Introducing a Novel Minimum Accuracy Concept for Predictive Mobility Management Schemes

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*—Over the past few years there has been an emergence in predictive algorithms proposing to manage user mobility issues in cellular networks. Such methods are limited by accuracy bounds causing decreased performance when compared to non-predictive processes, in cases where precision is sub-optimal. In this paper, an analytical model for the minimum required accuracy for predictive methods is derived in terms of both handover (HO) delay and HO signaling cost. After that, the total HO delay and signaling costs are derived for the worst-case scenario (when the predictive process has the same performance as the conventional one), and simulations are conducted using a simulated cellular environment to reveal the importance of the proposed minimum accuracy framework. In addition to this, three different predictors; Markov Chains, Artificial Neural Network (ANN) and an Improved ANN (IANN) are implemented and compared. The results indicate that under certain circumstances, the predictors can occasionally fall below the applicable level. Therefore, the proposed concept of minimum accuracy plays a vital role in determining this corresponding threshold.**

I. Introduction

As the expectations for next generation cellular networks are becoming increasingly demanding, i.e. very low latency (order of 1ms) and higher cell density [1], there is pressure to accommodate state of the art technologies, such as machine learning, in order to make cellular networks more adaptive, dynamic, and resilient [2].

One important issue in cellular networks is the case of mobility management, as network users are constantly moving. Handover (HO), the change of user equipment's (UE) access point (AP) or base station (BS) when in active mode, is very important aspect in the context of mobility management. In this regard, prediction on upcoming HOs is a common method in HO management, since it enables the BS to reserve resources in advance, reducing both HO delay and network signaling cost.

With the help of accurate predictions, HO delay and signalling cost can be decreased significantly [3]. This is important for providing a seamless communication for next generation cellular networks, since delay-intolerant applications; e.g. live video streaming and tactile internet, are in significant demand and will be used intensely in the future. However, the focus point for a predictive HO process should be accuracy, as this can lead to improved prediction capabilities, ultimately improving the fluidity of network resources. On the other hand, the predictive HO process is very prone to escalating both HO delay and signaling cost if the accuracy level of the predictor is low, as each incorrect prediction brings a subsequent penalty.

There have been numerous studies, which propose a HO prediction based method in order to optimize the network, amongst these are: Markov chains [3]–[6] and ANNs [7]–[9] being the most popular due to high efficiency, ease of implementation, and good performance. However, none of these studies provides any information about the minimum accuracy needed for each predictor. In addition, having this information would provide knowledge on how much performance degradation these systems can tolerate in order to continue employing predictive processes. Effects, such as revisit and ping-pong, can make the predictive processes unimplementable by degrading the prediction performances significantly. Therefore, it is apparent that there is a lack of work addressing the design parameters around the predictor.

This study proposes a more standardized framework to the issue of predictive mobility management, by introducing a minimum accuracy requirement for every predictor. First, an analytical model of minimum accuracy requirements for any predictor in terms of HO delay and signaling cost are derived by considering the worst-case scenario, which is given when the performance of the predictor is equal to the conventional case (no prediction). After that, the total HO delay and signaling cost equations are also derived considering the same worst-case scenario.

Furthermore, simulations are conducted to showcase that the performance of predictors are subject to both improvements and degradations. With these simulations, rather than proposing a specific predictor, the intent is to emphasize the usefulness and importance of the proposed minimum accuracy concept. To do this, first, a revisit problem belonging to Markov chains is introduced. Then, negative effects of this problem on the HO delay and signaling cost are also presented. After that, an ANN algorithm is designed to show that ANN is a better alternative when revisits are considered. This example reveals that switching the method of prediction can be a viable solution if the current method is no more applicable due to a performance degradation. Therefore, the proposed concept allows making these kind of decisions easier by evaluating the applicability criteria.

Finally, two different attempts on the designed ANN are offered to increase performance. In the first attempt, one additional input is considered, the orders of the HOs, while in the second attempt, the n-priori locations are added to the ANN. By doing this, it is emphasized that the accuracy improvement is vital for such processes as more accuracy leads

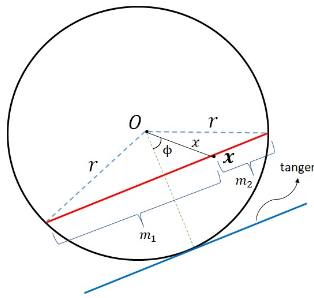

Fig. 1. Circular cell area with RWP.

to improvements in terms of HO delay and signaling cost.

The remainder of this paper is organized as follows: Section II presents the system model and provides detailed information about both the minimum accuracy bounds and the designed ANN. Section III provides the evaluation of the model and discusses its results. Finally, Section IV concludes the paper.

II. System Model

A. Minimum Accuracy in terms of HO Delay

In this paper a single circular cell was considered in order to determine the minimum accuracy bounds. In addition, users would move inside this cell according to the Random Way Point (RWP) mobility model.

Let us manipulate the unit disk example from [10] to make it convenient for a general circular area. Assuming two different points P_1 and P_2 are placed in a circular area with a distance of l as in Fig. 1. First, an arbitrary point $\mathbf{x}$, with a length of $|\mathbf{x}| = x$, is selected in the interior of the considered circular cell. Let m_1 and m_2 be the distances between x and the borders of the circle, such that

$$m_1 = x \sin\phi + \sqrt{r^2 - x^2 \cos^2\phi}, \tag{1}$$

$$m_2 = -x \sin\phi + \sqrt{r^2 - x^2 \cos^2\phi} \tag{2}$$

where r denotes the radius of the circle, and ϕ denotes the angle between $\mathbf{x}$ and the normal of a tangent.

The next step is obtaining the general expression of the probability density function (pdf)[1] in terms of $\mathbf{x}$ [10]:

$$f(x) = \frac{1}{E[l]S^2} \int_0^{\pi} m_1 m_2 (m_1 + m_2) d\phi, \tag{3}$$

where $E[l]$ is the expected value of length, and S is the area of the circle. After performing the required calculations, it can be found that $m_1 m_2 = r^2 - x^2$ and $m_1 + m_2 = 2\sqrt{r^2 - x^2 sin\phi}$. One can evaluate (3) and integrate the result to find the expected value, or an alternative aproach is provided in [11], in which the pdf of the distance between two points in a hypersphere is given in a general manner, and this could be utilized in order to generate the solution for a circular area, as follows:

[1]The conditional probability of finding P_2 within a distance of x from P_1 when the location of P_1 is known.

$$f(\lambda) = \frac{16}{\pi}\lambda\Big[\arccos\lambda - \lambda\sqrt{(1-\lambda^2)}\Big], \tag{4}$$

where $\lambda = \dfrac{l}{2r}$. Hence, (4) becomes

$$f(l) = \frac{4l}{\pi r^2}\left[\arccos\frac{l}{2r} - \frac{l}{2r}\sqrt{1 - \left(\frac{l}{2r}\right)^2}\,\right], \tag{5}$$

as it is also reported in [12]. By integrating (5), the expected value of length can be obtained, such that

$$E[l] = \int_0^{2r} l f(l) dl = \frac{128}{45\pi} r \cong 0.905r. \tag{6}$$

Based on that, (6) represents the average length of one move from two consecutive nodes. However, there might be many movements between many way points in the RWP, hence the number of movements is required to derive the total length. To this end, let $N(S)$ be a Poisson random variable, which represents the number of way points in a cell with an area of S. If the homogeneous Poisson process is assumed, the probability of having n points in the circle is denoted by

$$P\{N(S) = n\} = \frac{(\lambda S)^n}{n!} e^{-\lambda S}, \quad \lambda > 0. \tag{7}$$

The mean of $N(S)$ is also given by

$$E[N(S)] = \lambda S, \tag{8}$$

and the total expected length can be approximated as:

$$E[N(S)]E[l] \cong (0.905)\lambda\pi r^3 \tag{9}$$

where $S = \pi r^2$.

Once (9) is obtained, we can find the model for the expected time $\big(E[T]\big)$ spent in a cell, since there is a direct proportionality between length and time when considering rectilinear motion. Hence, it is now convenient to derive the equation for $E[T_{HO}]$ as

$$E[T_{HO}] = \frac{E[n]E[l]}{V} \cong \frac{(0.905)\lambda\pi r^3}{V} \tag{10}$$

where V is assumed to be a constant velocity.

Note that T_{HO} is the time spent in a cell before handing over to another. From (10) it can be seen that $E[T_{HO}]$ is a function of n, r, and V such that $E[T_{HO}](n, r, V)$. Moreover, it can be observed that $E[T_{HO}]$ has a direct proportionality with n and r, while it has an inverse proportionality with V. This implies that HO management could be more challenging in future cellular networks, specially considering the case of ultra-dense small cell deployments and/or high speed users, such as autonomous vehicles, since $E[T_{HO}]$ might be very short. Once (10) is obtained, the number of HOs can now be derived, since it gives the average time a user spends in a cell. Let T_{mob} be a random variable, which represents the duration that a user is mobile, with a continuous uniform distribution, $U(0, t)$. The pdf of T_{mob} is given by

$$f_t(T_{mob}) = \frac{1}{t}, \quad t \in \mathbb{N}^+, \tag{11}$$

and the expected value of T_{mob} becomes

$$E[T_{mob}] = \frac{t}{2}. \tag{12}$$

Then, the number of HOs that a user experiences during this time is

$$N_{HO} = \left\lfloor \frac{E[T_{mob}]}{E[T_{HO}]} \right\rfloor = \left\lfloor \frac{tV}{(1.81)\lambda\pi r^3} \right\rfloor, \quad N_{HO} \in \mathbb{Z}^+. \tag{13}$$

Note that $floor(\bullet)$ function is utilized, as the number of HOs should be an integer. Based on this, it is now convenient to introduce the concept of HO delay (D_{HO}), which can be defined as the time required for the whole HO process [13]:

$$D_{HO} = \begin{cases} t_p + t_e + t_c + 100ms, & \text{if } cell\ unknown \\ t_p + t_e + t_c + 20ms, & \text{otherwise} \end{cases} \tag{14}$$

where t_p, t_e, and t_c are the time required in without-prediction case for HO preparation, execution, and completion, respectively. If the target cell is unknown, a search delay, set to 80 ms by 3GPP [14], is added to the budget in addition to a 20 ms margin. Moreover, since the HO delay is being investigated from the UE's point of view, t_c should be counted as zero, as its Radio Resource Control (RRC) connection is performed when the HO completion phase starts [13].

One of the alternative ways to reduce D_{HO} is employing a predictive HO process. If, for example, the future locations of a user are known, preparations for upcoming HOs can be done in advance by skipping some of the steps performed in a conventional HO.

The expected HO delay can be expressed as

$$E[D_{HO}] = AD_{HO_c} + (1 - A)D_{HO_i}, \tag{15}$$

where A is the prediction accuracy, D_{HO_c} is the HO delay in case of correct prediction, and D_{HO_i} is the HO delay in case of incorrect prediction. Prediction accuracy (A) can be defined as

$$A = \frac{N_p^c}{N_p^c + N_p^i}, \tag{16}$$

where N_p^c and N_p^i are the numbers of correct and incorrect predictions. It is mentioned in [3] that a predictive HO with a correct prediction is better than the conventional process in terms of number of steps taken during HO process. However, whenever an incorrect prediction occurs, due to resources being allocated to the wrong cell, the conventional process is better than the predictive one. This implies that making an incorrect prediction incurs a penalty in terms of HO delay and signaling cost. In summary, the number of incorrect predictions should be minimized in order to maximize the effectiveness of the system.

According to [3], we can only decrease the delay of preparation phase, thus, (15) can be rearranged for t_p by considering t_e as constant, and setting t_c to zero:

$$E[t_p] = At_{p_c} + (1 - A)t_{p_i} \tag{17}$$

where t_{p_c} and t_{p_i} are the HO preparation delays in case of correct and incorrect predictions, respectively.

Since the main idea behind the predictive HO algorithm is to decrease the HO delay, (17) should be smaller than t_p ($E[t_p] \leq t_p$). From this point, a minimum required prediction accuracy, in terms of HO delay (A_{delay}), can be derived if a worst-case scenario is considered (t_p equals to $E[t_p]$)

$$A_{delay} = \frac{t_p - t_{p_i}}{t_{p_c} - t_{p_i}}. \tag{18}$$

If (13), (15), and (18) are combined as follows, we can derive the equation for overall delay, which can be defined as a multiplication of the single HO delay and the number of HOs during $E[T_{mob}]$, for the worst-case scenario

$$\begin{aligned} total_{DH} &= E[D_{HO}]N_{HO} \\ &= \frac{(t_p - t_{p_i})}{(t_{p_c} - t_{p_i})} \left\lfloor \frac{tV}{(1.81)\lambda\pi r^3} \right\rfloor \left(D_{HO_c} - D_{HO_i} \right) \\ &\quad + \left\lfloor \frac{tV}{(1.81)\lambda\pi r^3} \right\rfloor D_{HO_i}. \end{aligned} \tag{19}$$

B. Minimum Accuracy in terms of HO Signaling Cost

The HO signaling cost is defined as a combination of transmissions costs caused by the messages between BSs, between BS and UE, between BS and Mobility Management Unit (MME); processing costs at the BS, MME, and Service Gateaway (S-GW); and UE's detaching and access costs. Therefore, it is worth noting that the signalling cost mentioned here is from both UE's and the network's point of view [3].

Let us define the original signaling cost of a HO (without prediction) as follows [15]:

$$C = C_{search} + C_{movement}, \tag{20}$$

where C_{search} and $C_{movement}$ are the signaling costs for search and movement, respectively. Since this is a predictive process, it is better to write (20) in terms of the expected signaling cost [3]:

$$E[C] = AC_c + (1 - A)C_i, \tag{21}$$

where C_c and C_i are the signaling costs in case of correct and incorrect predictions, respectively.

Similarly, in order to comply with the idea of predictive HO management, the condition of $E[C] \leq C$ has to be satisfied. For consistency with the previous section, the minimum required prediction accuracy (A_{sig}) can be obtained by considering the worst case, which happens when the actual and expected signaling costs are equal

$$A_{sig} = \frac{C - C_i}{C_c - C_i}. \tag{22}$$

Moreover, the total $E[C]$ during $E[T_{mob}]$ period can be written as

$$\begin{aligned} total_C &= E[C]N_{HO} \\ &= \frac{(C - C_i)}{(C_c - C_i)} \left\lfloor \frac{tV}{(1.81)\lambda\pi r^3} \right\rfloor \Big(C_c - C_i\Big) \\ &\quad + \left\lfloor \frac{tV}{(1.81)\lambda\pi r^3} \right\rfloor C_i. \end{aligned} \tag{23}$$

As seen from (18) and (22), there are two different requirements for prediction accuracies; thus, the overall requirement for the minimum accuracy, by considering both (18) and (22), can be expressed as follows:

$$A_{min} = \max(A_{delay}, A_{sig}) \tag{24}$$

Equation (24) implies that the predictor accuracy is bounded by the maximum between A_{delay} and A_{sig}. Besides, other minimum accuracies can be derived by considering additional metrics.

C. Artificial Neural Network (ANN)

In the proposed study, a three-layer feed-forward conjugate gradient ANN is utilized. In this type of ANN, transitions between layers are performed with an activation function in the forward pass. In this study, a hyperbolic tangent ($tanh$) function is employed, so that the activation functions of each neuron are given by

$$f(z) = \frac{2}{1 + e^{-2z}} - 1. \tag{25}$$

Rather than using a classical backpropagation (BP) process for a backward pass with a constant learning rate, one of the conjugate gradient algorithms, the Scaled Conjugate Gradient (SCG) [16] algorithm, is employed. This method is preferred as the algorithm requires less computation power, since it dynamically calculates the learning rate at each iteration. In addition, the mean squared error (MSE) function is utilized as a cost function of the backwards pass.

Fundamentally, HO prediction is treated as a classification problem in which each different BS constitutes different classes. In other words, the ANN learns the patterns belonging to input layer and tries to predict which BS a user belongs to within a certain confidence level. In this paper, the designed ANN is utilized as a classification tool as well.

The network is composed of 1, 2, 3, and 5 input units, according to the version of the ANN, as it will be discussed later in this section. The inputs represent the serving BS of the user, past n location of the user where $n \in \{1, 2, 4\}$, and the orders of HOs. In addition, 15 hidden layer units and 10 outputs, representing the classes of BSs that a user can be assigned to, are used in the proposed architecture.

In order to validate the number of hidden layer units, the considered artificial data set was split into three, mainly training, cross validation and test set in a 70-15-15 percentage. Intuitively, when the number of units in the hidden layer increases, the error in the training set decreases. In contrast, this increase may cause a generalization problem (overfitting). Although 15 units gives suitable results for our study, this number may vary according to an application type and/or data size due to a consideration of aforementioned trade-off between the training error and generalization.

In addition, some stopping criteria for the proposed ANN were considered. First, when the algorithm meets 10^{-6} performance gradient, it stops computing. If this is not satisfied there is a second stopping criteria whereby, if six increases occur in the validation error value after the last decrease during iterations, the algorithm also stops.

In addition, in order to improve the performance of the proposed ANN, two different approaches were also considered:

- **Appending the orders of HOs:** orders of the HOs within a day are treated as an extra input unit of the ANN. In this regard, each HO is labelled according to its order in a particular day. Consider a mobile user who traverses 6 cells between his home and supermarket, for example, the HOs between cells would be labelled as 1, 2, 3, etc.
- ***n*-priori location:** previous locations of a user can also be considered as extra inputs, acting like a memory for the system. In this attempt, n-priori locations, with $n \in \{1, 2, 4\}$ were considered.

III. Performance Evaluation

In this study, MATLAB is used to create the cellular simulation environment. In addition, the MATLAB ANN Toolbox is also employed in order to model the Neural Network (NN).

In this paper, we start with the Markov predictor by showcasing its revisit problem: the Markov predictor becomes confused and starts making wrong predictions when a user visits the pre-visited cells during a movement period. To see why this problem occurs, it is best to analyze the Markov process briefly. In the Markov process, each transition probability from one state to another is stored in a transition matrix (TM). Then, when a user is located in one specific state, the predictor selects the state, which holds the highest transition probability from that specific state, as a next state. However, these predictors are not very good whenever the probabilities are equal or very close to one another, which can be caused by a user revisiting the same cell via another cell.

The simulation models a cellular environment with 10 BSs and the network is monitored for 200 days. The first 140 days are used to train the ANN, the next 30 days are used for cross-validation, and the remaining days are utilized for testing. In each day, the HOs that a single user trigged can be from either a predefined path, or a random neighbouring cell, similar to the model considered in [3]. To do this, neighbouring lists for each BS were created primarily, and the user was allowed to travel to the neighbouring cells only. Furthermore, various randomness degrees representing the rate of random path selection were considered; starting from 0% until 100% with an 10% incremental step. For 0% randomness, for example, the user follows the same predefined route for 200 days, and for 100% randomness degree, the user randomly selects its path independent of the pre-defined path. In addition, based

TABLE I
SIMULATION PARAMETERS

Parameter	Value
λ	0.8
π	3.14
t	10 hours
V	5 km/h *(pedestrian)*
r	1 km
Number of base stations	10
Number of consecutive days	200

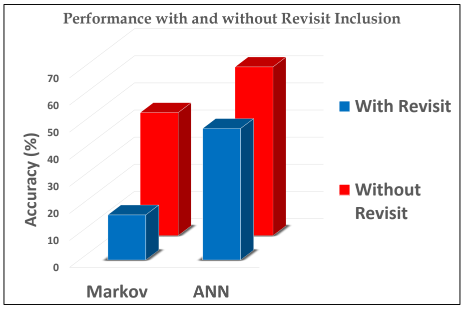

Fig. 2. Performance comparison of Markov chains and ANN.

on the revisit problem, two different data sets were generated, one considering revisits and another without revisits. Note that ping-pong HOs between neighbouring cell are not allowed during data generation process.

With the simulation parameters in Table I, number of HOs per day is found as 10 by calculating (8), (9), (10), and (13).

In Fig. 2 the optimal Markov-based predictor proposed in [3] and the designed ANN results are compared. These values are averages of the different degrees of randomness. As seen from Fig. 2, the ANN's response when revisits are included is not as significant as the response of the Markov-based predictor. There is almost 22% degradation for the ANN while it is around 63% for the Markov-based predictor. This deep performance degradation of Markov-based predictor is due to its highly dependence on the TM, which is prone to have similar probabilities when revisits occur. On the other hand, the ANN does not suffer from the revisit case as much as Markov chains since it has a more complex structure. This result acknowledges that the designed ANN is more appropriate for the dataset including revisits. However, its performance degradation is still an area for improvement.

Fig. 3 shows the performance of ANN achieved after augmenting the input using aforementioned approaches, when revisits are considered. Note that all these figures are obtained by averaging different randomness degrees like Fig. 2. Fig. 3 reveals that augmenting the input with HO orders improves the performance of the ANN by increasing its accuracy by almost 28%. This performance increase can be explained because now the ANN is able to distinguish days from each other with the help of labels. Moreover, these labels are also considered as

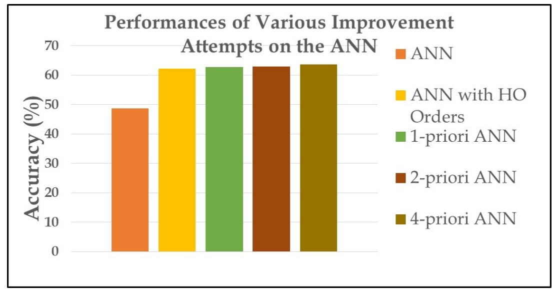

Fig. 3. Improvement on the proposed ANN.

rough representations of time within a day.

Similarly, using n-priori location(s) the performance of the ANN improved significantly; 1-priori ANN raises the accuracy from 48.6% to 62.7%, 2-priori ANN enhances it to 63%, and 4-priori improves to 63.7%. In other words, there is a 29.8% increase on average. However, it was unexpected to see no significant performance improvement for 2-priori and 4-priori ANNs over the 1-priori ANN. One possibility for this observed behaviour might be the simple structure of the dataset, as it only consists of current and next locations. On the other hand, 1-priori ANN seems reasonable since it provides almost same accuracy with 2 and 4-priori ANNs, but lower computational complexity in which each priori adds another unit to the input layer of the ANN.

For the HO delay and signaling cost calculations, the LTE X2 HO process is assumed as in [3]. The HO preparation phase consists of 4 steps in the non-predictive case, 1 step in the correct prediction scenario, and 6 steps when an incorrect prediction is made. Without loss of generality, we assumed each step takes the same amount of time (k). Based on that, A_{delay} can be calculated through (18) as

$$A_{delay} = \frac{t_p - t_{p_i}}{t_{p_c} - t_{p_i}} = \frac{4k - 6k}{1k - 6k} = \frac{2}{5} = 40\%. \quad (26)$$

For the signaling cost calculations, the following value of A_{sig} is obtained if values used in [3] are adopted:

$$A_{sig} = \frac{C - C_i}{C_c - C_i} = \frac{94 - 112}{67 - 112} = 40\%. \quad (27)$$

From (26) and (27), it is obvious that at least 40% accuracy is required by considering both the HO delay and signaling cost in order to make the predictive process better than the conventional one. The reason of why results of (26) and (27) are equal is because the signaling cost is represented in terms of signaling delay in [3]. However, (24) would be employed again if results of (26) and (27) were not equal to each other.

If Fig. 2 is reconsidered with (26) and (27), it is clear that the Markov-based predictor cannot meet the minimum accuracy criteria in the case of revisits (its accuracy is around 16%, far away from the minimum required). Moreover, the Markov-based predictor's accuracy is very close to A_{min} even in the revisit-free case (around 45%). Thus, these results showcase the superiority of the designed ANN over the Markov-based one since it improves the accuracy in both cases.

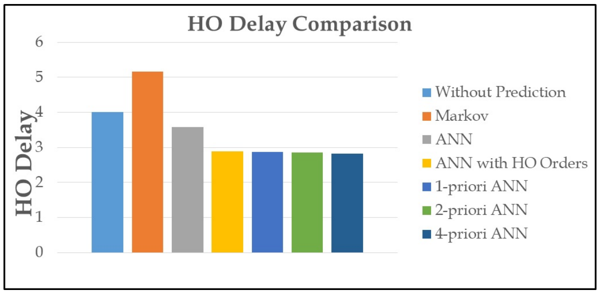

Fig. 4. Handover Delay.

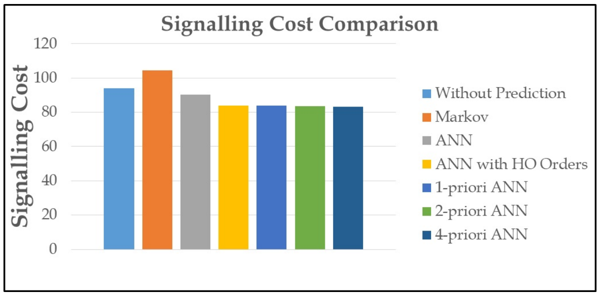

Fig. 5. Signaling cost.

These results highlight that the performance of prediction methods may decay in some circumstances. Although there are ways to improve it, however it is sometimes advantageous to change the prediction method to one that is more appropriate to the environment. Therefore, the proposed minimum accuracy concept provides an obvious benefit in the evaluation of the applicabilities of predictors in the decision process.

It is now convenient to present the results for the HO delay and signaling costs for each predictor in case of revisits in order to show how performance degradations can make predictive processes worse than the conventional one. As seen from Figs. 4 and 5, all versions of the ANN offer better results than the without-prediction case. Moreover, an incremental performance increase is provided by augmenting different features of the ANN. In addition, Markov-based predictor performs worse than the without-prediction case in terms of HO delay and signaling cost, demonstrating that the Markov-based predictor proposed in [3] is not applicable to the predictive HO process in case of revisits in its current form. Intuitively, it is better to use either the ANN with HO orders or 1-priori ANN, since they have similar performance.

IV. Conclusion and Future Work

This work has proposed an improved predictive architecture, contributing to advanced mobility prediction schemes. The required minimum prediction accuracy in terms of both the HO delay and the signaling cost were introduced and fully derived. Moreover, the maxima of the HO delay and the signaling cost were also investigated considering a worst case scenario for the predictive process. Further to this, an ANN with scaled gradient was implemented in order to compare its performance with Markov chains in the case of revisits. Finally, two different ways to increase the performance of the ANN through augmenting the input with HO orders and priori locations were presented.

Results show that if the predictor accuracy is not sufficient due to any effect, such as revisits, or ping-pongs, for example, it is better to improve the performance of the current predictor, if possible, or to switch the prediction method. Hence, this study provides a quantitative and standardized metric to evaluate the sufficiency of obtained accuracy levels.

In future work, we intend to propose a more detailed design parameter for predictors by considering more aspects of the network, such as radio resources. In addition, we plan to undertake the studies with a larger and more comprehensive dataset, in order to investigate and evaluate our solution further.

References

[1] N. Alliance, "5G white paper," *Next generation mobile networks, white paper*, 2015.

[2] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Communications Surveys Tutorials*, vol. 19, pp. 2392–2431, Fourthquarter 2017.

[3] A. Mohamed, O. Onireti, S. A. Hoseinitabatabaei, M. Imran, A. Imran, and R. Tafazolli, "Mobility prediction for handover management in cellular networks with control/data separation," in *Communications (ICC), 2015 IEEE International Conference on*, pp. 3939–3944, IEEE, 2015.

[4] H. Abu-Ghazaleh and A. S. 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.

[5] A. Nadembega, A. Hafid, and T. Taleb, "A destination and mobility path prediction scheme for mobile networks," *IEEE transactions on vehicular technology*, vol. 64, no. 6, pp. 2577–2590, 2015.

[6] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Predictive and core-network efficient RRC signalling for active state handover in RANs with control/data separation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1423–1436, 2017.

[7] A. Quintero, "A user pattern learning strategy for managing users' mobility in UMTS networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 6, pp. 552–566, 2005.

[8] N. Sinclair, D. Harle, I. A. Glover, J. Irvine, and R. C. Atkinson, "An advanced SOM algorithm applied to handover management within LTE," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 1883–1894, 2013.

[9] A. Quintero and O. Garcia, "A profile-based strategy for managing user mobility in third-generation mobile systems," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 134–139, 2004.

[10] E. Hyytia, P. Lassila, and J. Virtamo, "Spatial node distribution of the random waypoint mobility model with applications," *IEEE Transactions on mobile computing*, vol. 5, no. 6, pp. 680–694, 2006.

[11] J. M. Hammersley, "The distribution of distance in a hypersphere," *The Annals of Mathematical Statistics*, pp. 447–452, 1950.

[12] E. Hyytia, P. Lassila, and J. Virtamo, "Spatial node distribution of the random waypoint mobility model with applications," *IEEE Transactions on mobile computing*, vol. 5, no. 6, pp. 680–694, 2006.

[13] K. Alexandris, N. Nikaein, R. Knopp, and C. Bonnet, "Analyzing x2 handover in LTE/LTE-A," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2016 14th International Symposium on*, pp. 1–7, IEEE, 2016.

[14] 3GPP, "Evolved Universal Terrestrial Radio Access (E -UTRA); Requirements for support of radio resource management," TS 36.133, 3rd Generation Partnership Project (3GPP), Sept. 2009.

[15] J. S. Ho and I. F. Akyildiz, "Local anchor scheme for reducing signaling costs in personal communications networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 4, no. 5, pp. 709–725, 1996.

[16] M. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Computer science Department, University of Aaurhus, Denmark*, 1990.