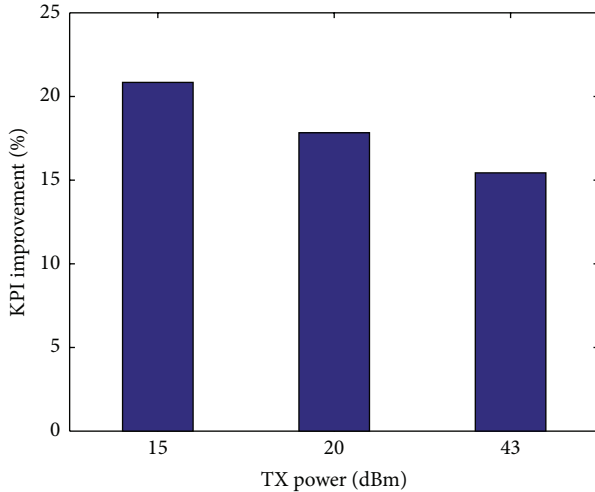


FIGURE 9: KPI improvement with various TX power.

characteristics. We propose a data-driven handover optimization (DHO) approach, which aims to mitigate mobility problems. The DHO approach consists of four parts:

- (1) Mobility problem identification.
- (2) KPI design.
- (3) KPI function estimation.
- (4) Handover parameter optimization.

The DHO approach collects data from the mobile communication measurement results and provides a model to estimate the relationship between the KPI and features from the collected dataset. Based on the model, the handover parameters, including the HOM and TTT, are optimized to minimize the KPI, which is a weighted average of different mobility problem ratios. Simulation results show that the proposed DHO approach could effectively mitigate mobility problems and provide better connectivity.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported in part by Ministry of Science and Technology (MOST), Taipei, Taiwan, under Grants nos. MOST 103-2221-E-155-005, MOST 104-2221-E-155-017, and MOST 105-2218-E-155-001.

References

- [1] J. G. Andrews, S. Buzzi, W. Choi et al., "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [2] K. Kitagawa, T. Komine, T. Yamamoto, and S. Konishi, "A handover optimization algorithm with mobility robustness for LTE systems," in *Proceedings of the IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '11)*, pp. 1647–1651, Toronto, Canada, September 2011.
- [3] 3GPP, "Self-configuring and self-optimizing network (SON) use cases and solutions," TR 36.902, V9.3.1, 2011.
- [4] A. Awada, B. Wegmann, I. Viering, and A. Klein, "A SON-based algorithm for the optimization of inter-RAT handover parameters," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 1906–1923, 2013.
- [5] T. Jansen, I. Balan, J. Turk, I. Moerman, and T. Kürner, "Handover parameter optimization in LTE self-organizing networks," in *Proceedings of the IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall '10)*, pp. 1–5, IEEE, Ottawa, Canada, September 2010.
- [6] T. Jansen, I. Balan, S. Stefanski, I. Moerman, and T. Kürner, "Weighted performance based handover parameter optimization in LTE," in *Proceedings of the IEEE 73rd Vehicular Technology Conference (VTC '11)*, pp. 1–5, May 2011.
- [7] W. Li, X. Duan, S. Jia, L. Zhang, Y. Liu, and J. Lin, "A dynamic hysteresis-adjusting algorithm in LTE self-organization networks," in *Proceedings of the IEEE 75th Vehicular Technology Conference (VTC Spring '12)*, pp. 1–5, Yokohama, Japan, May 2012.
- [8] A. Lobinger, S. Stefanski, T. Jansen, and I. Balan, "Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks," in *Proceedings of the IEEE 73rd Vehicular Technology Conference (VTC '11)*, pp. 1–5, May 2011.
- [9] V. Capdevielle, A. Feki, and E. Sorsy, "Joint interference management and handover optimization in LTE small cells network," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 6769–6773, IEEE, Ottawa, Canada, June 2012.
- [10] T. Ali and M. Saquib, "WLAN/cellular handover analysis for different mobility models," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 2758–2762, Ottawa, Canada, June 2012.
- [11] A. Awada, B. Wegmann, I. Viering, and A. Klein, "Cell-pair specific optimization of the inter-RAT handover parameters in SON," in *Proceedings of the IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '12)*, pp. 1168–1173, Sydney, Australia, September 2012.
- [12] D. Giacomini and A. Agarwal, "Vertical handover decision making using QoS reputation and GM(1,1) prediction," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 5655–5659, IEEE, Ottawa, Canada, June 2012.
- [13] D. López-Pérez, I. Guvenc, and X. Chu, "Theoretical analysis of handover failure and ping-pong rates for heterogeneous networks," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 6774–6779, IEEE, Ottawa, Canada, June 2012.
- [14] A. Rath and S. Panwar, "Fast handover in cellular networks with femtocells," in *Proceedings of the International Conference on Communications (ICC '12)*, pp. 2752–2757, IEEE, Ottawa, Canada, June 2012.
- [15] D. Xenakis, N. Passas, and C. Verikoukis, "A novel handover decision policy for reducing power transmissions in the two-tier LTE network," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 1352–1356, Ottawa, Canada, June 2012.
- [16] T. Taleb, K. Samdanis, and A. Ksentini, "Supporting highly mobile users in cost-effective decentralized mobile operator

- networks,” *IEEE Transactions on Vehicular Technology*, vol. 63, no. 7, pp. 3381–3396, 2014.
- [17] A. Imran, A. Zoha, and A. Abu-Dayya, “Challenges in 5G: how to empower SON with big data for enabling 5G,” *IEEE Network*, vol. 28, no. 6, pp. 27–33, 2014.
- [18] H.-D. Bae, B. Ryu, and N.-H. Park, “Analysis of handover failures in LTE femtocell systems,” in *Proceedings of the Australasian Telecommunication Networks and Applications Conference (ATNAC '11)*, pp. 1–5, IEEE, Melbourne, Australia, November 2011.
- [19] S. Haykin, *Neural Networks. A Comprehensive Foundation*, Prentice-Hall, 1999.
- [20] L. Behera, S. Kumar, and A. Patnaik, “On adaptive learning rate that guarantees convergence in feedforward networks,” *IEEE Transactions on Neural Networks*, vol. 17, no. 5, pp. 1116–1125, 2006.
- [21] J. A. Schnörr, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, vol. 97 of *Applied Optimization*, Springer, New York, NY, USA, 2005.
- [22] M. Gudmundson, “Cell planning in Manhattan environments,” in *Proceedings of the 42nd Vehicular Technology Conference*, pp. 435–438, IEEE, Denver, Colo, USA, May 1992.
- [23] P.-E. Ostling, “Implications of cell planning on handoff performance in manhattan environments,” in *Proceedings of the 5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications: Wireless Networks—Catching the Mobile Future*, vol. 2, pp. 625–629, The Hague, The Netherlands, September 1994.
- [24] M. Karabacak, D. Wang, H. Ishii, and H. Arslan, “Mobility performance of macrocell-assisted small cells in Manhattan model,” in *Proceedings of the 79th IEEE Vehicular Technology Conference (VTC-Spring '14)*, pp. 1–5, Seoul, South Korea, May 2014.
- [25] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, “Simulating LTE cellular systems: an open-source framework,” *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 498–513, 2011.

