3D Transition Matrix Solution for a Path Dependency Problem of Markov Chains-Based Prediction in Cellular Networks

Metin Ozturk, Paulo Valente Klaine, and Muhammad Ali Imran
School of Engineering - University of Glasgow
James Watt South Building, Glasgow G12 8QQ
{m.ozturk.1, p.valente-klaine.1}@research.gla.ac.uk, muhammad.imran@glasgow.ac.uk

***Abstract*—Handover (HO) management is one of the critical challenges in current and future mobile communication systems due to new technologies being deployed at a network level, such as small and femtocells. Because of the smaller sizes of cells, users are expected to perform more frequent HOs, which can increase signaling costs and also decrease user's performance, if a HO is performed poorly. In order to address this issue, predictive HO techniques, such as Markov chains (MC), have been introduced in the literature due to their simplicity and generality. This technique, however, experiences a path dependency problem, specially when a user performs a HO to the same cell, also known as a re-visit. In this paper, the path dependency problem of this kind of predictors is tackled by introducing a new 3D transition matrix, which has an additional dimension representing the orders of HOs, instead of a conventional 2D one. Results show that the proposed algorithm outperforms the classical MC based predictors both in terms of accuracy and HO cost when re-visits are considered.**

I. Introduction

According to [1], the number of mobile cellular subscribers by the end of 2015 was at around 7 billion, while by the end of 2000, it was only around 738 million. Furthermore, each user is very prone to adopt new technologies as it can be seen by 3G population, which has increased from 45% to 69% between 2011 and 2015 [1]. The main reason for this significant usage of mobile communications is its mobility support, which allows people to make calls and transfer data while they are commuting.

Mobile cellular networks have had a significant change over the years, i.e. from voice-based (1G) to data applications (4G). In addition to this, advancements related to Next Generation Networks (NGNs), such as 5G, have been continuing. In France, for example, speeds of up to 10 Gbps were achieved by Ericsson and Orange companies in a 5G partnership [2]. The evolution of the mobile cellular networks brings many challenges to network providers and researchers. One of this challenges is a handover (HO) management. HO is the change of base station (BS) or access point (AP) when a user has an ongoing call. In the future, it is expected that HO management will become a more difficult task due to the dense deployment of small cells. This, in turn, makes HO management difficult, since users have more possibilities of HOs. Furthermore, expectations for NGNs, in terms of HO management, mobility and network automation make its management crucial [3].

In the conventional HO procedure, mobile terminals (MTs) gather some parameters, i.e. signal-to-interference-plus-noise ratio (SINR), received signal strength indicator (RSSI), and reference signal received power (RSRP), from both serving and neighboring BSs in order to decide whether a HO is required or not. Not only this process can cause latency issues, due to the time required to perform this procedure (preparation and cancellation phases), but also a resource wastage can occur, in case a HO is not performed. In order to overcome these problems, predictive HO schemes have been proposed in the literature [4]–[8].

These schemes offer savings in both resource and time by predicting the future behavior of MTs in advance, which allows BSs to no longer need a fixed resource reservation for possible HOs. Accordingly, the saving on HO signaling costs, which can also be called as decrease in signaling overhead, is very important for the backhaul optimization of radio access networks (RAN) since a correct prediction reduces the number of messages required for HO between BSs and the core network (CN). Thus, predictive HO schemes play an important role in both the network and user side.

HO prediction can be provided by either machine learning or data mining algorithms. Among others, the most popular algorithms found in literature are MC and neural networks (NN). Despite its popularity, MC still has some problems that needs to be addressed in order to provide a reliable performance under different conditions. One of the main problems, which degrades the performance of MC based predictors, is their path dependency; i.e. when a pre-defined path for a user is altered, such as inserting re-visits to pre-visited locations, the accuracy drops dramatically. Hence, this limitation makes the scenario not complying with real applications. This problem is mostly caused by a transition matrix (TM) that MC based predictors use in order to make its predictions. A conventional TM is two-dimensional (2D) and includes transition probabilities from one state to another, and, whenever a user traverses from one cell to another, the state with the highest probability is assigned. Hence, if two states have very similar probabilities (which can occur when re-visits happen), conventional TMs might get confused and may make incorrect predictions.

In this study, in order to address this limitation of MC based predictors, a three-dimensional (3D) TM is proposed.

The conventional 2D TM represents the transition probabilities from one particular state to another one without considering the orders of HOs, i.e. it combines all the probabilities in the same matrix, making it susceptible to the re-visits problem. On the other hand, the 3D TM introduced in this study is a combination of 2D TMs, and every single HO has its own 2D TM. By adding a new dimension to the conventional TMs, it is shown that the proposed solution can mitigate the effect of re-visits. To the best of the authors' knowledge, this is the first attempt to change a structure of the TM in MC based predictors.

The remainder of this paper is organized as follows: works related to HO prediction for mobile networks are presented in Section II. Section III demonstrates the system model, while Section IV provides the evaluation of the model and discusses its results. Lastly, Section V concludes the paper.

II. Related Work

There are numerous studies performed in the literature related to MC based predictors for wireless communications. In [10], [11], authors employed MC for mobility prediction. In these studies, a classical MC was employed and they demonstrated the usability and suitability of the prediction via the MC for LTE networks with femtocell deployment.

In [9], the authors introduced a discrete-time MC (DTMC) in order to manage handovers for a control/data separation architecture (CDSA) in LTE networks. In the proposed method, the authors tried to predict the next place of a user and to reduce the signaling costs of the network. The LTE X2 HO procedure was assumed in this study and it indicates that a no prediction case is better than an incorrect prediction in terms of the signaling cost. This study is used in this paper for comparison purposes since it is an example of a traditional 2D TM, and from now on, it is called as DTMC.

In [12], authors proposed a Markov renewal process (MRP), which is a semi-Markov process, by considering both the location and sojourn time of a user. As a key point belonging to this study, they were able to predict *N* future locations of a user instead of only the next location. In [13], an *Order-k* Markov was proposed for 3G cellular networks. In contrast to classical Markov predictors which only depends on the current state of a user, *Order-k* Markov predictors consider *k* past states in addition to the current state in order to improve the prediction performance. In [14], authors presented a novel algorithm based on MC. In this algorithm, the authors tried to enhance the prediction performance of the MC for both new users and users with a high randomness.

In general, MC based predictors choose the state with the highest probability from the TM as a next state. All aforementioned studies use conventional 2D matrices when they make a prediction. However, in this study, an improvement on the prediction performance of the MC based predictors is proposed by introducing a novel 3D TM. As a main contribution, the 3D matrix can mitigate the effects of re-visits problem of classical MC based predictors.

III. System Model

A. MC based Prediction Concept

In general, MC, which can be fundamentally defined as a stochastic process, defines a finite number of states and related transition probabilities. When it is applied to cellular networks, each cell that a user traverses, can be considered as a state. Also, because of the Markov property, these predictors are only dependent on the current state of the MT in order to determine its next state, therefore, MC predictors are considered memoryless [11].

For a conventional MC predictor, assume $S = \{s_1,\ s_2,\ ...,\ s_n\}$ is the state space, where *n* represents the total number of states. Hence, the transition probability of the next state can be defined as:

$$\begin{aligned} P(S_{n+1} = s_{n+1} | S_n = s_n, ...,\ S_1 = s_1) \\ = P(S_{n+1} = s_{n+1} | S_n = s_n). \end{aligned} \tag{1}$$

As mentioned earlier in this section, the MC based predictors are based on probability theory because of their stochastic nature. All these probabilities are stored in a matrix, the transition matrix (TM). In other words, the TM includes the transition probabilities between states, i.e. one can find out the transition probability from a particular state to another by investigating the TM. It was observed in [10] that the TM is the only parameter affecting the prediction accuracy. The mathematical definition of the TM can be written as

$$T = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}, \tag{2}$$

where p_{ij} represents the transition probability from state *i* to state *j* and $\sum_{k=1}^{n} p_{i,k} = 1\ \forall i$. From now on, this conventional TM is named as 2D TM due to its two dimensional structure.

Overall, the prediction process of the MC based predictors may be represented as in (3)

$$p_m = p_0 T^m, \tag{3}$$

where p_m and p_0 are the probability vector belonging to the *mth* HO and the initial distribution vector, respectively. So, in order to obtain the prediction for the next state of a user, only the initial state, s_0, and the transition matrix are needed.

By using (3), the predictor calculates the transition probabilities between the current state and all other possible states and then assigns the next place as the state with the highest probability. An example of MC based transition between four states is shown in Fig. 1. As it can be seen, there is a transition probability between each state, and the predictor makes its decision according to these values.

B. Problem Definition

In order to start defining the problem belonging to the MC based predictors, re-visits should be described first. A re-visit occurs when the MT performs a transition to the pre-visited state in its daily route.

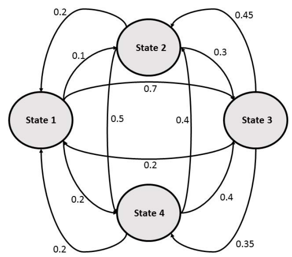

Fig. 1. Sample MC transitions between four states.

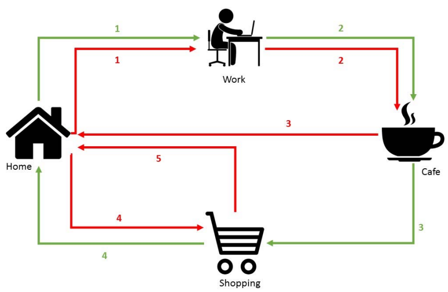

Fig. 2. Sample path with re-visit (red) and without re-visit (green).

In some studies, such as [9], related to the MC based prediction in the literature, these re-visits are not considered. These studies assumed pre-defined paths for its users, since including re-visits would degrade the accuracy of the predictors.

This performance degradation, when re-visits are included, occurs due to the structure of the conventional 2D TM, which is prone to have equal probabilities when re-visits occur. When re-visits are considered, the 2D TM does not know which state to choose from, leading to a wrong prediction. Because the TM is the only parameter affecting the performance of the MC based predictors [10], the structure of the conventional 2D TM should be improved in order to mitigate this problem.

This process is illustrated in Fig. 2 with two cases: green, denoted as *Case 1*, and red, as *Case 2*. As demonstrated in Fig. 2, there is no re-visit in *Case 1* while in *Case 2* re-visits occur (transition 3: *Cafe* to *Home*). In *Case 1*, since the user always moves from *Home* to *Work*, its transition probability is 1, while its 0 from *Home* to other states. Hence, the predictor selects *Work* as a next state when the MT is located at *Home*. Thus, there will be no confusion about the next place decision in this scenario.

On the other hand, in *Case 2*, the transition probabilities from *Home* to *Work* and from *Home* to *Shopping* are equal to each other, which is 50%. In this case, the predictor gets confused. As it can be understood from this case study, re-visits deeply affect the performance of the MC based predictor by building equal probabilities. The main reason for this problem is that the 2D TM combines all the probabilities together without considering anything else.

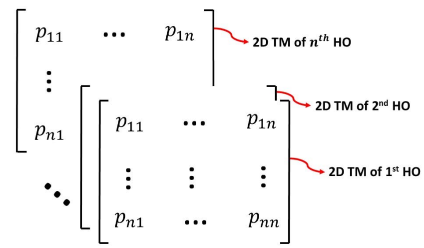

Fig. 3. Proposed 3D transition matrix (3D TM).

C. Proposed Method

In order to alleviate the effects of re-visits, the 3D TM is introduced in this study. In addition to the transition probabilities, the orders of HOs are also considered in the proposed 3D TM. In other words, the 3D TM is composed of an aggregation of multiple 2D TMs, considering the order of the HOs, as shown in Fig. 3.

Assume the previous case study again for the proposed 3D TM model. There was no problem for *Case 1* but the problem occurred for *Case 2* when re-visits took place. Since the 3D TM stores the transition probabilities for each individual HO separately, the transition probability from *Home* to *Work* becomes 1 in the first 2D TM of the 3D TM because this was the first HO performed in the day. Similarly, the probability from *Home* to *Shopping* is also 1 in the third 2D TM of the 3D TM. By doing this, the structure of the 3D TM significantly decreases the possibility of occurrence of equal probabilities.

More generally, the proposed system builds individual TMs for each HO within a day, then combines all TMs in order to construct the 3D TM. When the predictor performs a prediction, it should first investigate the order of HO. Then, it uses the TM which is assigned to that particular HO. The remaining part of the procedure is the same as the classical MC based predictors, i.e. the probability vector is obtained and then the predictor selects the state with the highest probability. The proposed process is shown in Algorithm 1.

Algorithm 1 3D TM for MC based Prediction

```
Initialize the 3D TM with equal probabilities.
Build individual 2D TMs for each HO with an online
    learning process.
Determine the 2D TM within the 3D TM according to the
    order of the HO.
Apply (3) to the determined 2D TM.
Obtain the probability vector.
Make a prediction by selecting the state with the highest
    probability.
```

IV. Performance Evaluation

The simulation models a cellular environment with 19 cells. In addition, the proposed model considers the movement of a single user through the system over a period of 100 days and a total of 10 HOs per day. For each day, the HOs that the user perform can be either from a fixed path, predefined before, or to a neighbor random cell, similar to the model considered in [9]. Furthermore, 4 different scenarios were considered, one without randomness and three others with 10%, 20% and 30% of random HOs, respectively. Simulations were performed in a computer with an Intel Core i7-6700@3.40 GHz processor and a RAM size of 16 GB.

Previously, each TM is constructed with equal probabilities, i.e. $1/(number\ of\ neighbouring\ cells)$, then each HO increases the corresponding probability and decreases the others as the user traverses into the system. In contrast to the model proposed in [9], in this work a three dimensional TM was considered. This enables the algorithm to identify not only the cells that the user is moving from and to, but also to keep track of which HO it is, for example: first of the day, second, third, etc.

By considering this new information (number of the HO), the TMs of each HO will not be contaminated by the other HOs, which makes the algorithm more robust, as it will be seen in the results section. The prediction accuracy used in this study is defined as a ratio between number of correct and total amount of predictions, respectively. Fig. 4(a) shows the results for prediction accuracy for DTMC (2D TM) and 3D TM with different number of re-visits, i.e. from 0 to 7. As shown from Fig. 4(a), although the proposed model outperforms the DTMC in case of no re-visit, this improvement is not significant since results are very close to each other. Hence, both the DTMC and 3D TM give acceptable results in terms of the prediction accuracy when there is no re-visit.

On the other hand, the 3D TM model shows a significant improvement on the prediction accuracy when re-visits are inserted in the system. However, the prediction performance of DTMC is degraded dramatically, and it shows many fluctuations when the number of re-visits changes. These results indicate that the proposed predictor outperforms the DTMC. As such, using the 3D TM instead of the conventional 2D TM mitigates the path dependency problem of the MC based predictors by making them more robust to path alterations. Specifically, while the condition of the conventional MC based predictors, i.e., not including re-visits, does not comply with real world scenarios, the proposed study suits well to real applications.

Intuitively, an increase in the number of HOs results in a better training phase for a predictor, hence it leads to more accurate predictions. On the other hand, too much data may cause an overfitting problem, in which the predictor will not be able to make correct predictions for new unseen paths.

Signaling cost calculations, which were also performed in the DTMC study, are demonstrated in Fig. 4(b). As it can be seen, the proposed 3D TM model also improves the performance of the DTMC in terms of the signaling cost. Similarly, in case of no re-visits, the performance of both DTMC and 3D TM are very close to each other, and they perform better than the no-prediction case.

In the scenario considering re-visits, however, the 3D TM outperforms the DTMC significantly. Moreover, it can be seen that the DTMC performs worse than no-prediction in this case. Hence, this does not comply with the main idea behind the predictive HO, which is to achieve a better performance than without predicitons. The main reason of this is that the signalling cost is directly proportional to the prediction accuracy and an expected signalling cost is given in [9] as

$$C_{pred} = A_p C_c + (1 - A_p) C_i \tag{4}$$

where A_p is the prediction accuracy, C_c is the signaling cost of making a correct prediction, and C_i is the signaling cost of making an incorrect prediction. From (4), the prediction accuracy and expected signaling cost have a direct proportionality, hence degradation in the prediction accuracy affects the signaling cost in a negative way. This is the reason why the proposed 3D TM model can improve the performance of DTMC in terms of both the prediction accuracy and expected signaling cost.

As a result, it was observed from both Fig. 4(a) and Fig. 4(b) that the conventional 2D TM based MC based predictors have a path dependency problem that affects the system performance significantly in terms of both signaling cost and prediction accuracy, while the proposed 3D TM model reduces this issue.

After obtaining the results, which showcase a significant improvement when switching a TM from 2D to 3D, an analysis of both the computational complexity and memory requirements of the proposed solution was performed. Fig. 5 shows both the elapsed time and memory requirements for both DTMC and 3D TM. Here, the elapsed time is considered as an indicator of computational complexity since more complex networks require more time for computation.

From Fig. 5 it can be seen that 14% more time is required to run DTMC than 3D TM. This is because dividing data into several TMs and choosing only one TM makes the computation easier. Thus, the proposed method does not increase the computational complexity, even decreased by 14%. Furthermore, as understood from this figure, 3D TM almost quadruples DTMC in terms of required memory. This is the main drawback of the proposed solution, which is fundamentally caused by the fact that storing an individual TM for each HO of each user increases the memory size drastically. Therefore, the proposed method is not suitable for networks with a storage scarcity. If the memory takes place on the network side, then the solution of the proposed method does not make sense for congested areas since they have many users and they are very prone to have a storage scarcity. On the other hand, this model is more useful for rural areas due to the fact that they have limited number of users, in which causes a more free memory.

V. Conclusion

In this study, the 3D TM is introduced in order to solve the path dependency problem of classical MC based predictors.

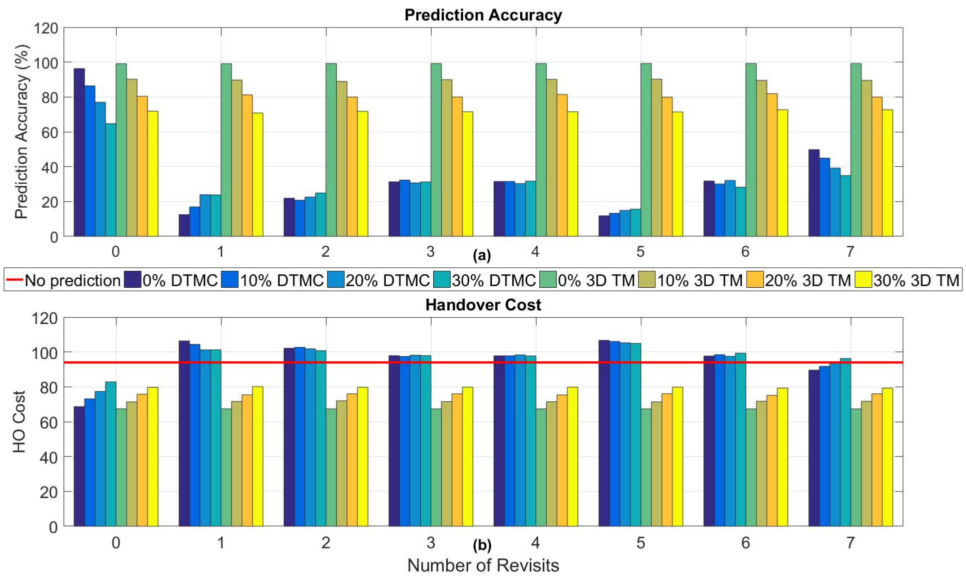

Fig. 4. (a) Accuracies and (b) Signaling costs.

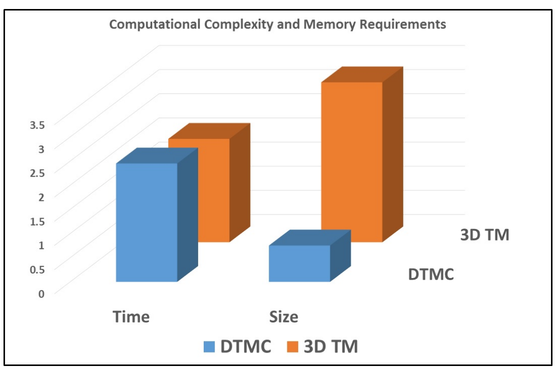

Fig. 5. Computational complexities and memory sizes.

A case study was considered evaluating scenarios with and without re-visits. Results show that MC based predictors experience severe degradation in performance as re-visits in an user's path are considered. On the other hand, the proposed 3D TM model mitigates this issue and also outperforms classical MC models in terms of both prediction accuracy and HO cost.

VI. Acknowledgement

We acknowledge the support of EPSRC (GCRF) funds under the grant no. EP/P028764/1. The first author is supported by Republic of Turkey Ministry of National Education.

References

[1] B. Sanou, "ICT Facts and Figures. The world in 2015", International Telecommunication Union, Geneva, 2015.

[2] "Ericsson and Orange demonstrate speeds beyond 10Gbps in live 5G field trial", Ericsson.com, 2017. [Online]. Available: https://www.ericsson.com/news/170125-ericsson-and-orange-demonstrate-speeds-beyond-10-gbps_244010065_c. [Accessed: 03-Mar- 2017].

[3] N. Panwar, S. Sharma and A. Singh, "A survey on 5G: The next generation of mobile communication", Physical Communication, vol. 18, pp. 64-84, 2016.

[4] L. Lu, J. Wu and W. Chen, "The study of handoff prediction schemes for resource reservation in mobile multimedia wireless networks", International Journal of Communication Systems, vol. 17, no. 6, pp. 535-552, 2004.

[5] Ming-Hsing Chiu and M. Bassiouni, "Predictive schemes for handoff prioritization in cellular networks based on mobile positioning", IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, pp. 510-522, 2000.

[6] K. Lu and J. Wu, "Handoff prediction by mobility characteristics in wireless broadband networks", in Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks., 2005.

[7] Q. Huang, S. Chan and M. Zukerman, "Improving handoff QoS with or without mobility prediction", Electronics Letters, vol. 43, no. 9, p. 534, 2007.

[8] P. Fazio and S. Marano, "Mobility prediction and resource reservation in cellular networks with distributed Markov chains", in 8th International-Wireless Communications and Mobile Computing Conference (IWCMC), 2012.

[9] A. Mohamed, O. Onireti, S. Hoseinitabatabaei, M. Imran, A. Imran and R. Tafazolli, "Mobility Prediction for Handover Management in Cellular Networks with Control/Data Separation", in 2015 IEEE International Conference on Communications (ICC), 2015.

[10] N. Amirrudin, S. Ariffin, N. Malik and N. Ghazali, "Mobility Prediction via Markov Model in LTE Femtocell", International Journal of Computer Applications, vol. 65, no. 18, 2013.

[11] N. Amirrudin, S. Ariffin, N. Malik and N. Ghazali, "Users mobility history-based mobility prediction in LTE femtocells network", in IEEE International RF and Microwave Conference (RFM), 2013.

[12] H. Abu-Ghazaleh and A. Alfa, "Application of Mobility Prediction in Wireless Networks Using Markov Renewal Theory", IEEE Transactions on Vehicular Technology, vol. 59, no. 2, pp. 788-802, 2010.

[13] K. Zhang and L. Cuthbert, "Traffic pattern prediction in cellular networks", in 11th IEEE Singapore International Conference on Communication Systems, 2008, 2008.

[14] S. Bellahsene and L. Kloul, "A new Markov-based mobility prediction algorithm for mobile networks", in 7th European performance engineering conference on Computer performance engineering, Bertinoro, Italy, 2010.

[15] S. Gambs, M. Killijian and M. Cortez, "Next place prediction using mobility Markov chains", in First Workshop on Measurement, Privacy, and Mobility (ACM), 2012.