IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 17, 2020, accepted October 31, 2020, date of publication November 5, 2020, date of current version November 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3036070

A Survey on Next-Cell Prediction in Cellular Networks: Schemes and Applications

LINYU HUANG, (Member, IEEE), LIANG LU, AND WEI HUA
College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

Corresponding author: Wei Hua (huaw@scu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61801318.

ABSTRACT Mobility prediction is a powerful tool for network operators to optimize network performance. From cell level, if network operators know the cells to which the users will be connected in advance, wireless resources can be pre-allocated to improve network performance and better user experience can be provided in location-based services. Many next-cell prediction models and methods have been suggested and implemented. This article is devoted to next-cell prediction (cell level mobility prediction) in cellular networks, and provides a thorough survey of the prediction schemes and applications. Particularly, a two-level classification methodology was proposed and applied. We first divided the prediction schemes into three categories based on the mobility data used for prediction, i.e. Current Movement State based Approaches (CMSA), Historical Movement Pattern based Approaches (HMPA), and HybriD Approaches (HDA). Prediction schemes in each category were further classified based on the used prediction methods. The typical application scenarios were introduced as well, including handover management, resource allocation, etc. Finally, current challenges and potential trends in the near future were further discussed.

INDEX TERMS Next-cell prediction, mobility prediction, handover management, resource allocation, location-based service.

I. INTRODUCTION

Recently, new frequency in mid- and high-bands has been allocated for fifth generation mobile network (5G). To deal with the coverage problem caused by shorter signal penetration and range, smaller cells have to be densely deployed for 5G cellular networks. Smaller cell coverage implies more frequent handover for mobile users, which poses great challenges for radio resource management among cells.

To avoid connection termination during handover, one conventional scheme is to make each base station (BS) reserve a fixed amount of resources for new coming users in the future. However, if resources are not reserved properly, it may cause resource waste or bad user experience. To alleviate this negative impact of user mobility, an efficient way is the implementation of mobility prediction [1]. If user mobility can be successfully forecasted, network operators can adopt a passive (or in-advance) reservation policy [2] to ensure service continuity without wasting huge amounts of resources.

The key to the successful implementation of passive reservation is the capability of user mobility prediction. The accuracy of prediction will have a direct impact on the performance of resource allocation in wireless networks. It is well known that humans tend to have similar patterns of cyclic behavior, which makes it possible to predict individual movement from previous historical location information. It has been shown that the predictability of user mobility can be reached at 93% [3]. However, such a high precision prediction has many limitations, such as sufficient context data and special movement patterns, which makes it hard to achieve. Fortunately, to the best of our knowledge, cell level prediction is sufficient to achieve continuous connection during handover, instead of accurate position prediction. Therefore, this article concentrates on the problem of next-cell prediction and provides a systematic survey on prediction schemes and applications.

II. RELATED WORKS

From the perspective of mobility big data analytics, the authors of [4] review geolocation prediction methods. They mainly focus on three topics: methods for mining popular geological regions (such as home and workplaces) from raw geological data collected by smart mobile devices, methods for mining personal trajectories that

The associate editor coordinating the review of this manuscript and approving it for publication was Mostafa Zaman Chowdhury.

consist of popular geological regions, and geolocation prediction models. Four types of geolocation prediction methods are explained, enveloping Markov-based methods, Bayesian network-based methods, regression-based methods, and neural network-based methods. However, the survey mainly considers how to deal with mobility big data and mine useful information for movement prediction, and a few prediction approaches are introduced. Besides, application scenarios are not introduced.

More comprehensive mobility prediction surveys have been made in [5]–[7]. The authors of [5] present a thorough survey of anticipatory mobile networking. They introduce and classify different types of mobile context that can be predicted and exploited, including geographic context, link context, traffic context, and social context. Some selected prediction methods are explained for these contexts in four main categories: time series methods, similarity-based classification, regression analysis, and statistical methods for probabilistic modeling. Optimization techniques adopted by anticipatory networking solutions are also introduced. In [7], the authors provide an overview of location prediction, including basic concepts, data sources, and trajectory pre-processing approaches (i.e. data cleaning, noise removal, trajectory compression, and feature extraction), and applications. From the perspective of prediction methods, the paper focuses on deep learning, pattern-based, semantic-based, distributive big data, social network, representative, and content-based approaches. Additionally, a survey based on geo-social networking data is discussed in [6], which is expanded from the following five aspects: problem categories, data sources, feature extraction, mathematical models, and evaluation metrics. Except for mobility data, the article mainly describes an investigation on social information, such as check-in data and social relations. From the view of prediction model and approaches, matrix factorization, Markov chain, neighborhood model, statistical learning model, ranking-based model, and embedding techniques are described. However, these articles did not consider the impact of user mobility on network performance and did not provide a detailed overview of mobility prediction based applications.

Some cell level surveys have been presented in [8], [9]. The authors in [8] describe the latest cellular technologies which can be optimized with mobility prediction, including heterogeneous networks and femto-cells with 3G, 4G, 4G+, 5G, and IEEE 802.11ac wireless local area network (WLAN). Signaling and reservation protocols for resource management are presented as well. Various mobility prediction approaches are surveyed in detail, including Markov processes, Kalman filters, neural networks, data mining, time series and bio-inspired approaches. In [9], the authors describe in detail on mobility prediction, such as how to obtain user's position information, prediction methods, and prediction outputs. Prediction schemes with different methods are surveyed, including Markov chain, Hidden Markov Model (HMM), Artificial Neural Network (ANN), Bayesian network, and data mining. However, these two surveys mainly deal with prediction methods and pay less attention to types of mobility data used for prediction in separate schemes.

In the existing literature, numerous prediction approaches have been suggested, with different types of mobility information have been utilized. To better understand the advantages and disadvantages of these methods, an appropriate classification methodology needs to be implemented to classify these works. In [10], mobility prediction schemes are mainly classified into two categories according to the mobility pattern used for prediction, i.e. individual user mobility information and group mobility patterns.

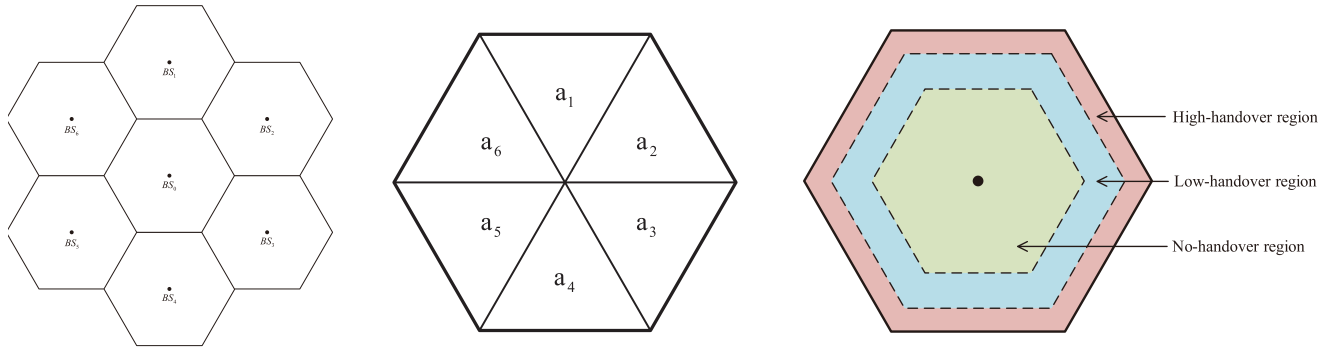
In this survey, we focus on cell-level mobility prediction in cellular networks. We propose a two-level classification methodology, which first divides the prediction methods into three categories based on the mobility data used for prediction, i.e. Current Movement State based Approaches (CMSA), Historical Movement Pattern based Approaches (HMPA), and HybriD Approaches (HDA). Prediction methods in each category are further classified based on the used prediction methods. Specifically, angle-based algorithms, distance-based algorithms, angle-distance combined algorithms, Markov chain, HMM, Bayesian network, Support Vector Machine (SVM), ANN, and data mining are discussed in detail. Considering cell structure is an important attribute while not discussed in detail in the aforementioned articles, we also pay special attention to the utilization of cell structure in prediction. The main contributions of this survey are listed as follows:

- The predictability of cell-level mobility prediction is discussed;
- A two-level classification methodology is proposed and applied;
- The works on next-cell prediction are classified and overviewed;
- Next-cell prediction based applications are reviewed;
- Current challenges and future potential trends on next-cell prediction are discussed.

The rest of this article is organized as follows. Section III provides an outline of basic concepts and characteristics of mobility prediction. Then, next-cell prediction schemes based on current movement state and historical movement pattern separately are introduced in Section IV and Section V, respectively. Hybrid prediction approaches that utilize both current and historical movement information are introduced in Section VI. After that, next-cell prediction based applications, current challenges, and future potential trends are presented in Section VII. Finally, Section VIII concludes the paper.

III. NEXT-CELL PREDICTION: CONCERNS AND CHARACTERISTICS

In this section, we discuss the related issues and characteristics of next-cell prediction, including predictability of next-cell, the types of commonly used mobility data, and how to



(a) Hexagonal cells. (b) A 6-sector cell. (c) A multi-tier cell.

FIGURE 1. Cell structures.

obtain the important mobility parameters (i.e. current position and moving direction).

A. PREDICTABILITY OF NEXT-CELL

1) REGULARITY OF USER MOBILITY

In daily lives, our movement is not completely random, but directed or destination oriented. People often move along a specific route and move regularly where they live for a long time. For instance, students go to school usually follow a particular route, also the commute route of office workers on weekdays will hardly change. In these trajectories, the cells to which the mobile users will be connected are fixed with high probabilities. By observing these users' mobility patterns for a certain period, it is possible to obtain the regularity of their movements and further to predict the succeeding cells while they are moving. The work of [3] analyzes the extent to which human behavior is predictable. By measuring the entropy of numerous trajectories of mobile phone users, they found that the potential predictability of user mobility could be up to 93% and the predictability is lack of variability. In other words, there exists strong regularity in human mobility, and it is theoretically possible to develop accurate prediction models. The predictability of vehicular mobility has also been investigated in [11], which shows that about 78% - 99% of the location and above 70% of the staying time are predictable, respectively. They also revealed that there is strong regularity in everyday vehicular mobility, which can be exploited in the development of practical prediction algorithms.

Actually, user movement can be regarded as a combination of regular movement and random movement [12]. For the regular movement, the regularity can be discovered through users' long-term (weeks or months) historical trajectories. Numerous prediction approaches have been suggested to discover the regularity for next-cell prediction in literature, such as Markov chain [13], HMM [14], Bayesian network [15], SVM [16] and ANN [17]. These methods will be discussed in detail in later sections. However, when a user comes to a new place, there is no sufficient historical data to mine its mobility pattern, and its movement shows irregularity. Though the prediction of random movements is relatively more difficult due to irregularity, next cell can also be determined to some extent by various strategies, such as real-time monitoring [18]. Approaches for predicting random movements will be introduced in later sections as well.

2) COARSE LOCATION PREDICTION

From the perspective of space granularity, unlike position and trajectory predictions, next-cell prediction is a coarse location prediction. It only needs to determine to which cells users will be connected, instead of accurate locations. Therefore, it is generally easier to predict a cell or cell sequence than a fine position or trajectory.

3) CELL STRUCTURE

For wireless networks, the coverage of BSs is modeled as regular hexagons generally. Therefore, researchers can investigate cell geometry with various cell structures to assist forecasting the next cell. The most commonly used cell structures are illustrated in Fig. 1, including ideal hexagonal cells [19], [20], 6-sector cell [21], and multi-tier cell [22].

Hexagonal cells can ensure that the minimum number of BSs are deployed to cover a certain region. By equally dividing a hexagonal cell into 6 parts, 6-sector cell structure further improves the granularity of prediction and increases the length of mobility data. Based on multi-tier cell structure, the timing of prediction can be better managed. If predicting users' next cell while they enter high-handover region, the forecast cost can be substantially reduced.

B. OBTAINING USER MOBILITY INFORMATION

User mobility information of both intra-cell and inter-cell movement can be utilized for next-cell prediction. The commonly used mobility data includes current and historical position, moving direction, speed, acceleration, and traverse history. Since historical information is accumulated from current mobility information, so we only consider how to obtain current mobility information in this article. The most important parameters are current position and moving direction, which can directly affect related next cell. Whereas the additional parameters, such as speed and acceleration, can be calculated by a series of positions [23]. Now, we are going to introduce how to obtain the current position and moving direction.

1) CURRENT POSITION

In cellular networks, each BS has been assigned a Cell Identity (Cell-ID), and the corresponding location is known to the network operator. The Cell-ID can be used to identify to which cell a user is connected. It can also be used as a coarse positioning method, which can achieve a precision of several hundred meters, instead of specific location coordinates. It is convenient to get users' Cell-ID for network operators without setting any extra devices on User Equipment (UE). Since Cell-ID history reflects users' inter-cell movements, it is useful for determining future cell sequence.

Although Cell-ID can provide rough location information, many prediction applications need a more accurate position to investigate intra-cell movements. The Received Signal Strength (RSS) based positioning method can provide a more precise location. Based on RSS, it can determine the distance between a user and a BS based on signal propagation models [24]. When the distances from a user to multiple BSs are obtained, user position can be calculated by cell tower trilateration [25]. The authors of [26] introduce a method to determine the position, speed, and direction of users based on measured RSS.

Furthermore, with the wide application of Global Positioning System (GPS), most smart UEs are equipped with GPS because of its convenience and high accuracy. By periodically reporting GPS information from UE to BSs, network operators can collect user position with an accuracy of several meters [27].

There also exist some additional methods to determine user locations, such as positioning with APP check-in or social dynamics [28]. Additionally, indoor positioning mechanisms such as WLAN, Radio Frequency Identification (RFID), and Bluetooth, will not be introduced in this article, we refer the interested readers to [9].

2) MOVING DIRECTION

Once user positions have been determined, the moving direction can be obtained by various methods. The simplest way is to get the vector of two consecutive positions measured over a short time [29].

Additionally, the authors in [30] introduce different direction estimation methods as depicted in Fig. 2. A virtual circle is established in the cell, and user position will be recorded within the circle at a fixed time interval. Based on the collected position, the direction of motion can be determined by various algorithms. Method 1 first determines direction angle between every two consecutive positions, then calculates the average of all angles to estimate the direction of movement. Method 2 sets another small circle within the big circle, and only considers the average of user angles recorded within the circular strip. Method 3 performs exponential moving average on the user angles recorded in the circle, which gives higher weight to later angles and then makes the average, as the later positions are theoretically more decisive to future direction. Method 4 uses only instantaneous angle of users before leaving the virtual circle.

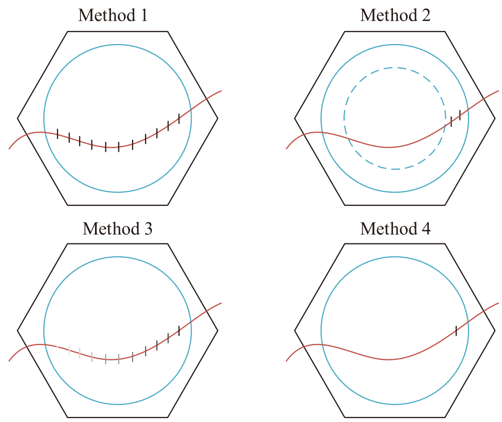


FIGURE 2. Direction estimation methods.

The moving direction can also be predicted by Lagrange interpolation. The position reported by a user is a series of discrete coordinates stored in the cache, then Lagrange polynomials are established by fitting these co-ordinates to represent the trajectory. The moving direction can be predicted by analyzing the slope of the corresponding trajectory [18].

Except for mobility data, there also exists some works using other types of information, such as social relation [6], Channel State Information (CSI) [31], road topology [32], user behavior [33], and so on, which will not be expanded in this article.

C. CLASSIFICATION OF MOBILITY DATA

Mobility data are commonly divided into two classes: individual mobility information and group mobility pattern [10]. However, this classification cannot reflect the relationship between movements and time series well. To forecast the next cell and future crossing cells, both short- and long-term patterns should be taken into consideration. In this survey, based on time attributes, mobility data for next-cell prediction is mainly classified into two categories: users' Current Movement State (CMS), and Historical Movement Pattern (HMP).

1) CURRENT MOVEMENT STATE (CMS)

CMS is defined as the information related to real-time movement, including the current position, moving direction, speed, acceleration, etc., where position and direction are most commonly used. CMS can be obtained by monitoring user position in real-time via some methods. For example, UEs can periodically report their location coordinates to BS, and BS can further calculate other CMS parameters, e.g. speed, direction, and acceleration, based on positions. Once the CMS is obtained, both regular and random movements in the near future can be predicted.

2) HISTORICAL MOVEMENT PATTERN (HMP)

HMP information includes Cell-ID with CMS history, traversed cell sequence, handover history, historical sectors, and so on. By observing users' long-term (in weeks or months) movements, HMP can be obtained and stored in a database. Besides, periodically reported CMS will also be stored in the database to update HMP. By analyzing large volumes of HMP data via some tools, such as machine learning or data mining, the relationship between connected Cell-ID and movement parameters or cell transitions of users can be discovered, which are known as movement patterns or mobility rules. Based on these patterns or rules, the cells that a user will cross in the near future can be predicted.

D. CLASSIFICATION OF PREDICTION METHODS

Based on mobility information, a detailed survey of next-cell prediction algorithms for the last two decades are given in this article. For better understanding and management of these methods, we propose a classification methodology based on the types of mobility data used by prediction. All prediction approaches can be divided into three categories: CMSA, HMPA, and HDA.

Particularly, CMSA exploits users' real-time movement states to calculate the most possible adjacent cell that will be connected to. Whereas HMPA analyzes users' movement history to construct mathematical models, and further mines movement patterns or mobility rules to predict future crossing cells. Considering both current movement state and historical movement pattern, HDA is a kind of complementary approaches. It is worth noting that HMPA also needs to take CMS as the input of prediction models.

IV. CURRENT MOVEMENT STATE BASED APPROACHES (CMSA)

By monitoring a user's position constantly, network operators can calculate that user's speed, moving direction, etc., so as to obtain the full CMS information. Many next-cell prediction approaches have been proposed by exploiting CMS. In order to predict the next cell, context information, such as cell geometry, is also considered in general. Based on different determining factors (i.e. angle and distance), we divided CMSA into three categories: angle-based, distance-based and angle-distance combined ways.

A. ANGLE-BASED APPROACHES

It is expected that users' movements are directed to their intended future cells. Consequently, the subsequent cell can be determined by investigating moving direction. In this subsection, we introduce angle-based approaches that utilize angles between the target user and neighboring BSs or Access Points (APs) to determine the adjacent cell into which the user will enter, as shown in Fig. 3.

The work of [20] consider 6-sector cells where each BS is equipped with six 60-degree directional antennas. Each sector is adjacent to only one neighboring cell. In this way, the next

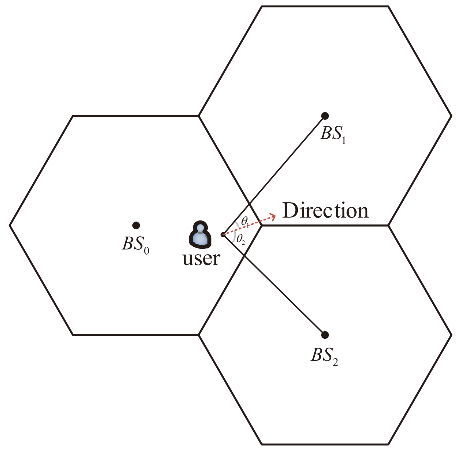


FIGURE 3. Angle-based approach.

cell(s) that a mobile user will move into is limited by the last sector that user passed, which makes it easy to predict. At the same time, the concepts of non-critical region (handover probability is fairly low) and critical region (handover probability is fairly high) are introduced to further divide each sector into 2-tier-like structures. The critical region is dynamic and determined by historical handover points. If a user enters the critical region, the system determines his/her next cell based on moving direction by tracking his movements within the sector.

In [22], the authors introduce 6-sector and 3-tier based cell structures. Each cell is divided into reservation zone, handoff zone, and non-reservation-zone based on RSS level from the BS to users. In the proposed scheme, the next cell is predicted according to the RSS measurement from the current position and extrapolation of the UE direction when the users in the reservation zone, handoff zone, and leaving the non-reservation-zone, respectively.

Whereas, the authors in [18] explicitly propose a scheme that predicts mobile direction by fitting a polynomial equation along the trajectory based on Lagrange's interpolation. Similarly, the paper considers hexagonal cells within two-region structures. Within the incircle of the cell, handover will not be considered, while the circumcircle is regarded as a handover region. When a user moves from the incircle to the handover region, the motion direction at the handover region is predicted by calculating the slope of the trajectory at the starting point of handover based on a polynomial equation. Finally, the next cell is determined by investigating the relationship between the angular range and neighboring cells.

B. DISTANCE-BASED APPROACHES

Except for angle, the distance between a user and neighboring BSs can also be utilized to predict the next cell, as shown in Fig. 4. The distances are generally calculated by laws of signal propagation. The larger the distance between a user

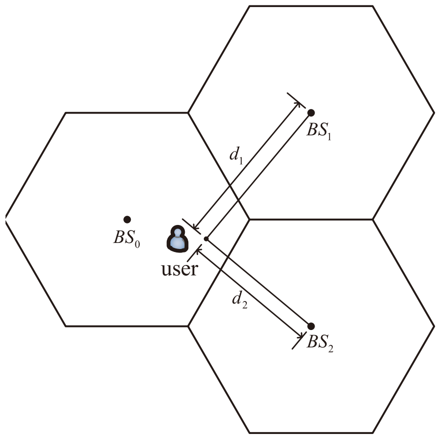


FIGURE 4. Distance-based approach.

and a cell is, the smaller probability that the user comes into that cell will be [19].

Assuming that the wireless propagation condition is homogeneous over the service area of the system, the authors of [19] consider a uniform grid of hexagonal cells system model, and present an adaptive fuzzy inference prediction system to predict the next cells based on the distance. The system consists of two parts, a fuzzy inference system and a recursive least square predictor. Particularly, the fuzzy inference system estimates the probability that a mobile user will be active at the moment based on real-time measured pilot signal power. Then the recursive least square predictor predicts the probability in the future.

C. ANGLE-DISTANCE COMBINED APPROACHES

Considering both angle and distance information, combined prediction methods were invented. The authors of [30] introduce an angle-distance combined approach to predict cell transition. Specifically, they propose a probability equation, which is determined by both moving angle and distance, to predict the succeeding cell. Considering Ultra-Dense Networks (UDNs) have small diameters and the network environment is more complex, which makes the conventional prediction schemes cannot play an effective role, the authors of [34] present a combined approach to predict Transmission Points (TPs) in UDNs. Similar to [18], the paper exploits Lagrange's interpolation to estimate moving direction. But it considers not only angle but also the distance between a user and adjacent TPs. It first determines a group of possible TPs whose angle with user moving direction is less than a threshold, then calculates the transition probabilities based on the distance between TPs and users to determine a target TP.

In summary, for the determining factors, distance is more important. Because the closer to a BS, the more likely it is to enter the corresponding cell. On the other hand, The users may also change their directions and enter other cells. Therefore, both angle and distance should be considered when predicting.

V. HISTORICAL MOVEMENT PATTERN BASED APPROACHES (HMPA)

Being different from CMSA, HMPA constructs mathematical models or mines the relationships of cell transitions to predict future crossing cells, based on a large volume of historical movement patterns. In this article, we mainly explain three categories of HMPA: probabilistic models, discriminative models, and data mining. Specifically, we introduce Markov chain, HMM, Bayesian network, SVM, ANN, and association rules. HMPA can predict next cell, especially for regular movements, with relatively high accuracy based on user mobility pattern.

A. PROBABILISTIC MODELS

A probabilistic model is based on the theory of probability and statistics, which incorporates random variables and probability distributions into the model of an event or phenomenon. Prediction methods have the advantages of high efficiency and flexibility in dealing with large-scale sample classification and linear system learning. Commonly used models are Markov chain, HMM, and Bayesian network. By calculating the joint probability, the possible next cell can be predicted.

1) MARKOV CHAIN

A Markov chain is a discrete stochastic process with the Markov property (also called memoryless property) which means that the next state is only related to the current state, but not to previous states. Markov chains greatly simplify the complexity of stochastic models, reduces the difficulty of calculation, and are commonly used in mobility prediction [35].

The work of [36] applies Markov chains to predict the next cell for both simple and complex environments. Specifically, for simple scenarios (e.g. highway or street, the user mainly moves in one direction), the model divides the next possible cells around a user into two parts: left and right. Then use a three-state Markov chain to predict the next cell, in which stationary, left move and the right move states are used. For complex urban environments, a seven-state (i.e. the current cell and six adjacent cells) Markov chain is proposed for next-cell prediction. The transition probability matrix can be calculated by fitting historical data. Based on the initial distribution and transition probability matrix, next cell distribution can be forecasted.

In IEEE 802.16m networks, the authors of [37] consider the case of three BSs to establish Markov chains. Each cell is regarded as a state, and user movement between cells is regarded as state transitions. The most likely next cell can also be determined according to the initial distribution and probability transition matrix.

Considering a system with 4 APs, a second-order Markov chain is created in [38]. Target AP is predicted based on the current state and the previous states. A simulation with the data trace of CRAWDAD database is also performed to verify the model.

Although the Markov chains are applicable to those users moving regularly in a region, it is difficult to predict the movements of newly entered users. To solve this problem, the work of [39] introduces the concept of visit frequency. Given a particular neighboring cell, the visit frequency is defined as the ratio of the total visited number of this cell to the total visited number of all adjacent cells that appeared in historical data. When the historical data is sufficient, a second-order Markov chain is utilized. Otherwise, a first-order Markov chain is employed for prediction. If there is no available historical data (such as new users in the region), Markov chains is not available and the visit frequency is used for prediction.

The authors of [13] use a Markov chain to predict movements of vehicles between APs under vehicular networks. When building the model, the initial distribution takes vehicle positions and speed into account. They also use data mining to analyze and process the historical data, and then establish a transactional database from collected historical data, so as to obtain the transmission probability matrix.

2) HIDDEN MARKOV MODEL

Hidden Markov Model (HMM) is a doubly stochastic process with an underlying stochastic process that is not observable (which is hidden), but can only be observed through another set of stochastic processes that produce the sequence of observed symbols [40]. Different from Markov chains, HMM contains two state sequences: a hidden state sequence $\mathbf{S}$ and an observable one $\mathbf{O}$. For normal Markov chains, we can only see one state, and cannot make full use of user information. While in HMM based schemes, two kinds of context information (e.g. traversed cell sequence and geographical location) can be taken into consideration simultaneously to achieve better prediction accuracy. There is also a Markov chain between the hidden states $\mathbf{S}$. The next state is only related to the current state, but not to the previous states. At the same time, hidden states $\mathbf{S}$ determine the distribution of observable states $\mathbf{O}$. An HMM has three parts, including initial distribution π, transition probability matrix A among hidden states $\mathbf{S}$, and observation transition probability matrix B from hidden states $\mathbf{S}$ to observable states $\mathbf{O}$. A and B can be obtained by fitting history movement data. Accordingly, an HMM can be represented as $\lambda = \{A, B, \pi\}$ [40]. In HMM, there exist three key problems: evaluation, decoding, and learning. To achieve different purposes, we can solve the corresponding problems.

a) Evaluation
Given the observation sequence $O = \{O_1, O_2, \ldots, O_T\}$ and model $\lambda = \{A, B, \pi\}$, where T is the time index, computing the probability of the observation sequence $\Pr(O|\lambda)$. This problem can be solved by the forward-backward algorithm [40].

b) Decoding
Given the observation sequence $O = \{O_1, O_2, \ldots, O_T\}$, determining optimal hidden state sequence $\mathbf{S}$ in some meaningful sense. This problem can be solved by the Viterbi algorithm [40].

c) Learning
Adjusting the model parameters $\lambda = \{A, B, \pi\}$ to maximize $\Pr(O|\lambda)$. The Baum-Welch algorithm can be utilized to solve this problem [40].

In HMM based next-cell prediction approaches, one of the key issues is to choose appropriate hidden states and observable states. Considering cells and call duration time as hidden states and observable states, respectively, the authors of [41] propose an HMM based next-cell prediction method. Call duration time is decomposed into a sequence by the channel holding time, which dominates the number of visited cells and the cell dwell time during a call duration life. Given the call duration time distribution, HMM can be used to calculate the cell distribution, so as to determine the target next cell. This is an optimal state sequence problem (decoding problem), which can be solved by the Viterbi algorithm [40]. The work of [42] regards femtocell AP as hidden states and users' geographical location as observable states, the next cell is predicted by solving the decoding problem. Similarly, regarding cells as hidden states, and the cell distribution can be established by solving the decoding problem. The interested readers can refer to [43], [44].

HMM can also be combined with other algorithms for next-cell prediction. The work of [45] proposes a vehicle network movement prediction approach based on the combination of HMM with Kalman filtering. The authors take the mobility information of vehicles as the observable states and the connected AP as the hidden states. A Kalman filter is used to predict the direction, speed, and position of vehicles, which will be used as observable states for HMM to predict possible next cell. Definitely, an HMM can combine another HMM as well. Dual HMM is utilized for optimal Wi-Fi AP prediction in [46], which two hidden states and two observable states. Users' current locations are input into the first HMM to get predicted locations, which will be the input of the second HMM to determine the next AP. In particular, the first HMM takes the future geographical coordinates of the mobile user as hidden states, and current geographical coordinates as observable states, respectively. This HMM is used to predict the user location. While in the second HMM, optimal Wi-Fi AP and geographic location represent hidden states and observable states respectively, which is applied to predict the next Wi-Fi AP.

Except for predicting the next cell by solving the decoding problem, from another perspective, cells can also be regarded as observable states for prediction by solving the evaluation problem [14], [47]–[49]. From the perspective of AP controller, the work of [47] designs an HMM based algorithm by taking APs as observable states and users' real geographical locations as hidden states. The cell distribution is obtained by solving the evaluation problem. The works of [48], [49] introduce the method of HMM to predict the target cell and treat the prediction as an evaluation problem. Forward-backward algorithm is used to solve the problem and target next cell is predicted. The authors in [14] introduce the method of transforming both regular and irregular arrangemented cell networks into an undirected graph to build HMMs. They divided the prediction into two parts, learning and prediction. First, fitting historical data obtains the optimal parameters to complete learning. Then the prediction is completed in three steps: 1) according to the observable states (i.e. the cells at the current time), the distribution of hidden states at the current time is calculated by solving the decoding problem; 2) according to the hidden states at the current time, the distribution of the hidden state at the next time is calculated; 3) according to the distribution of hidden states at the next time, the observable states are determined by solving the evaluation problem (i.e. cell distribution at the next moment). Furthermore, they also introduce two improved schemes, normalized probability distribution and logarithmic summary, to improve the prediction accuracy.

3) BAYESIAN NETWORK

A Bayesian network is a directed acyclic graph probability model [8], where nodes represent random vectors and directed edges between nodes represent the relationship between nodes. Bayesian networks are very friendly to the modeling of complex environments with lots of impact factors. It can be used to predict the next cell by taking various mobility related information as input.

A Bayesian network based on the current location, direction, and street structure of vehicles is proposed in [50] to predict the next cell of vehicles. The established Bayesian network includes five random variables: lane, street section, GPS location, direction, and the next cell. In this Bayesian network, the lane is the parent node, street section, direction, and the next cell are its child nodes. GPS location is the child node of street section. The Bayesian network predictor in [51] considers many other factors, such as cell distribution, road structure, user mobile information (i.e. position, velocity, and acceleration), and random factors. The establishment of Bayesian network is divided into four processes: 1) establish Cell and Road Topology (CRT) based on cell environment, 2) establish Road State Transition (RST) via CRT, 3) establish Probability Distribution Network (PDN) via RST, and 4) finally establish Bayesian network model via PDN. At the same time, the paper also analyzes the distribution of prediction parameters, and predicted the residence time according to the current position, speed, and predicted position. Based on the next cell distribution, an adaptive paging scheme is proposed in [52]. Because the location of specific user groups has a strong correlation, the work of [33] introduces the collective behavior patterns, and proposes a mobility prediction method based on the collective behaviors, which uses the location of other users to predict another user's movement. Based on the collective behaviors mobility model and Bayesian network, the authors of the paper propose the collective behaviors patterns-based Bayesian predictor and construct the hybrid schemes with Markov based predictor. In [15], a user profile prediction method is proposed, including user location and service pattern prediction. Taking cells as the minimum units of prediction, the posterior probability of user entering target cell is calculated by using Bayesian theory.

Limited by the length of the article, some other approaches, such as Markov Renewal Process [53]–[56], Kalman Filter [57], [58] will not be described in detail.

B. DISCRIMINATIVE MODEL

Since the next-cell prediction is to choose one cell from multiple possible cells, it can be regarded as a classification problem that divides the candidate cells into two classes. A discriminative model is based on the discriminant function generated by limited samples to find the best discriminant boundary or classification surface between different categories, reflecting the differences between different types of data.

1) SUPPORT VECTOR MACHINE

A Support Vector Machine (SVM) is suitable for solving small-sample, non-linear, and high-dimensional problems. It has a strong anti-noise ability for data noise and can reduce the influence of outliers on the models. The basic idea of its classification is to transform the nonlinear input space into a high-dimensional space by defining an appropriate kernel function and to find the support vector in the high-dimensional space to form the optimal hyperplane.

In the scenario of next-cell prediction, since the predicted location has many possibilities, it needs to use multi-value SVM for prediction. Considering the constraints of geographical topology, the work of [59] considers a scenario with an urban center, regular streets, and rail areas. In the paper, mobility sequence, generated by mobility model of integrated path follower, gravity, and random walk models, are fed into multiple classes SVM for training and further to predict the most likely next cell. Their simulation results show that the prediction accuracy can exceed 90%. To improve prediction accuracy, the authors of [12] proposed a multiclass SVM method with considering regular and random movements separately. Two different multiclass SVM location sample vectors are designed to treat the two mobility patterns differently. In their models, heterogeneous network system model of Long-Term Evolution (LTE) and WLAN is considered, where each irregular LTE cell contains a round WLAN cell. Target region is also set in each cell, which is similar to the 2-tier structure. When users enter the target region, prediction is started.

In [16], a SVM based cell forecast method under 5G UDN is proposed. The prediction is divided into two stages: sample acquisition stage and SVM prediction stage. In the former stage, user ID, time, location coordinates, and speed are collected and preprocessed. In the latter stage, SVM model is trained with historical data and results are predicted.

The work of [31] proposes a SVM based next-cell prediction scheme by using short-term CSI and long-term handover history. In the paper, the prediction is formulated as a classification problem, where the CSI sequence serves as an input vector, and the next cell index as the label. In the training phase, training CSI sequences, their associated previous cells and their next cell indices derived from handover history are exploited. While in the prediction phase, testing CSI sequence and associated previous cells will be inputted to obtain next possible cell. Since that CSI feedback and handover history are readily available in cellular networks, no signaling costs are added to the radio link and implementation effort is limited. Their simulation, conducted from a Manhattan grid scenario and a realistic radio map of downtown Frankfurt, shows that SVMs predict the next cell substantially more accurately with CSI than with handover history alone, and can reach 100% prediction accuracy with not more than 60% of the CSI input values are required.

2) ARTIFICIAL NEURAL NETWORK

An Artificial Neural Network (ANN) has obvious advantages in dealing with random data and non-linear data. It is particularly suitable for systems with large data scale and unclear information. Moreover, the model structure is simple. When sample data is sufficient, ANNs can be effectively used for mobility prediction, with strong robustness and fault tolerance. However, one of the disadvantages of ANNs is that, in most cases, it needs a lot of parameters and the training process takes a long time [60].

In [61], a next-cell prediction method based on neural network is proposed, with a 2-tier cell structure based on pilot signal strength. It starts to predict after users entering the tier-2 area, to save resource consumed by prediction. Taking user coordinates as input, the feedforward neural network is selected for training and prediction. By such a method, when a user suddenly changes its direction, the correct prediction results can still be obtained and the prediction accuracy can be improved. In [23], the inter-cell movement is treated as a function of user position, velocity, acceleration, and direction can be calculated by three consecutive available position points. This implies that only three consecutive user locations needed to be considered in the prediction process. Two-group-complete-bipartite-meshed feed-forward neural networks are selected to address the problem, which takes the coordinates of three consecutive positions as input, and three future consecutive cells as outputs. Based on user movement direction, backpropagation neural networks (with low complexity) based prediction method is proposed in [62], which can handle the uniform, regular or deterministic movements. In the paper, a rectangular cell structure is considered and an 8×8 array of cells were defined. The work of [63] presents a more realistic mobility model called Smooth Random Mobility Model. It depends on the change of speed and direction, which can make the trajectory smoother. Two kinds of neural networks are proposed for the case with and without historical handover data, respectively. For the case with no historical data, the proposed neural network only depends on current position and directions, which takes cluster identifier, cell identifier, and movement direction as input. The other neural network is based on historical data, which recursively calls the previous prediction results as input for prediction. In [64], a regular rectangular system network is considered, and users may connect to one of the nearest three BSs. The distance between a user and a BS is calculated by RSS and path loss model. Using Recurrent Neural Networks with long and short-term memories, the sequences of RSS values are used as the input of the neural network for training and prediction.

Besides, the authors of [17] use Recurrent Neural Networks with both long and short-term memory for next-cell prediction. The authors of [65] propose supervised learning algorithms to forecast future Wi-Fi APs by exploiting historical connected information. The information includes instants of time when the user begins and ends a connection to each AP, the average SNR, the average Received Signal Strength Indicator (RSSI), and the number of bytes transmitted/received during the connection of the user to each AP. In particular, neural networks and random forests are exploited for training and prediction. At the same time, the authors study the influence of historical data granularity (i.e. daily and weekly) on prediction. Their results show that the neural networks based approach can get higher prediction accuracy than random forest based one, at the expenses of an increase in the computation time. Also, the prediction accuracy is higher when jointly considering time-period, daily, and weekly historical information than using separate data only. The work of [66] uses the real mobility model to test the mobility prediction method based on extreme gradient boosting trees and deep neural networks to predict the future cell.

Some other discriminative approaches, such as clustering [67], [68] and Random Forest algorithm [69], will be skipped in this article.

C. DATA MINING

Data mining, in terms of association rules, is widely used in mobility prediction. Association rules can find hidden links between different data. These links could be useful and available. The famous discovered relationship is that if a customer bought bread, butter, and coffee, it is likely that he/she would also buy milk. In our case, this technology can be used to discover the relationship between cells and obtain some types of information, such as users' movement history, roads, and locations of BSs [70].

In mobility prediction, most methods only consider the spatial factors, as it is directly related to users' future locations. However, research results show that time characteristics also have a significant impact on mobility prediction. Most conventional studies do not consider spatial and temporal attributes of data simultaneously. The work of [71] suggests a spatiotemporal data mining approach to predict the next cell. Data mining is divided into two parts: discovering frequent movement patterns, and then discovering frequent movement rules. Two algorithms, Allmop and Maxmop, are introduced to mine all frequent patterns and maximum patterns respectively. Finally, the user movement is predicted based on the movement rules. Similarly, the work of [72] also proposes a mobility prediction method based on spatiotemporal data mining, which includes four stages. In the first stage, a transactional database is generated from the historical log file; in the second stage, all frequent mobility patterns meeting the minimum support threshold are mined from the transactional database; in the third stage, the mobility rule is generated from the frequent mobility pattern; in the last stage, all eligible mobility patterns are used for mobility prediction. Based on [72], an improved scheme is suggested in [73]. The authors think that the recently mined mobility rule is more important than the foregoing ones. Therefore, in the prediction stage, different weights are allocated to the mobility rules based on temporal attributes, so as to improve the prediction accuracy. Another five-step spatiotemporal data mining based next-cell prediction method is introduced in [74]. In the first step, the reduction of noise/outliers depicting random movements is performed. The second phase is to partition frequent mobility patterns into similar groups or clusters. Group determination of current trajectory is performed in the third phase. The fourth step is to use the discovered frequent mobility patterns to generate mobility rules. Lastly, find the next cell-ID in the predicted region using the matched rule in preceding phases. Besides, the authors of [70] investigate using data mining to find the relationship between user location and other information, such as users' movement history, roads, and locations of BSs. The Apriori algorithm is used to get the association rules and predict the target cell.

D. OTHERS

There also exist some other methods based on historical movement patterns that are out of the scope of the above categories. In [75], an irregular network structure model is considered and a graph model is used to represent the relationship among users in cells. Derived from data compression techniques, the Ziv-Lempel algorithms are used to predict next cell, based on historical user location and handoff time. In [76], a cell matching method is proposed. It stores users' movement pattern history and current movement information in each user's local device. During prediction, the short-term current movement is compared with the user's pattern history. If there exists a matching trajectory, the next cell is predicted according to the historical data. Otherwise, no prediction is performed and normal handover is used.

A hybrid method is employed in [77]. Two kinds of user profiles are stored: local profiles containing personal information and global profiles containing a certain percentage of user information under the network. Accordingly, two prediction schemes, Local Profile based Prediction Algorithm (LPPA) and Global Profile based Prediction Algorithm (GPPA), are proposed. LPPA matches the local profile when the user's historical data is sufficient. Otherwise, it uses Markov chains or visit frequency to predict the user's historical track. GPPA calculates the possibility of users following the global profile to determine the target cell through the current cell. LPPA is preferred in prediction, and GPPA is used if LPPA fails.

Based on the historical mobility information of a user itself and its neighbors, the work of [78] exploits ant colony algorithm to predict the next cell. Considering five parameters, including user ID, period, source cell, destination cell, and date, the authors of [78] establish the system model based on ant colony system. The possible next cell is then determined through the pheromone field and visibility field. This method can be applied to both regular and random motions. Three prediction methods based on user history data are introduced in [79], which are probabilistic predictor based on Bayesian theory, group predictor based on ant colony optimization algorithm, and spatial predictor based on road topology structure. Combining these three schemes, this article also proposes a hybrid mobility prediction strategy, which can get higher prediction accuracy and avoid resource waste.

Based on the cell sequence of history, the work of [80] proposes a sequence model to predict next Cell-ID. By using graph embedded algorithm to determine the correlation between cells, a spatial loss function and a spatial cross entry loss function are proposed to predict the future Cell-ID.

VI. HYBRID APPROACHES

Although both CMSA and HMPA can predict the next cell, they still have some limitations. For example, CMSA can only carry out short-term prediction and may lead to a significant burden on UE due to constant monitoring. HMPA depends on the large volume of movements history. It must construct a model to predict next cell by consuming high computation complexity. Meanwhile, it cannot perform well when lacking history data, especially in the scenario that users enter a new region or move randomly. Consequently, HDA that investigate both long-term and short-term mobility prediction based on CMS and HMP emerged. On one hand, HDA exploits HMP to find the user's regularity of movements. On the other hand, real-time estimation is performed for instantaneous motions as well. Theoretically, HDA can deal with both regular and random movements while alleviating the negative impacts of CMSA and HMPA.

A method of sectorized mobility prediction algorithm is proposed in [21]. A 6-sector and 3-tier network structure are adopted in the system model. Consequently, a user's historical cell sequence can be replaced by the historical section sequence. Ideally, the data length will be 6 times the original. Due to the increase in data length, the prediction accuracy of regular movements is also improved. At the same time, a prediction method named cell sector numbering scheme is proposed for random movements. If the regular prediction fails, the real-time movement information will be monitored. When the user enters high handover regions, the target sector will be determined according to moving direction, so that target cell is also be determined. To solve the problem of poor accuracy when users move towards the corner of hexagonal cells in the sectorized method of [21], the authors of [81] propose an improved scheme based on genetic algorithm.

A 2-tier cell structure is applied and each sector is divided into two regions: high handover region and low handover region. Based on real-time RSS, the region in which user is located can be determined. If the user enters into the high handover regions, the target cell is determined by the genetic algorithm. Otherwise, the traditional sectorized method is used. It improves prediction accuracy and reduces the consumption of resources, and does not need to rely on historical data.

In order to solve the problem of fuzzy and irregular cell boundary prediction, road topology information is considered in [32], and second-order Markov chains are used to predict the current position and speed of users. The work of [82] introduces the User Mobility Profile (UMP), including quasi-stationary UMP with long-term information and dynamic UMP with short-term information. Firstly, the concept of zone is proposed to narrow the prediction range. Then, a mobility prediction model is proposed, which considers historical records, predictive future locations, moving direction, moving speed, and cell residence time. Finally, the most likely cell can be predicted based on historical information, current velocity, and dwell time.

Under the two-node all-IP network, the authors in [83] propose a prediction algorithm combining the Global Prediction Algorithm (GPA) and the Local Prediction Algorithm (LPA). A 6-sector, 3-tier system model is established. In GPA, gateway saves user's historical movement track, i.e. cell sequence, and uses the second-order Markov chain to predict the next cell. In LPA, the user's motion information is monitored in real-time to obtain the velocity in motion direction, so as to predict the possible cell. The combination of these two methods can predict regular and random movements.

A hybrid scheme based on handover history table and real-time GPS for Markov process prediction is described in [84]. In [85], the authors also introduce hybrid schemes. The next cell is determined by matching user's current path from the historical connection patterns. If the historical data volume is insufficient, the real-time GPS positioning method is used for prediction, based on moving angle and distance from the adjacent BSs.

Based on users' long-term and short-term trajectories, the authors of [86] propose a hybrid cell prediction approach. For the long-term trajectory, i.e. cell sequence in weeks or months, the Markov update process is used to mine the regular movements. The short-term trajectory, which records users' short-term movement in the cell, has a certain degree of randomness. It is used to determine user's movement direction through empirical moving average, so as to determine the next cell. Finally, the prediction results of Markov renewal process and short-term trajectory are combined by Dempster Shafer theory.

A method combined ant colony system and the sectorized approach is proposed in [87]. A 6-Sector and 3-tier system model is established. The prediction is divided into two stages: the first stage is ant prediction engine, which uses ant colony system to predict the user movements; the second stage uses stored dual mobility model to predict real-time movements. The average angle method is used to determine user movement, and the cell sector numbering method is used to predict the next sector. These two methods are combined to improve prediction accuracy. All the next-cell prediction models and schemes are described in Table 1.

VII. APPLICATIONS AND FUTURE

Widely deployed next-cell prediction schemes have been explored in previous parts. In this section, we introduce some application scenarios for next-cell prediction, including handover management, load balancing, resource reservation, and Location-Based Services (LBSs). Besides, although next cell can be predicted with high accuracy using those approaches, there still exist flaws that can be further improved. Accordingly, the current challenges and future potential research directions on next-cell prediction will also be discussed.

A. APPLICATIONS OF NEXT-CELL PREDICTION

Based on next-cell prediction schemes, wireless network operators know the adjacent cell or even a sequence of crossing cells which users will enter into in advance, then it may optimize network performance to provide better services.

1) HANDOVER MANAGEMENT

Handover is a critical issue in cellular networks. When a user leaves a cell and enters into another one, a handover process will be executed. A successful handover requires sufficient resources at target cell to ensure that the connection will not be terminated. Handover management is the key to achieve continuous service for cellular networks. Since next-cell prediction can estimate target BS (or AP) to which the UE will be connected in advance, the handover can be predicted as well. Therefore, handover management is one of the most suitable and intensively used applications. BS controller can investigate handover procedure to achieve a more efficient handover and eliminate unnecessary handovers, further to improve network performance via next-cell prediction. Here we will introduce three folds of handover management applications, enveloping enhanced handover mechanism, optimal target handover BS (or AP) selection, and ping-pong effect mitigation.

Next-cell prediction can be used to enhance handover procedure performance. Prior to next-cell information obtained, the serving BS can prepare a suitable handover strategy, i.e. proactive and reactive handover [37], to reduce handover interruption time and increase handover efficiency. Besides, to handover between the same or different types of networks, known as horizontal and vertical handover, next-cell prediction can play important roles. The authors in [88] investigate horizontal and vertical handover in LTE femtocell based on cell prediction. They analyze three handover scenarios: handover from Macro Base Station (MBS) to Femto Access Point (FAP) (hand-in), handover from FAP to MBS (hand-out), and handover from FAP to FAP (inter-FAP). In the paper, the corresponding proactive and reactive handovers are also

TABLE 1. A summary of next-cell prediction schemes.

Mobility data	Prediction algorithms		Advantages	Drawbacks	Scenarios	Works
CMS	Angle-based approaches		High real-time performance; moving direction sensitive.	Large overhead on UE.	Irregular movements; scarce mobility data; direction change frequently.	[18], [20], [22]
	Distance-based approaches		High real-time performance; high accuracy.	Large overhead on UE; moving direction is not considered.	Irregular movements; scarce mobility data; little change in direction.	[19]
	Angle-distance combined approaches		High real-time performance. Both direction and position are considered for determining next cell.	Large overhead on UE; relatively high complexity to compute distance and angle.	Irregular movements; scarce mobility data.	[30], [34]
HMP	Probabilistic model	Markov chain	Modeling and calculation are simple.	Performance depends on historical data. Only one kind of information (i.e. cell sequence) can be processed to forecast next cell.	Regular movements; sufficient historical data; simple network environment.	[13], [36]–[39]
		HMM	Indirect factors can be utilized to determine next cell; higher prediction accuracy than Markov chain.	Sufficient historical data is needed; relatively high complexity.	Regular movements; sufficient historical data. There exist indirect factors that will influence target cell.	[14], [41]–[49]
		Bayesian network	It can cope with lots of impact factors.	The process of model construction is cumbersome; sufficient historical data is needed.	Regular movements; sufficient historical data; complex environments with a lot of impact factors.	[15], [33], [50]–[52]
	Discriminative model	SVM	It has good generalization capabilities which prevent it from over-fitting. Handles non-linear data efficiently. Stability.	Not friendly to large-scale data; sufficient historical data is needed.	Regular movements; small samples, nonlinear, and high dimensional problems.	[12], [16], [31], [59]
		ANN	Good performance. It can cope with large data scale and unclear information.	High computational complexity; long training time; sufficient historical data is needed.	Regular movements; sufficient historical data; large data scale.	[17], [23], [61]–[66]
	Data mining	Association rules	Based on big data, it is able to find the hidden links between different data.	High computational complexity; sufficient historical data is needed.	Regular movements; sufficient historical data; massive, noisy, incomplete, and random data.	[71]–[74]
CMS and HMP	HDA		High accuracy. It can perform long-term and short-term mobility prediction. To some extent, both random and regular movements can be predicted.	High computational complexity and large overhead on UE.	Sufficient historical data; sufficient system resources.	[21], [32], [81]–[87]

be analyzed. Similarly, the authors in [46] consider using next-cell prediction to enhance handover performance between Wi-Fi APs and cellular networks, and present an enhanced handover mechanism with mobility prediction. In the test scenarios, including homogeneous (i.e. handover between Wi-Fi APs) and heterogeneous (i.e. handover between Wi-Fi and cellular network) networks, their scheme can improve network throughput and decrease retransmission rate.

In handover region, it may be covered by multiple BSs. Accordingly, it is necessary to determine the most suitable BS to connect. Based on accurate users' movement, BS controller can management next-cell prediction and further encourage mobile users to connect to a better BS. Many introduced prediction approaches in last Section can be used to optimize BS selection [17], [44], [56], [85]. Based on user mobility information, the authors in [89] propose a cell selection algorithm, which first predicts candidate next cells and further chooses the one with maximum SINR as the target cell. This scheme reduces the number of handovers and mitigates the degradation of communications quality. Similarly, the authors in [90] propose a scheme to determine the next femto access point based on mobility prediction integrated with received power and quality of reference signal. Their results show that the proposed scheme can increase handover success rate.

Because of the extensive deployment of BSs (or APs), a large number of handovers may occur. More handover leads to more power consumption of UE and decreases system efficiency. Hence, it is necessary to reduce unnecessary handover. Meanwhile, due to the nature of user movement and signal fluctuation at cell border, UE may perform frequent handover back and forth in a short time between serving and adjacent cells. This potentially undesirable phenomenon is known as ping-pong effect [91], which severely increases traffic delay, consumes more power, and degrades network performance. This problem can be alleviated by next-cell prediction, which can reduce unnecessary handover and relieve ping-pong effect [92], [93]. For example, the authors in [94] set a power offset based on mobility prediction, to determine target BS and avoid unnecessary handovers. Since traditional handover decisions based on the received quality of reference symbols will lead to the ping-pong effect, the authors in [95] introduced Vehicular Location Prediction Handover Algorithm, which considering two parameters, i.e. predicted cell and reference symbols received quality, to reduce unnecessary handover and ping-pong effect.

If network operators cannot know users' mobility information in advance, they can only perform traditional handover. In this situation, mobile users may suffer from undesirable termination due to resource inefficiency, unnecessary handover, and ping-pong effect. As described above, the prediction of user motion can improve handover performance undoubtedly. Since it is unlikely that a user will hand over frequently with short intervals, real-time prediction is unnecessary in general. Besides, handover management has relatively low requirements for prediction accuracy. Although other prediction techniques with high precision or real-time features can also be applied to handover management, they usually have higher computation complexity or require more feedback from users, which bring extra processing and energy overhead to both the access point and users. In such scenarios, next-cell prediction has its special advantages. Since network operators only need to know which cells the users will enter into, next-cell prediction is good enough, in terms of both precision and delay tolerance. Therefore, next-cell prediction is able to strike a good balance between precision and cost, which makes it suitable for handover management.

2) LOAD BALANCING

In an imbalance scenario, some cells may not have enough resources to support all requests, while other lightly loaded neighboring cells still have extra unused resources. An imbalanced traffic load of wireless networks will cause longer packet delay and throughput degradation. Load balancing based on next-cell prediction can result in a tremendous increase in network performance under high load [59].

The works of [96]–[98] introduce load balancing schemes based on next-cell prediction. In particular, the authors in [96] propose a mobility prediction technique to solve load balancing problem through an adaptive handover approach. The proposed approach adaptive sets handover hysteresis threshold for different neighboring cells based on both signal strength and load information, where mobility prediction is used to reduce the unnecessary handovers. Except for using handover for load balancing, network operators can also achieve load balance via content caching management based on cell prediction. A Proactive Load Balancing (PLB) framework is investigated in [97]. Specifically, the authors exploit users' trajectory to predict their future crossing cells and model users' content profile to predict their most expected future data. The PLB framework proactively caches users' future contents during their stay at lightly loaded cells, considering cell prediction and data demand jointly. Their results show that the proposed scheme can improve cell load fairness. Similarly, the works of [98] investigate a proactive load balancing method called OPERA to solve the imbalance issue between macro and small cells. OPERA estimates users' future cells and further predicts future loads of the cells, then proactively optimizes key antenna parameters and cell individual offsets to preempt congestion before it happens.

In order to balance the load of BSs, user's long-term mobility information (i.e. subsequently traversed cells) needs to be predicted. Therefore, long-term forecast schemes, such as ANN, association rules, etc. should be applied.

3) RESOURCE ALLOCATION

To maintain a continuous service during handover, the bandwidth of target cell must be reserved in advance. Otherwise, it may decrease the Quality of Service (QoS) or terminate the connection. Traditional fixed reservation usually permanently reserve a part of bandwidth in all neighboring cells. However, when a user enters into one of the adjacent cells, the bandwidth reserved in other neighboring cells will be wasted. Thus, adaptive reservation should be implemented to improve bandwidth utilization efficiency. Through next-cell prediction, the target cell can be determined, and an adaptive reservation can be derived. It reserves the required bandwidth only in those cells where the users are expected to visit in the near future [99], [100]. Also, BSs compute the amount of bandwidth need to be reserved based on cell estimation to reduce unnecessary bandwidth reservation [29]. To some extent, if user mobility can be predicted, the network load can also be predicted. The authors of [101] solve the problem of user association and resource allocation in virtual small cell aided multi-tier heterogeneous networks based on mobility prediction.

In addition to allocating bandwidth, BSs power can also be dynamically configured based on mobility prediction [102]. In a high-speed railway wireless communication environment, the work of [103] study a scheme which uses Fuzzy C-Means algorithm and user mobility prediction model to classify mobile users into center and edge users. Based on classification results, differential power distribution schemes are implemented for center and edge users. This scheme effectively manages interference and significantly improves network performance.

In wireless cellular networks, radio resources such as bandwidth, time slots, antennas and transmit power, can be dynamically allocated based on context information [104]. It can be performed on different levels, i.e. inter-cell and inner-cell levels. From inter-cell level, all the users inside a cell are considered a whole. The resource allocation puts more emphasis on reserving enough resources for new coming users in advance. Therefore, it does not need precise and real-time prediction. For rural cells, which in general have large coverage and relatively rare handovers, even simple next-cell prediction methods, such as Markov chain, can meet the requirements. However, for urban and sub-urban cells, which usually have smaller cell radius, they have to deal with handover frequently with higher accuracy. Next-cell prediction methods with higher precision, such as CMAS and HDA, should be applied. On the contrary, from inter-cell level, mobility-aware resource allocation usually treats users inside a cell as individuals and considers more detailed management problems, such as the assignment of sub-channels or scheduling of time slots [105]. Both high precision and near-real-time features are required when making position and trajectory prediction. For such applications, next-cell prediction has its limitation and cannot be applied.

4) LOCATION-BASED SERVICE

Providing a sequence of cell trajectory, next-cell prediction can be efficiently applied in Location-Based Services (LBSs) [71], [106], such as advertising recommendation, content prefetching, and warning for traffic jams.

LBSs usually need a relatively accurate position of mobile users. It is important to exploit positioning mechanism with less energy consumption. Based on mobility prediction results, it is reasonable to determine users' current position. In [107], the authors introduce a positioning mechanism to reduce energy consumption by utilizing mobility prediction in mobile networks. This scheme proposes that disabling the high power GPS based positioning and determining current position with mobility prediction to save battery usage when a mobile user follows his/her historical mobility pattern. Otherwise, GPS is used to retrieve the current true position as usual. This solution has been evaluated in real-life for a period of three months. Results show that the energy consumption is reduced 60% compared to continuous GPS positioning, and accuracy is increased by around 76% compared to network based positioning only. Further, [108] proposes a novel localization method based on neighbor RSS and mobility prediction to provide higher accuracy of localization with lower calibration requirements in smart building environments.

The predicted information can be further applied to location-based social activity predictions. Due to festivals or holidays, a large number of mobile users may gather to form a crowd, which poses high a load situation to the respective serving BS and leads to congestion possibly. The work of [109] predicts crowd formation using the users' mobility pattern in neighboring cells. Similarly, the authors of [110] analyze the mobility behavior of vehicular users to predict the traffic status of a cell. This information of crowd formation and road traffic status can be used to proactively trigger load balancing schemes and design efficient radio resource management techniques.

Through cell prediction, content recommendation can be performed as well. In recent years, short video application has been growing rapidly. In order to improve the Quality of Experience (QoE) in high-speed scenarios, the work of [111] investigates a short video recommendation approach based on cell prediction. If the network operator knows the user's interests and future crossing cells in advance, it can push a user's preferred short video content to the most likely BS that user will be connected to. Their experiments show that better recommendations can be provided and the waiting time for short video will be eliminated.

For LBSs, assist positioning mechanism needs an accurate prediction model. While, the precision requirements of crowd formation and content recommendation are relatively lower, but they expect long-term estimation. Hence, prediction models should be considered according to different scenarios.

In summary, next-cell prediction algorithms play a crucial role in wireless cellular networks. The main applications of next-cell prediction and suitable prediction models are summarized in Table 2.

B. CURRENT STATUS AND CHALLENGES

Next-cell prediction has great potential to estimate mobile users' future trajectory, in terms of crossing cells, and has been intensively applied in many fields, e.g. handover management, resource allocation, and LBSs. However, there still exist many challenges with the rapid development of wireless networks. In this subsection, we discuss some problems

TABLE 2. The main applications of next-cell prediction.

Applications	Descriptions
Handover management	Enhanced handover strategies, selections of optimal target handover BS (or AP), and ping-pong effect mitigation.
Load balancing	Proactively caching users' future contents during their stay at lightly loaded cells.
Resource allocation	Resource allocation on bandwidth, time slots, transmit antennas, and transmit power.
Location-based service	Assisted positioning mechanism, crowd formation prediction, warning for traffic jams, and content recommendation.

which can be improved and several potential future research directions.

- More information could be utilized, e.g. social activities, weather, and traffic status, instead of movements parameters only, to further improve prediction accuracy;
- Fast and efficient predictions need to be studied. The prediction should be completed before users enter new a cell, for high data rate networks (i.e. 5G cellular network) with high speed environments (i.e. high speed railway). Therefore, balancing prediction accuracy and computational complexity are equally important;
- Pay more attention to random movements prediction. Regular users can be predicted very well using history profile. However, accurate prediction of users' random movements is still a difficult problem. Besides, most current random movements predictors need to monitor users' positions in real-time, which leads to high energy consumption. So researchers may work on a more accurate and energy efficient approach for next-cell predictions of random movements;
- More consideration of spatiotemporal prediction. Ideal cell prediction should estimate where users are heading to and when users will arrive. While many existing schemes, e.g. Markov chains, can predict cell transition well but cannot estimate dwell time;
- Consider data incompleteness and errors. Lots of current prediction schemes are based on GPS data to determine user position. However, affected by buildings or obstacles, indoor positioning of GPS cannot work well, and part of data may lose or be error. This severely impact prediction results, which is not considered in most works;
- Privacy and data security. In the process of cell prediction, mobile users' trajectory data may face the risk of disclosure. This may threat the security of users. Thus, network operator need more carefully to handle mobility prediction to prevent users' personal data leakage;
- Explore more applications based on next-cell prediction in the future.
- Standardization activity regarding next-cell prediction could be considered in the future.

VIII. CONCLUSION

This article provides a thorough survey of next-cell prediction problem in cellular networks. We first introduce concerns and characteristics of next-cell prediction, including predictability of the next cell, how to obtain user mobility information, classification of mobility data, and classification of prediction methods. To better understand and manage next-cell prediction models and methods, we propose a classification based on types of mobility data used for prediction (i.e. CMSA, HMPA, and HDA). CMSA needs to monitor users' movements parameters such as moving direction and speed in real-time, to make a short time prediction and determine adjacent cell it will be connected to, which can handle random movements but energy costing. Whereas, HMPA exploits user movement history to mine movement patterns and predicts future crossing cells. HMPA can predict regular movement very well and be extensively studied but poor random capacity. HDA considers both types of data to carry out estimation. Theoretically, it can get a better accuracy performance, as it can make both short-term and long-term predictions. Lots of prediction approaches, enveloping angle-based, distance-based, Markov chain, HMM, Bayesian network, ANN, SVM, etc., can be used to predict the next cell, and have been introduced in detail in the paper. Then the applications of cell prediction are presented, especially for handover management, resource allocation, and LBSs. Finally, we discuss current existed issues and future potential research directions.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this article.

REFERENCES

[1] A. Nadembega, A. Hafid, and T. Taleb, "Mobility-prediction-aware bandwidth reservation scheme for mobile networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2561–2576, Jun. 2015.
[2] F. De Rango, P. Fazio, and S. Marano, "Utility-based predictive services for adaptive wireless networks with mobile hosts," *IEEE Trans. Veh. Technol.*, vol. 58, no. 3, pp. 1415–1428, Mar. 2009.
[3] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, Feb. 2010.
[4] G. Xu, S. Gao, M. Daneshmand, C. Wang, and Y. Liu, "A survey for mobility big data analytics for geolocation prediction," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 111–119, Feb. 2017.
[5] N. Bui, M. Cesana, S. A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1790–1821, 3rd Quart., 2017.
[6] S. Xu, X. Fu, J. Cao, B. Liu, and Z. Wang, "Survey on user location prediction based on geo-social networking data," *World Wide Web*, vol. 23, pp. 1621–1664, Jan. 2020.
[7] V. Koolwal and K. K. Mohbey, "A comprehensive survey on trajectory-based location prediction," *Iran J. Comput. Sci.*, vol. 3, no. 2, pp. 65–91, Jun. 2020.

[8] P. Fazio, F. De Rango, and M. Tropea, "Prediction and QoS enhancement in new generation cellular networks with mobile hosts: A survey on different protocols and conventional/unconventional approaches," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1822–1841, 3rd Quart., 2017.

[9] H. Zhang and L. Dai, "Mobility prediction: A survey on state-of-the-art schemes and future applications," *IEEE Access*, vol. 7, pp. 802–822, 2019.

[10] R. Doss, A. Jennings, and N. Shenoy, "A review on current work in mobility prediction for wireless networks," in *Proc. 3rd Asian Int. Mobile Comput. Conf.*, 2004, pp. 1–7.

[11] Y. Li, D. Jin, P. Hui, Z. Wang, and S. Chen, "Limits of predictability for large-scale urban vehicular mobility," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2671–2682, Dec. 2014.

[12] J. Chen, L. Ma, and Y. Xu, "Support vector machine based mobility prediction scheme in heterogeneous wireless networks," *Math. Problems Eng.*, vol. 2015, pp. 1–10, Jan. 2015.

[13] A. Hasbollah, S. H. S. Ariffin, and N. Fisal, "Mobility prediction method for vehicular network using Markov chain," *Jurnal Teknologi*, vol. 78, nos. 6–2, Jun. 2016.

[14] H. Si, Y. Wang, J. Yuan, and X. Shan, "Mobility prediction in cellular network using hidden Markov model," in *Proc. 7th IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2010, pp. 1–5.

[15] Y. Qiu, X. Wang, F. Wang, and S. Bian, "Intelligent user profile prediction in radio access network," in *Signal and Information Processing, Networking and Computers* (Lecture Notes in Computer Science). Singapore: Springer, 2018, pp. 437–445.

[16] J. Yang, C. Dai, and Z. Ding, "A scheme of terminal mobility prediction of ultra dense network based on SVM," in *Proc. IEEE 2nd Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2017, pp. 837–842.

[17] C. P. Koushik and P. Vetrivelan, "Heuristic relay-node selection in opportunistic network using RNN-LSTM based mobility prediction," *Wireless Pers. Commun.*, vol. 114, no. 3, pp. 2363–2388, Oct. 2020.

[18] S. Chatterjee, D. Sarddar, J. Saha, S. Banerjee, A. Mondal, and M. K. Naskar, "An improved mobility management technique for IEEE 802.11 based WLAN by predicting the direction of the mobile node," in *Proc. Nat. Conf. Comput. Commun. Syst. (NCCCS)*, Nov. 2012, pp. 1–5.

[19] X. Shen, J. W. Mark, and J. Ye, "User mobility profile prediction: An adaptive fuzzy inference approach," *Wireless Netw.*, vol. 6, no. 5, pp. 363–374, 2000.

[20] M. Nkambule, G. Ojong, and M. Adigun, "Predicting mobility on the wireless Internet using sectored-cell approach," in *Proc. SATNAC*, 2004, pp. 1–6.

[21] R. Chellappa, A. Jennings, and N. Shenoy, "The sectorized mobility prediction algorithm for wireless networks," in *Proc. Int. Conf. Inf. Commun. Technol.*, 2003, pp. 8–11.

[22] M. Al-Sanabani, S. Shamala, M. Othman, and Z. Zukarnain, "Mobility prediction based resource reservation for handoff in multimedia wireless cellular networks," *Int. Arab J. Inf. Technol.*, vol. 5, no. 2, pp. 162–169, 2008.

[23] S.-C. Liou and Y.-M. Huang, "Trajectory predictions in mobile networks," *Int. J. Inf. Tech. Decis.*, vol. 11, no. 11, pp. 109–122, 2005.

[24] M. Page and T. L. Wickramarathne, "Enhanced situational awareness with signals of opportunity: RSS-based localization and tracking," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 3833–3838.

[25] L. A. M. Hernández, S. P. Arteaga, G. S. Pérez, A. L. S. Orozco, and L. J. G. Villalba, "Outdoor location of mobile devices using trilateration algorithms for emergency services," *IEEE Access*, vol. 7, pp. 52052–52059, 2019.

[26] A. Jayasuriya and J. Asenstorfer, "Mobility prediction model for cellular networks based on observed traffic patterns," in *Proc. IASTED Int. Conf. Wireless Opt. Commun. (WOC)*, 2002, pp. 386–391.

[27] L. Lee, M. Jones, G. S. Ridenour, S. J. Bennett, A. C. Majors, B. L. Melito, and M. J. Wilson, "Comparison of accuracy and precision of gps-enabled mobile devices," in *Proc. IEEE Int. Conf. Comput. Inf. Technol. (CIT)*, Dec. 2016, pp. 73–82.

[28] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2011, pp. 1082–1090.

[29] D.-S. Lee and Y.-H. Hsueh, "Bandwidth-reservation scheme based on road information for next-generation cellular networks," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 243–252, Jan. 2004.

[30] N. P. Kuruvatti, A. Klein, J. Schneider, and H. D. Schotten, "Exploiting diurnal user mobility for predicting cell transitions," in *Proc. IEEE Globecom Workshops, GC Wkshps*, Dec. 2013, pp. 293–297.

[31] X. Chen, F. Mériaux, and S. Valentin, "Predicting a user's next cell with supervised learning based on channel states," in *Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 36–40.

[32] W.-S. Soh and H. S. Kim, "QoS provisioning in cellular networks based on mobility prediction techniques," *IEEE Commun. Mag.*, vol. 41, no. 1, pp. 86–92, Jan. 2003.

[33] H. Xiong, D. Zhang, D. Zhang, and V. Gauthier, "Predicting mobile phone user locations by exploiting collective behavioral patterns," in *Proc. IEEE Int. Conf. Ubiquitous Intell. Comput. IEEE Int. Conf. Auton. Trusted Comput. (UIC-ATC)*, Sep. 2012, pp. 164–171.

[34] B. Li, H. Zhang, and H. Lu, "User mobility prediction based on Lagrange's interpolation in ultra-dense networks," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–6.

[35] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility Markov chains," in *Proc. 1st Workshop Meas., Privacy, Mobility*, 2012, pp. 1–6.

[36] P. Fulop, S. Szabo, and T. Szalka, "Location prediction methods with Markovian approach and extended random walk model," in *Proc. Int. Conf. Comput. Inf. Syst. Ind. Manage. Appl. (CISIM)*, 2007, pp. 53–58.

[37] A. Ulvan, M. Ulvan, and R. Bestak, "The enhancement of handover strategy by mobility prediction in broadband wireless access," in *Proc. Netw. Electron. Commer. Res. Conf.*, 2009, pp. 266–276.

[38] P. S. Prasad and P. Agrawal, "Mobility prediction for wireless network resource management," in *Proc. Annu. Southeast Symp. Syst. Theory*, 2009, pp. 98–102.

[39] S. Bellahsene and L. Kloul, "A new Markov-based mobility prediction algorithm for mobile networks," in *Computer Performance Engineering* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2010, pp. 37–50.

[40] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *IEEE ASSP Mag.*, vol. 3, no. 1, pp. 4–16, Jan. 1986.

[41] C.-F. Wu, L.-T. Lee, and D.-F. Tao, "An HMM prediction and throttling-based call admission control scheme for wireless multimedia networks," *Comput. Math. Appl.*, vol. 54, no. 3, pp. 364–378, Aug. 2007.

[42] A. Ben Cheikh, M. Ayari, R. Langar, G. Pujolle, and L. A. Saidane, "Optimized handoff with mobility prediction scheme using HMM for femtocell networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 3448–3453.

[43] A. B. Cheikh, M. Ayari, R. Langar, and L. A. Saidane, "OHMP-CAC: Optimized handoff scheme based on mobility prediction and qos constraints for femtocell networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf., IWCMC*, Sep. 2016, pp. 936–941.

[44] K.-L. Yap and Y.-W. Chong, "Optimized access point selection with mobility prediction using hidden Markov model for wireless network," in *Proc. Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2017, pp. 38–42.

[45] N. AlJeri and A. Boukerche, "An efficient movement-based handover prediction scheme for hierarchical mobile IPv6 in VANETs," in *Proc. ACM Int. Symp. Perform. Eval. Wireless Ad Hoc, Sensor, Ubiquitous Netw.*, 2018, pp. 47–54.

[46] K.-L. Yap, Y.-W. Chong, and W. Liu, "Enhanced handover mechanism using mobility prediction in wireless networks," *PLoS ONE*, vol. 15, no. 1, Jan. 2020, Art. no. e0227982.

[47] P. S. Prasad and P. Agrawal, "Movement prediction in wireless networks using mobility traces," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2010, pp. 1–5.

[48] P. Fazio and S. Marano, "Mobility prediction and resource reservation in cellular networks with distributed Markov chains," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, Aug. 2012, pp. 878–882.

[49] P. Fazio and S. Marano, "A new Markov-based mobility prediction scheme for wireless networks with mobile hosts," in *Proc. Int. Symp. Perform. Eval. Comput. Telecomm. Syst. (SPECTS)*, 2012, pp. 1–5.

[50] K. Schougaard, "Vehicular mobility prediction by Bayesian networks," DAIMI, Copenhagen, Denmark, Tech. Rep. 582, 2007.

[51] Y. Zhang, J. Hu, J. Dong, Y. Yuan, J. Zhou, and J. Shi, "Location prediction model based on Bayesian network theory," in *Proc. GLOBECOM IEEE Global Telecommun. Conf.*, Nov. 2009, pp. 1–6.

[52] Y. Yuan, Y. Zhang, Y. Huang, X. Jin, X. Zheng, and J. Shi, "An adaptive wireless paging scheme using Bayesian network location prediction model," in *Proc. IEEE Veh. Technol. Conf. Fall*, Sep. 2009, pp. 1–5.

[53] H. Abu-Ghazaleh and A. S. Alfa, "Mobility prediction and spatial-temporal traffic estimation in wireless networks," in *Proc. IEEE Veh. Technol. Conf.*, May 2008, pp. 2203–2207.

[54] H. Abu-Ghazaleh and A. S. Alfa, "Application of mobility prediction in wireless networks using Markov renewal theory," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 788–802, Feb. 2010.

[55] H. Farooq and A. Imran, "Spatiotemporal mobility prediction in proactive self-organizing cellular networks," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 370–373, Feb. 2017.

[56] W. Fu, X. Li, H. Ji, and H. Zhang, "Optimal access points selection based on mobility prediction in heterogeneous small cell networks," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2018, pp. 421–425.

[57] P. N. Pathirana, A. V. Savkin, and S. Jha, "Robust extended Kalman filter applied to location tracking and trajectory prediction for PCS networks," in *Proc. IEEE Int. Conf. Control Appl.*, vol. 1, Sep. 2004, pp. 63–68.

[58] H. Li and G. Ascheid, "Mobility prediction based on graphical model learning," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2012, pp. 1–5.

[59] S. Michaelis, "Balancing high-load scenarios with next cell predictions and mobility pattern recognition," in *Proc. IEEE Int. Conf. Mobile Services*, 2011, pp. 164–169.

[60] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[61] S.-C. Liou and H.-C. Lu, "Applied neural network for location prediction and resources reservation scheme in wireless networks," in *Proc. Int. Conf. Commun. Technol. (ICCT)*, vol. 2, 2003, pp. 958–961.

[62] M. Stoyanova and P. Mahonen, "A next-move prediction algorithm for implementation of selective reservation concept in wireless networks," in *Proc. IEEE 18th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2007, pp. 1–5.

[63] P. Fazio, F. De Rango, and I. Selvaggi, "A novel passive bandwidth reservation algorithm based on neural networks path prediction in wireless environments," in *Proc. Int. Symp. Perform. Eval. Comput. Telecommun. Syst. (SPECTS)*, 2010, pp. 38–43.

[64] D. S. Wickramasuriya, C. A. Perumalla, K. Davaslioglu, and R. D. Gitlin, "Base station prediction and proactive mobility management in virtual cells using recurrent neural networks," in *Proc. IEEE Wireless Microw. Technol. Conf. (WAMICON)*, Apr. 2017, pp. 1–6.

[65] S. S. Shaabanzadeh and J. Sánchez-González, "On the prediction of future user connections based on historical records in wireless networks," in *Proc. IFIP Adv. Inf. Commun. Technol.*, 2020, pp. 84–94.

[66] M. Manalastas, H. Farooq, S. M. Asad Zaidi, and A. Imran, "Where to go next: A realistic evaluation of AI-assisted mobility predictors for HetNets," in *Proc. IEEE 17th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2020, pp. 1–6.

[67] D. Katsaros, A. Nanopoulos, M. Karakaya, G. Yavas, O. Ulusoy, and Y. Manolopoulos, "Clustering mobile trajectories for resource allocation in mobile environments," in *Advances in Intelligent Data Analysis V* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2003, pp. 319–329.

[68] T. Anagnostopoulos, C. Anagnostopoulos, and S. Hadjiefthymiades, "Efficient location prediction in mobile cellular networks," *Int. J. Wireless Inf. Netw.*, vol. 19, no. 2, pp. 97–111, Jun. 2012.

[69] S. Toufga, S. Abdellatif, P. Owezarski, T. Villemur, and D. Relizani, "Effective prediction of V2I link lifetime and Vehicle's next cell for software defined vehicular networks: A machine learning approach," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Dec. 2019, pp. 1–8.

[70] M. Daoui, M. Belkadi, L. Chamek, M. Lalam, and A. Berqia, "Mobility prediction and location management based on data mining," in *Proc. Next Gener. Netw. Services*, 2013, pp. 137–140.

[71] T. H. N. Vu, J. W. Lee, and K. H. Ryu, "Spatiotemporal pattern mining technique for location-based service system," *ETRI J.*, vol. 30, no. 3, p. 421–431, 2008.

[72] T. Duong, D. Q. Tran, and C. H. Tran, "Spatiotemporal data mining for mobility prediction in wireless network," in *Proc. Nat. Conf. Fundam. Appl. IT Res.*, 2011, pp. 224–235.

[73] T. V. T. Duong and D. Q. Tran, "An effective approach for mobility prediction in wireless network based on temporal weighted mobility rule," *Int. J. Comput. Sci. Telecommun.*, vol. 3, no. 2, pp. 29–36, 2012.

[74] A. K. Gupta and U. Shanker, "SPMC-PRRP: A predicted region based cache replacement policy," in *Advances in Data and Information Sciences*. Cham, Switzerland: Springer, 2019, pp. 313–326.

[75] F. Yu and V. Leung, "Mobility-based predictive call admission control and bandwidth reservation in wireless cellular networks," *Comput. Netw.*, vol. 38, no. 5, pp. 577–589, Apr. 2002.

[76] J.-I. Jung, J. Kim, and Y. You, "Mobility prediction handover using user mobility pattern and guard channel assignment scheme," in *Proc. Eur. Conf. Universial Multiservice Netw.*, 2004, pp. 155–164.

[77] D. Barth, S. Bellahsene, and L. Kloul, "Combining local and global profiles for mobility prediction in LTE femtocells," in *Proc. ACM Int. Conf. Modeling, Anal. Simulation Wireless Mobile Syst. (MSWiM)*, 2012, pp. 333–342.

[78] M. Daoui, A. M'Zoughi, M. Lalam, M. Belkadi, and R. Aoudjit, "Mobility prediction based on an ant system," *Comput. Commun.*, vol. 31, no. 14, pp. 3090–3097, Sep. 2008.

[79] A. I. M. Saleh, "A hybrid mobility prediction (HMP) strategy for PCS networks," *Pattern Anal. Appl.*, vol. 19, no. 1, pp. 173–206, 2016.

[80] M. Lv, D. Zeng, T. Chen, and L. Chen, "A sequence-to-sequence model for cell-id trajectory prediction," in *Proc. UbiComp/ISWC-Adjunct ACM Int. Joint Conf. Pervasive Ubiquitous Comput. ACM Int. Symp. Wearable Comput.*, 2019, pp. 137–140.

[81] C. Mala, M. Loganathan, N. Gopalan, and B. Sivaselvan, "A novel genetic algorithm approach to mobility prediction in wireless networks," in *Proc. Int. Conf. Contemp. Comput. (IC3)*, 2009, pp. 49–57.

[82] I. F. Akyildiz and W. Wang, "The predictive user mobility profile framework for wireless multimedia networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 6, pp. 1021–1035, Dec. 2004.

[83] S. Bellahsene, L. Kloul, and D. Barth, "A hierarchical prediction model for two nodes-based ip mobile networks," in *Proc. ACM Int. Conf. Modeling, Anal. Simulation Wireless Mobile Syst. (MSWiM)*, 2009, pp. 173–180.

[84] S. E. Abdalla and S. H. S. Ariffin, "A hybrid model for user history-based prediction with geolocation assisted handover in 5G," *Int. J. Simul.-Syst., Sci. Technol.*, vol. 18, no. 4, pp. 1–6, 2017.

[85] R. Ahmad, E. A. Sundararajan, N. E. Othman, and M. Ismail, "Efficient handover in LTE-A by using mobility pattern history and user trajectory prediction," *Arabian J. Sci. Eng.*, vol. 43, no. 6, pp. 2995–3009, Jun. 2018.

[86] M. Goodarzi, N. Maletic, J. Gutiérrez, V. Sark, and E. Grass, "Next-cell prediction based on cell sequence history and intra-cell trajectory," in *Proc. 22nd Conf. Innov. Clouds, Internet Netw. Workshops (ICIN)*, Feb. 2019, pp. 257–263.

[87] A. I. Saleh, M. S. Elkasas, and A. A. Hamza, "Ant colony prediction by using sectorized diurnal mobility model for handover management in PCS networks," *Wireless Netw.*, vol. 25, no. 2, pp. 765–775, 2019.

[88] A. Ulvan, R. Bestak, and M. Ulvan, "Handover procedure and decision strategy in LTE-based femtocell network," *Telecommun. Syst.*, vol. 52, no. 4, pp. 2733–2748, Apr. 2013.

[89] A. Kishida, Y. Morihiro, T. Asai, and Y. Okumura, "Cell selection scheme for handover reduction based on moving direction and velocity of UEs for 5G multi-layered radio access networks," in *Proc. Int. Conf. Inf. Netw.*, Jan. 2018, pp. 362–367.

[90] M. Mandour, F. Gebali, A. D. Elbayoumy, G. M. A. Hamid, and A. Abdelaziz, "Handover optimization and user mobility prediction in LTE femtocells network," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2019, pp. 1–6.

[91] Z. Feher, A. Veres, and Z. Heszberger, "Ping-pong reduction using sub cell movement detection," in *Proc. IEEE 75th Veh. Technol. Conf.*, May 2012, pp. 1–5.

[92] T.-H. Kim, Q. Yang, J.-H. Lee, S.-G. Park, and Y.-S. Shin, "A mobility management technique with simple handover prediction for 3G LTE systems," in *Proc. IEEE 66th Veh. Technol. Conf.*, Sep. 2007, pp. 259–263.

[93] J. Gu, S. J. Bae, M. Y. Chung, K.-Y. Cheon, and A.-S. Park, "Mobility-based handover decision mechanism to relieve ping-pong effect in cellular networks," in *Proc. 16th Asia–Pacific Conf. Commun. (APCC)*, Oct. 2010, pp. 487–491.

[94] R. Shi, Y. Peng, and L. Zhang, "A user mobility prediction method to reduce unnecessary handover for ultra dense network," in *Proc. 28th Wireless Opt. Commun. Conf. (WOCC)*, May 2019, pp. 1–5.

[95] A. A. Hasbollah, S. H. Ariffin, and N. E. Ghazali, "Optimal forwarding probability for vehicular location prediction handover algorithm," in *Proc. Commun. Comput. Inf. Sci.*, 2017, pp. 369–380.

[96] Q.-P. Yang, J.-W. Kim, and T.-H. Kim, "Mobility prediction and load balancing based adaptive handovers for LTE systems," *Int. J. Comput. Sci. Eng.*, vol. 4, no. 4, p. 638, 2012.

[97] S. Manzoor, S. Mazhar, A. Asghar, A. Noor Mian, A. Imran, and J. Crowcroft, "Leveraging mobility and content caching for proactive load balancing in heterogeneous cellular networks," *Trans. Emerg. Telecommun. Technol.*, vol. 31, no. 2, p. e3739, Feb. 2020.
[98] H. Farooq, A. Asghar, and A. Imran, "Mobility prediction based proactive dynamic network orchestration for load balancing with QoS constraint (OPERA)," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3370–3383, Mar. 2020.
[99] M. A. Al-Sanabani, S. K. Shamala, M. Othman, and Z. A. Zukarnain, "Multi-class bandwidth reservation scheme based on mobility prediction for handoff in multimedia wireless/mobile cellular networks," *Wireless Pers. Commun.*, vol. 46, no. 2, pp. 143–163, Jul. 2008.
[100] P. Fazio, M. Tropea, F. D. Rango, and M. Voznak, "Pattern prediction and passive bandwidth management for hand-over optimization in QoS cellular networks with vehicular mobility," *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2809–2824, Nov. 2016.
[101] Z. Cheng, N. Chen, B. Liu, Z. Gao, L. Huang, X. Du, and M. Guizani, "Joint user association and resource allocation in HetNets based on user mobility prediction," *Comput. Netw.*, vol. 177, Aug. 2020, Art. no. 107312.
[102] L. Huang, C. W. Sung, and C. S. Chen, "Context-aware wireless broadcast for next generation cellular networks," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Nov. 2014, pp. 349–353.
[103] J. Sheng, Z. Tang, Q. Zhu, C. Wu, Y. Wang, and B. Ai, "An improved interference alignment algorithm with user mobility prediction for high-speed railway wireless communication networks," *IEEE Access*, vol. 8, pp. 80468–80479, 2020.
[104] A. Ahmad, S. Ahmad, M. H. Rehmani, and N. U. Hassan, "A survey on radio resource allocation in cognitive radio sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 888–917, 2nd Quart., 2015.
[105] L. Lu, L. Huang, and W. Hua, "Energy efficient resource allocation and scheduling for delay-constrained multi-cell broadcast networks," *IEEE Access*, vol. 8, pp. 164844–164857, 2020.
[106] M. Vukovic, G. Vujnovic, and D. Grubisic, "Adaptive user movement prediction for advanced location-aware services," in *Proc. Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, 2009, pp. 343–347.
[107] X. Cai, L. Westberg, A. Mihály, and J. Christoffersson, "Prediction based energy efficient mobile positioning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 2593–2598.
[108] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1294–1307, Jul. 2016.
[109] N. P. Kuruvatti, A. Klein, and H. D. Schotten, "Prediction of dynamic crowd formation in cellular networks for activating small cells," in *Proc. IEEE Veh. Technol. Conf.*, May 2015, pp. 1–5.
[110] N. P. Kuruvatti, J. F. S. Molano, and H. D. Schotten, "Monitoring vehicular user mobility to predict traffic status and manage radio resources," in *Proc. IEEE Veh. Technol. Conf.*, Jun. 2017, pp. 1–6.
[111] Q. Li, Y. Zhang, H. Huang, and J. Yan, "Deep learning-based short video recommendation and prefetching for mobile commuting users," in *Proc. ACM SIGCOMM Workshop Netw. Emerg. Appl. Technol. (SIGCOMM)*, 2019, pp. 49–55.

LINYU HUANG (Member, IEEE) received the Ph.D. degree in electronic engineering from the City University of Hong Kong, in 2014. He joined the faculty as a Lecturer with the College of Electronics and Information Engineering, Sichuan University, Chengdu, China, in 2014. His research interests include wireless communications, resource allocation, power control, and network coding.

LIANG LU received the B.Sc. degree from the School of Information Engineering, Shaoyang University, Shaoyang, China, in 2018. He is currently pursuing the M.Sc. degree with the College of Electronics and Information Engineering, Sichuan University, Chengdu, China. His research interests include mobility prediction and resource allocation in wireless networks.

WEI HUA received the B.Sc. and M.Sc. degrees and the Ph.D. degree in biomedical engineering from Sichuan University, Chengdu, China, in 1990, 1997, and 2006, respectively. He is currently an Associate Professor with the College of Electronics and Information Engineering, Sichuan University. His research interests include multilateration, power line communications, and wireless communications.