

Proactive Mobility Management based on Virtual Cells in SDN-enabled Ultra-dense Networks

Qian Liu^{1,2,3}, Gang Chuai^{1,2}, Jingrong Wang³, Jianping Pan³, Weidong Gao^{1,2}, and Xuewen Liu^{1,2}

¹Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing, China

²Beijing University of Posts and Telecommunications, Beijing, China

³Department of Computer Science, University of Victoria, Victoria, BC, Canada

Abstract—Ultra-dense networking (UDN) is a promising technology to improve the network capacity in the next-generation mobile communication system. However, it brings in some new challenges to mobility management due to the frequent handovers and heavy signaling overhead. The problem becomes severe for vehicles owing to their fast moving speed, making it more sensitive to the handover delay with reactive handover decision. In this paper, driven by a real-world vehicle mobility dataset, we propose a proactive mobility management solution based on the virtual cell technique for vehicles. Assisted by a trajectory prediction framework based on the long short-term memory neural network, four function modules are designed in the centralized Software-Defined Networking controller to support the proactive solution. The corresponding signaling procedure is then carefully designed, working with virtual cells to reduce the signaling cost. The prediction framework can achieve satisfactory performance of predicting the next location. The proposed proactive solution eliminates the handover delay and reduces the handover signaling cost by 35% compared with the reactive approach.

Index Terms—Proactive mobility management, LSTM, trajectory prediction, signaling procedure, virtual cells, UDN

I. INTRODUCTION

Ultra-dense networking (UDN) is considered as a pillar technology for the next-generation mobile communication system to enhance the system capacity in hotspots [1]. However, mobility management becomes more complex because of the ultra-dense and irregular deployment of the next-generation NodeBs, i.e., gNBs. Handover happens frequently, and a coordinated multi-cell transmission scheme has been proposed to solve this problem [2]. Users can receive data from multiple gNBs as if there is a virtual cell around them [3]. The complex signaling interactions introduced by virtual cells deteriorate mobility management. The Software-Defined Networking (SDN) technology provides an innovation enabler [4]. The logically centralized SDN controller maintains a global view of the network and makes it more flexible [5]. Mobility management based on SDN-enabled virtual cells becomes a hot research topic in UDN.

In the existing works, mobility enhancements based on virtual cells were proposed, which are assisted by anchor base stations in the heterogeneous network [6], [7]. Some researches look at mobility management based on an architecture integrating the virtual cell technique and SDN [8], [9], which mainly focused on pedestrians. There are many bicycles, buses, and vehicles loaded with passengers in UDN. Other researches proposed the realizations of virtual cells con-

sidering vehicles-to-infrastructure communications [10], [11]. However, the formation of virtual cells is still reactive, i.e., gNBs form a virtual cell centering at the current location of the vehicle. Besides, they did not provide a complete mobility management solution for virtual cells. In the traditional reactive method, the selected gNBs start to provide services after the measurement, decision and signaling interaction procedures. Unfortunately, the new gNBs may become invalid soon because of the high-speed movement of cars and dense deployment of the gNBs. It will result in more useless handovers, and the increased handover signaling cost is a heavy burden for the network manager.

The mobility pattern of users can be predicted efficiently with the development of machine learning [12]. If we can start the handover preparation in advance and reasonably select the serving gNBs according to the predicted next location, we can reduce the frequent handover and the signaling cost. Therefore, we propose a proactive mobility management solution based on virtual cells in this paper. First, we propose a prediction framework based on the long short-term memory (LSTM) neural network to predict the trajectory of moving vehicles with a data-driven approach. Next, we design four function modules for the SDN controller to measure the quality of gNBs and build virtual cells in advance based on the predicted next location of vehicles. It eliminates the handover delay by starting the handover procedure before users reach the next location. Then, we carefully design the signaling procedure based on our virtual cell construction. The advantages of the virtual cell architecture and the optimized signaling procedure greatly reduce the signaling overhead in the mobility management process. The proposed solution reduces the handover signaling cost by 35% compared with the reactive approach.

The rest of the paper is organized as follows: the system model in Section II, our proactive mobility management solution in Section III, the analysis in Section IV, numerical results in Section V, and the conclusion in Section VI.

II. SYSTEM MODEL

In this section, we describe the system model and assumptions. As shown in Fig. 1, the considered area with radius R is covered by a centralized SDN controller and it manages all gNBs. The virtual cell of user j is designed as a circular area around the user with radius D , $V_j \triangleq \{gNB_i \in \Phi_{gNB}, |x_i - x_j| \leq D\}$. x_i and x_j denote the

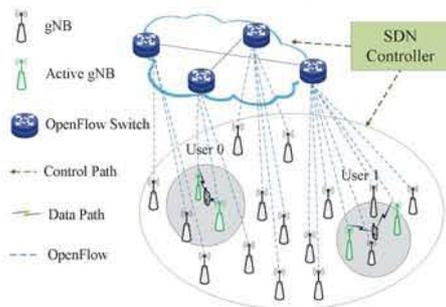


Fig. 1: The network architecture

location of gNB i and user j , respectively. The gNBs with the OpenFlow protocol are distributed with a stationary Poisson Point Process (PPP) Φ_{gNB} of density λ_{gNB} . $h_{ij} = \sqrt{l_{ij}}f_{ij}$ is the channel gain for user j from gNB i with the corresponding path-loss $l_{ij} = |x_i - x_j|^{-\alpha}$, and $\alpha > 2$ is the path-loss exponent. Assuming a Rayleigh fading environment, $\{f_{ij}\}$ are independently and identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance. Each channel is estimated independently.

The non-coherent joint transmission (NCJT) mechanism is used to construct virtual cells in this paper. All active gNBs in a virtual cell create a composite channel by NCJT to provide services for the user. The reference signal received power (RSRP) of active gNBs needs to be larger than a threshold T . The indicator $a_{V_j}^{ia_x} \in \{0, 1\}$ reflects whether gNB i is activated. $a_{V_j}^{ia_x} = 1$ means active, and $a_{V_j}^{ia_x} = 0$ is inactive. ia_x is the resource allocation scheme. ia_1 means the resource allocated for NCJT cannot be reused by other inactive gNBs in a virtual cell. ia_2 is a reuse scheme. The optimal D , T , and ia_x are determined in [13]. We consider the channel estimation error due to imperfect channel state information (CSI) as a new interference source, and it cannot be ignored. Therefore, the aggregate interference mainly includes three categories. The first is the interference caused by inactive gNBs in a virtual cell when the resource allocation scheme is ia_2 . It is noticed that it is zero when the resource allocation scheme is ia_1 . The second is the interference caused by gNBs outside a virtual cell. The third is the residual interference caused by the imperfect CSI.

S_j denotes the transmission signal to user j , $\sigma_{\text{MMSE},ij}^2$ is the minimum mean-square error of the i -th channel estimate for the j -th user, and n_j is the corresponding complex Gaussian noise with zero mean and variance σ_n^2 . The received signal at user j is given by

$$\begin{aligned}
 y_j = & \sum_{i \in \Phi_{\text{gNB}} \cap V_j} \sqrt{(1 - \sigma_{\text{MMSE},ij}^2)} h_{ij} S_j a_{V_j}^{ia_x} \\
 & + \sum_{i \in \Phi_{\text{gNB}} \cap \bar{V}_j} h_{ij} S_j (1 - a_{V_j}^{ia_x}) + \sum_{i \in \Phi_{\text{gNB}} \cap \bar{V}_j} h_{ij} S_j \quad (1) \\
 & + \sum_{i \in \Phi_{\text{gNB}} \cap V_j} \sigma_{\text{MMSE},ij} h_{ij} S_j a_{V_j}^{ia_x} + n_j.
 \end{aligned}$$

For pilot-based channel estimation, $\sigma_{\text{MMSE},ij}^2$ has the similar form with [14]. N_{pilot} is the total number of pilot resource

blocks dedicated to channel state estimation.

$$\sigma_{\text{MMSE},ij}^2 = \frac{1}{1 + E_{H_{ij}} [\text{SINR}_{\text{pilot},ij}] \frac{N_{\text{pilot}}}{\pi D^2 \lambda_{\text{gNB}}}}. \quad (2)$$

The signal-to-interference-plus-noise ratio (SINR) at user j is shown as

$$\text{SINR} = \frac{\sum_{i \in \Phi_{\text{gNB}} \cap V_j} (1 - \sigma_{\text{MMSE},ij}^2) |h_{ij}|^2 a_{V_j}^{ia_x}}{I_{\text{CSI}} + I_c + I_{\bar{c}} + 1/\eta}, \quad (3)$$

where I_{CSI} is the residual interference due to imperfect CSI, which is

$$I_{\text{CSI}} = \sum_{i \in \Phi_{\text{gNB}} \cap V_j} \sigma_{\text{MMSE},ij}^2 |h_{ij}|^2 a_{V_j}^{ia_x}, \quad (4)$$

I_c is the interference caused by inactive gNBs in a virtual cell when $ia_x = ia_2$, which is

$$I_c = \sum_{i \in \Phi_{\text{gNB}} \cap V_j} |h_{ij}|^2 (1 - a_{V_j}^{ia_2}), \quad (5)$$

$I_{\bar{c}}$ is the interference caused by gNBs outside a virtual cell, which is

$$I_{\bar{c}} = \sum_{i \in \Phi_{\text{gNB}} \cap \bar{V}_j} |h_{ij}|^2, \quad (6)$$

and η is the signal-to-noise ratio.

III. PROACTIVE MOBILITY MANAGEMENT

In the proactive mobility management solution, the serving base station of a user is a virtual cell, i.e., a cluster of active gNBs in the virtual cell area. It is quite different from the reactive approach, and we need to redefine the definition of handover. In the proposed solution, a handover is a change of the activation list of gNBs. When a user moves, the coverage of its virtual cell changes but the activation list may not. The benefit of this definition is to avoid useless handover and reduce handover frequency. If we have an accurate prediction of a user's trajectory, the SDN controller can obtain an optimal activation list of gNBs and start the signaling procedure of handover in advance. It will greatly reduce handover delay and improve the efficiency of handover. The proactive mobility management solution consists of three aspects: 1) the framework of trajectory prediction; 2) the design of function modules in the SDN controller; 3) the handover signaling procedure.

A. Framework of trajectory prediction

LSTM network is a kind of recurrent neural network (RNN). It is well-suited for classifying and making predictions based on time series data [15]. In this section, we build a prediction framework based on LSTM to infer the next location of taxis by their partial historical trajectories. The prediction will use the real dataset of taxi rides in Rome, Italy.

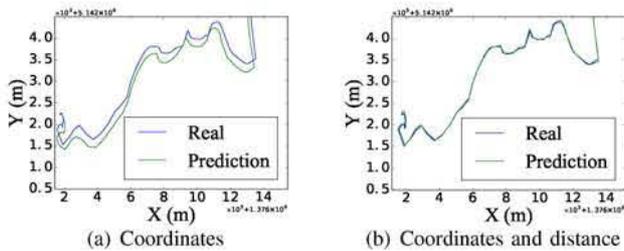


Fig. 2: Predicted trajectories

1) *Dataset*: The Mobility Traces of Rome Taxi Cabs dataset contains the GPS coordinates of approximately 320 taxis collected over 30 days. We randomly choose ten taxis (ID = 2, 17, 39, 55, 68, 87, 139, 196, 222, and 361) and use the following features as the inputs of our framework.

- Taxi ID: It contains a unique identifier for each taxi driver;
- Location: It contains a list of GPS coordinates of each taxi every 15 s.

2) *Prediction framework*: We establish a three-layer LSTM neural network with the first layer of 20 neurons, the second layer of 50 neurons, and a dense layer. A moving window with 19 records is used to predict the location of a taxi in the next time slot. These optimized parameters are selected with a grid search. We create two data sequences (coordinates: longitudes and latitudes) and use them as the input of the framework to predict the next location and distance. The first layer will return the results one by one to the next layer as training data. In the second layer, the first 18 values pass their output values to the layer itself, as the input value for the next operation. However, for the 19th input, the second layer passes its output to the next dense layer. Then the dense layer uses the real value for loss calculation and optimization.

Mean Haversine Distance (MHD) is introduced as the evaluation metric of the prediction performance. The distance d between two points is based on their latitude and longitude, which can be computed as [16]

$$d = 2 \times r \times \arctan \left(\sqrt{\frac{\rho}{1 - \rho}} \right), \quad (7)$$

$$\rho = \sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right), \quad (8)$$

where ϕ is the latitude, λ is the longitude, and r is the Earth's radius, i.e., 6,371 km.

However, the prediction errors with coordinate itself far exceed the average coverage of a gNB (radius=0.043 km). This cannot provide meaningful assistance for the proactive mobility management. Therefore, we add a data sequence of distances between two coordinates to assist the prediction and improve the accuracy. First, we extract the angle information between the current location and the predicted location. Then, we combine the extracted angle information with the predicted distance to generate a new prediction.

In Fig. 2, taking taxi 2 as an example, MHD using the original prediction method is 0.256 km. After considering the distance and angle information, MHD is reduced by 45%. MHD with the proposed method is about 0.1 km on average. From the above results, we can see that our prediction framework has improved the performance by adding the sequence of distances as well as the angle information.

B. Function modules in the SDN controller

In this paper, the SDN controller has four function modules for proactive mobility management. It is in charge of measuring the quality of gNBs, building virtual cells, making activation decision, and executing the handover process. The main function modules of proactive mobility management in the SDN controller include the following aspects.

1) *Measurement control*: In the measurement control module, the controller needs to determine measurement parameters, such as the measurement period, region, and report.

- Measurement period: Let $\tilde{\beta} = 10\%$ denote the maximal outage probability. The probability of $\text{SINR} < \tilde{\beta}$ is approximately near zero in the virtual cell region. So, we choose the mean residence time of a virtual cell as the maximum T_c . According to [17], the crossing rate of a gNB area for a moving user is derived as

$$\gamma_{\text{gNB}} = \frac{2v}{\sqrt{\pi\alpha}}, \quad (9)$$

where α is the circle coverage of a gNB and v is the speed with an average of 11.12 m/s. The mean residence time of the area is: $E[t] = \frac{1}{\gamma_{\text{gNB}}}$. When $D = 122$ m, the minimal $E[t] = 17.22$ s, which is larger than the prediction period 15 s. Thus, we set the maximum T_c to be 15 s. T_c not only affects the signaling cost, but also has a great influence on the transmission capacity. Function Q is designed to choose the optimal T_c , which aims to balance the trade-off between these two metrics.

$$Q(T_c) = \varepsilon \times \left(\frac{S_{T_c=1} - S_{T_c}}{S_{T_c=1}} \right) - \mu \times \left(\frac{C_{T_c=1} - C_{T_c}}{C_{T_c=1}} \right), \quad (10)$$

where ε and μ are the impact factors of signaling cost and transmission capacity, respectively.

- Measurement region: In order to build the next virtual cell ahead of time, we need to measure the potential serving gNBs in advance. The next location of a user at next T_c is predicted by our framework. In our previous work [13], we obtained the optimal radius D of virtual cells.
- Measurement report: The measurement report includes the state information (RSRPs) of all gNBs in the measurement region and the state information (locations and traffic loads) of users.

2) *Activation control*: The SDN controller updates the list of active gNBs periodically. The user compares the new list with the old one and divides its active gNBs into three classes: new gNBs, old gNBs, and ongoing gNBs. The old gNBs will just transfer the buffered data packet and then release

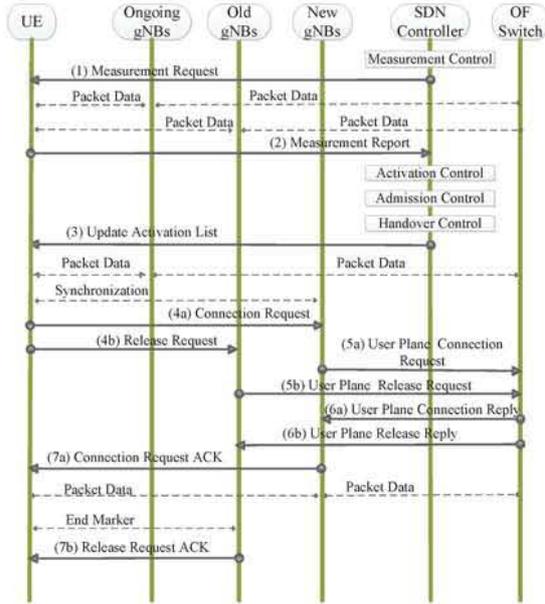


Fig. 3: Proactive handover signaling procedure

all resources. The new ones will synchronize with the user. The ongoing gNBs maintain their transmissions.

3) *Admission control*: The SDN controller determines the list of active gNBs. Then, the controller executes the admission control to avoid overload. If the traffic load of a gNB exceeds a threshold, the controller will remove it from the list. The next available gNB will take its place.

4) *Handover control*: When the SDN controller updates the list of active gNBs, the process of handover is triggered. However, the controller still needs to wait a time-to-trigger (TTT) time to deliver the new list to a user. T_{pro} is the time of a handover procedure by the proactive solution and its value is defined by (13). Every list has its TTT, and

$$\text{TTT} = T_c - T_{\text{pro}}. \quad (11)$$

C. Handover signaling procedure

We design a handover signaling procedure to support our proactive mobility management solution in this section. We show the main steps of the process in Fig. 3.

Step 1: The SDN controller executes the module of measurement control. It determines the measurement period and measurement region. Then, the controller sends a measurement request periodically to a user.

Step 2: The user receives the measurement request and measures the states of gNBs. Then, it creates a measurement report and sends it to the controller.

Step 3: The SDN controller executes the modules of activation control and admission control. Then, it updates the activation list of gNBs and triggers the handover control. After a TTT time, the controller sends the new list to the user.

Step 4: The user receives the activation list and divides the current active gNBs into three classes mentioned above.

Then, it sends connection requests and release requests simultaneously to new gNBs and old gNBs, respectively.

Step 5: The new gNBs which receive the activation request send connection requests to the OpenFlow switch. At the same time, the old gNBs which receive the release requests send release requests to the switch.

Step 6: The OpenFlow switch respectively sends connection replies and release replies to gNBs.

Step 7: The new gNBs build connections with the user and transfer the data. Then, they send connection acknowledgments (ACKs) to the user. Simultaneously, the old gNBs detach data paths. Then, they release ACKs to the user. The ongoing gNBs transfer the buffered data continuously.

Our proposed proactive solution will be compared with the reactive solution [8], in which a reactive handover signaling procedure is provided.

IV. PERFORMANCE METRICS

As a user moves, the activation list of gNBs will change. At this point, the SDN controller needs to perform the corresponding signaling to complete the handover. We define a handover cost S to evaluate the signaling overhead. Besides, we use the handover failure rate (HFR) to reflect the management efficiency of different solutions.

A. Handover cost S

The handover cost S is the transmission and processing time of handover in the whole simulation area. In other words, it is the number of handovers N_{handover} in the simulation area for a typical user multiplied by the time for a handover process.

$$S = E[N_{\text{handover}}] \times \mathcal{T}, \quad (12)$$

where \mathcal{T} is the time of transmission and processing latency in a handover process. From Fig. 3 and [8], we obtain T_{pro} and T_{rea} separately.

$$T_{\text{pro}} = 4T_{\text{OF-Switch}} + 8P_{\text{gNB}} + 3P_{\text{SDN-Controller}}, \quad (13)$$

$$T_{\text{rea}} = 4T_{\text{OF-Switch}} + 13P_{\text{gNB}} + 2P_{\text{SDN-Controller}}, \quad (14)$$

where $T_{\text{OF-Switch}}$ is a transmission latency between gNB and the OpenFlow-enabled switch, P_{gNB} is the processing latency at the gNBs, and $P_{\text{SDN-Controller}}$ is the latency at the SDN controller. N_{handover} is the number of handovers within the simulation area, i.e., the number of times of a mobile user crossing a particular gNB in the simulation area within the time interval between the data packet sessions. It is given by

$$E[N_{\text{handover}}] = \gamma_{\text{gNB}} \times \frac{1}{\lambda_s} \times p, \quad (15)$$

where γ_{gNB} is the border crossing rate of a gNB for a user and is obtained by (9). γ_s is the Poisson session arrival rate. With the density of gNBs λ_{gNB} , we can rewrite (9) as

$$\gamma_{\text{gNB}} = \frac{2v}{\sqrt{\pi}a} = \frac{2v\sqrt{\lambda_{\text{gNB}}}}{\sqrt{\pi}}. \quad (16)$$

p is the handover probability within the simulation area and is given by

$$p = 1 - \frac{1}{\pi R^2 \lambda_{\text{gNB}}}. \quad (17)$$

The handover cost S is rewritten as

$$E[N_{\text{handover}}] = \frac{2v\sqrt{\lambda_{\text{gNB}}}}{\sqrt{\pi}} \times \frac{1}{\lambda_s} \times \left(1 - \frac{1}{\pi R^2 \lambda_{\text{gNB}}}\right). \quad (18)$$

Handover costs of the proactive and reactive solutions are

$$\begin{aligned} S_{\text{pro}} &\leq \left(\frac{1}{T_c}\right) \left(\frac{2v\sqrt{\lambda_{\text{gNB}}}}{\sqrt{\pi}}\right) \left(\frac{1}{\lambda_s}\right) \left(1 - \frac{1}{\pi R^2 \lambda_{\text{gNB}}}\right) \\ &\quad \times (4T_{\text{OF-Switch}} + 8P_{\text{gNB}} + 3P_{\text{SDN-Controller}}). \\ S_{\text{rea}} &\leq \left(\frac{2v\sqrt{\lambda_{\text{gNB}}}}{\sqrt{\pi}}\right) \left(\frac{1}{\lambda_s}\right) \left(1 - \frac{1}{\pi R^2 \lambda_{\text{gNB}}}\right) \\ &\quad \times (4T_{\text{OF-Switch}} + 13P_{\text{gNB}} + 2P_{\text{SDN-Controller}}). \end{aligned} \quad (19)$$

The ongoing activation list at two adjacent T_c may not change. T_{pro} and T_{rea} are the maximal numbers of handovers. Thus, the handover signaling overhead S_{pro} and S_{rea} are also the upper limit of the signaling overhead.

B. Handover failure rate

The handover failure occurs when the user equipment does not have sufficient capacity of the current serving cells. Let C and $C_{\text{req}} = \omega C_{\text{sys}}$ denote the received capacity from the current virtual cell and the required capacity for a successful handover, respectively. $\omega \in (\beta, 1]$ is the proportion of C_{req} in the average system capacity C_{sys} , which can be calculated as

$$\begin{aligned} C_{\text{sys}} &= E[\log_2(1 + \text{SINR})] \\ &= \int_0^\infty (\text{SINR} \geq 2^\tau - 1) d\tau. \end{aligned} \quad (20)$$

Handover failure happens when

$$\frac{C}{C_{\text{req}}} < 1 - \delta. \quad (21)$$

In this work, we do not differentiate when the handover failure occurs. The handover failure rate (HFR) is calculated as

$$\text{HFR} = \frac{\# \text{ of handover failures}}{\# \text{ of handovers}}. \quad (22)$$

V. NUMERICAL RESULTS

In this section, our proposed proactive mobility management solution is compared with the reactive one, which is the baseline. The resource allocation scheme is ia_1 . Other simulation parameters are shown in Table I.

TABLE I: System parameters in the simulation

Parameter	Value	Parameter	Value
R	6 km	λ_{gNB}	174 /km ²
D	0.122 km	λ_s	1 session/s
N_{pilot}	92	C_{sys}	4.6 bps/Hz
α	4	$T_{\text{OF-Switch}}$	1 ms [4]
η	162	P_{gNB}	4 ms [18]
T	0 dBm	$P_{\text{SDN-controller}}$	3 ms [18]
ω	0.5		

A. The optimal measurement period T_c

From the analysis in the previous section, the maximal T_c is 15 s. The largest T_c greatly reduces the signaling cost. It is straightforward for this relationship. When T_c is large, it means that we do not need to make handover decisions very often, and the number of handovers is naturally reduced. The handover signaling cost S is reduced as well.

However, a large T_c does not always bring benefits. As shown in Fig. 4, taking taxi 2 as an example, the system capacities under different speeds of the vehicle are investigated. When $v = 0.01$ m/s, the vehicle is nearly stationary. T_c does not affect the system capacity. With a speed of 10.93 m/s and the increase of T_c , after slightly fluctuating between 4 bps/Hz and 5 bps/Hz, the capacity continuously decreases. The capacity loss between $T_c = 8$ s and 10 s is 50%. Under the high speed, the capacities dramatically dropped by 85% when T_c increases to 5 s. It is obvious that the QoS requirements cannot be guaranteed when we choose a large T_c .

We consider not only the gains of signaling cost but also the loss of capacity. Thus, we need to find an optimal T_c to balance the signaling cost gains and capacity loss. The proactive solution needs to guarantee the performance of taxis, especially when they are under the high mobility. Let $\varepsilon = \mu$. Under the high mobility, i.e., $v = 38.24$ m/s, the optimal T_c is 3 s, which is the default in the sequel as shown in Fig. 5.

B. Handover performance

1) *Signaling cost*: According to (19), handover signaling cost S is affected by both the number of handovers and the handover execution time. When comparing the handover signaling procedure of the proposed proactive and existing reactive solutions, our approach simultaneously executes the process of activating new gNBs and releasing old gNBs (step 4, 5, 6, and 7 in Fig. 3). Thus, T_{pro} can save 17% in handover execution time when compared with $T_{\text{rea}} = 0.098$ s. The proactive solution also helps reduce frequent handovers by accuracy trajectory prediction. Therefore, the proposed solution can efficiently decrease the handover cost as shown in Fig. 6. We classify the trajectories of all taxis into four patterns according to their characteristics, i.e., traveling in straight lines (P1), making turns (P2), making U-turns (P3) and circling around (P4). In the same duration (450 s), the cumulative handover signaling costs of the two solutions decrease from P1 to P4. As the areas taxis passing through shrink, fewer handovers and signaling costs are needed. On the other hand, the gaps between two solutions also decrease from P1 to P4, i.e., 2.28, 2.11, 1.63, and 1.62 s. The trajectory prediction

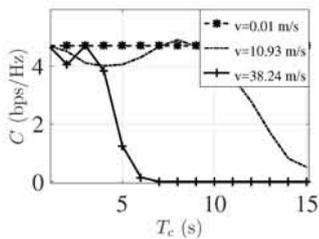


Fig. 4: Capacity C vs T_c

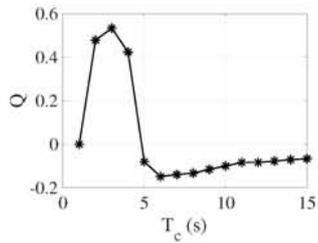


Fig. 5: The optimal T_c

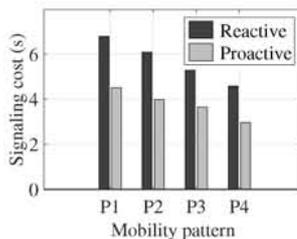


Fig. 6: Signaling cost S

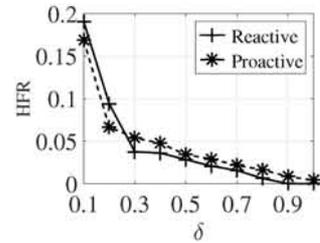


Fig. 7: HFR

accuracy decreases when taxis make more turns, thus resulting in the performance degradation in the proactive solution. Nevertheless, under different mobility patterns, the signaling cost can be efficiently reduced in the proactive approach. This manifests the robustness of the proposed solution.

2) *HFR*: Fig. 7 shows the relationship between HFR and the pre-defined capacity threshold $1 - \delta$. A small δ indicates that the current capacity should be large enough to guarantee a successful handover. Thus, with the increase of δ , the requirement relaxes and thus HFR decreases. The proactive solution takes advantage of trajectory prediction in advance, thus leading to a lower HFR which is reduced by 2%. In addition, prediction deviation in the proposed proactive solution also results in a lower capacity at certain places especially when taxis change their direction sharply. However, the two solutions perform similarly with the difference of only 0.19%. The proposed solution provides an acceptable HFR while bringing gains in the handover delay and signaling overhead.

VI. CONCLUSION

UDN has been a promising direction of network infrastructure densification in the next-generation mobile communication system. However, the ultra-dense deployment of gNBs brings a huge challenge in mobility management, especially for vehicles. In this paper, we proposed a proactive mobility management solution based on the virtual cell technique. The SDN-enabled controller proactively measures the quality of gNBs assisted by the trajectory prediction with a data-driven approach. The corresponding signaling procedure is designed to support the proposed proactive management. Simulation results showed that the proposed solution can greatly decrease the handover delay and handover signaling cost when compared with the reactive approach. In the future work, we will further consider improving the accuracy of the trajectory prediction.

ACKNOWLEDGEMENT

This work is supported by the National Science Technology Major Project of the Ministry of Science and Technology of China (Grant No. 2018ZX03001029), NSERC, CFI, and BCKDF. We are grateful to Syeda Mahfuza Begum for the prediction model analyses.

REFERENCES

- [1] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: challenges, methodologies, and directions," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 78–85, 2016.
- [2] B. Yang, X. Yang, X. Ge, and Q. Li, "Coverage and handover analysis of ultra-dense millimeter-wave networks with control and user plane separation architecture," *IEEE Access*, vol. 6, pp. 54 739–54 750, 2018.
- [3] J. Kim, H. Lee, and S. Chong, "Virtual cell beamforming in cooperative networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1126–1138, June 2014.
- [4] T. D. Assefa, R. Hoque, E. Tragos, and X. Dimitropoulos, "SDN-based local mobility management with X2-interface in femtocell networks," in *Proc. IEEE CAMAD*, 2017, pp. 1–6.
- [5] Z. Zaidi, V. Friderikos, Z. Yousaf, S. Fletcher, M. Dohler, and H. Aghvami, "Will SDN be part of 5G?" *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3220–3258, 2018.
- [6] N. Meng, H. Zhang, and H. Lu, "Virtual cell-based mobility enhancement and performance evaluation in ultra-dense networks," in *Proc. IEEE WCNC*, 2016, pp. 1–6.
- [7] R. Balakrishnan and I. F. Akyildiz, "Local mobility anchoring for seamless handover in coordinated small cells," in *Proc. IEEE GLOBECOM*, 2013, pp. 4489–4494.
- [8] S. Costanzo, R. Shrivastava, D. Xenakis, K. Samdanis, D. Grace, and L. Merakos, "An SDN-based virtual cell framework for enhancing the QoE in TD-LTE pico cells," in *Proc. QoMEX*, 2015, pp. 1–6.
- [9] K. Samdanis, R. Shrivastava, A. Prasad, D. Grace, and X. Costa-Perez, "TD-LTE virtual cells: an SDN architecture for user-centric multi-eNB elastic resource management," *Computer Communications*, vol. 83, pp. 1–15, 2016.
- [10] M. Joud, M. Garca-Lozano, and S. Ruiz, "User specific cell clustering to improve mobility robustness in 5G ultra-dense cellular networks," in *Proc. WONS*, 2018, pp. 45–50.
- [11] T. Sahin, M. Klugel, C. Zhou, and W. Kellerer, "Virtual cells for 5G V2X communications," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 22–28, 2018.
- [12] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [13] Q. Liu, G. Chuai, W. Gao, and K. Zhang, "Fuzzy logic-based virtual cell design in ultra-dense networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 87, 2018.
- [14] A. Lapidoto and S. Shamai, "Fading channels: how perfect need "perfect side information" be?" in *Proc. IEEE ITCW*, 1999, pp. 36–38.
- [15] R. Hug, S. Becker, W. Hübner, and M. Arens, "Particle-based pedestrian path prediction using LSTM-MDL models," in *Proc. ITSC*, 2018, pp. 2684–2691.
- [16] A. De Brébisson, É. Simon, A. Auvolet, P. Vincent, and Y. Bengio, "Artificial neural networks applied to taxi destination prediction," *arXiv preprint arXiv:1508.00021*, 2015.
- [17] F. V. Baumann and I. G. Niemegeers, "An evaluation of location management procedures," in *Proceedings of 1994 3rd IEEE International Conf on Universal Personal Communications*, Sep. 1994, pp. 359–364.
- [18] A. L. Aliyu, P. Bull, and A. Abdallah, "Performance implication and analysis of the OpenFlow SDN protocol," in *Proc. WAINA*, 2017, pp. 391–396.