A Framework for Mobility Prediction and High Bandwidth Utilization to Support Mobile Multimedia Streaming

Apollinaire Nadembega[1], Abdelhakim Hafid[1], Tarik Taleb[2]
[1]Network Research Lab, University of Montreal, Canada
{nadembea,ahafid}@iro.umontreal.ca

[2]NEC Europe, Heidelberg, Germany
talebtarik@ieee.org

ABSTRACT – With the increasing number of mobile users on one hand, and the scarce resources in wireless networks on the other hand, there is need for efficient bandwidth utilization in order to avoid/reduce calls blocking/dropping. In this paper, we propose a framework that aims at ensuring QoS in terms of bandwidth reservation along a user's path (i.e., from his source location to his destination). For this objective, knowledge of user's path from source to destination and his handoff times is required. Therefore, we exploit user mobility prediction schemes [1, 2] to predict users' mobility and handoff time windows estimation schemes [3] to predict bandwidth availability for efficient call admission; however, these schemes [1-3] allow for processing individual user requests. Thus, to make the proposed framework scalable (with the number of users), we also propose a group-based path prediction model that estimates path to destination for a group of users (not only for a single user). The simulation results show that the proposed framework is effective in reducing the dropping rate of handoff calls and new call blocking rate, and in improving bandwidth utilization.

I. INTRODUCTION

Recent years have seen an explosion in the number and diversity of mobile devices due to recent trends in the development of chip technologies and their miniaturization. These devices support several multimedia applications, including voice and video, that are bandwidth intensive. These applications may experience performance degradation due to the large number of mobile users, using wireless cellular networks, and their intrinsic characteristics of mobility. Thus, to support QoS from source to destination, the dynamics of every mobile user, such as his path to destination and his arrival/departure times in/from each cell along the path, should be known in advance. Having this knowledge a priori is not possible in realistic scenarios; thus, a solution is to predict, as accurately as possible, the mobility of users, and accordingly perform bandwidth allocation. Indeed, the authors have proposed several predictive schemes in the literature [1-3].

In this paper, we intend using our contributions [1-3] for the prediction of user's mobility and bandwidth availability in order to build a framework, called Mobility Prediction and Bandwidth Reservation Framework (MPBRF), that reflects mobility awareness in the bandwidth reservation to a mobile user's calls along his path to destination. However, these schemes [1-3] allow for processing individual user requests. Thus, to make MPBRF scalable (with the number of users), we propose a Group-Based Path Prediction Model (i.e., GB-PPM) that estimates path to destination for a group of users (not only for a single user). GB-PPM takes into account (a) users' preferences in terms of road characteristic, (b) spatial conceptual maps, (c) users' current locations and (d) their estimated destination; the estimated destination is determined by the Destination Prediction Model (DPM) proposed in [1]. More specifically, GB-PPM divides the navigation map into subzones and determines paths between them, named profile-based paths. A profile is the combination of preferences in terms of road characteristics. The road characteristics are (1) highway, (2) multi-lane, (3) one-way, (4) without traffic light and (5) without stop sign. Then, for each profile, GB-PPM selects the sequence of road segments from a source subzone to a destination subzone that best meets the profile.

The MPBRF architecture consists of two blocks, namely the User Equipment (UE) and the Controller (CTL) which is located in the network system (NS). UE is responsible for predicting user's destination (making use of DPM [1]) and his path to destination when the navigation zone is lightly dense (making use of PPM [2]). UE consists of two modules, namely Destination Predictor (DP) and Path Predictor-User (PP-U). CTL is responsible for predicting path for a group of users when the navigation zone is highly dense (making use of PPM and GB-PPM), the handoff times of the group of users (making use of HTEMOD – Handoff Time Estimation Model [3]) and the available bandwidth. It then enforces a call admission control. CTL consists of four modules, namely Path Predictor-Network (PP-N), Handoff Time Estimator (HTE), Available Bandwidth Estimator (ABE), and Call Admission Inspector (CAI). To the best of our knowledge, this is the first framework which takes into account users' navigation zone density to select the block which shall perform users' path prediction. The advantage of this method is to reduce the amount of messages exchanged between UE and NS; which ultimately avoids the network traffic overload due to the addition of mobility management framework. Also, GB-PPM is the first mobile users group path prediction model.

The remainder of this paper is structured as follows. Section II presents some related research work. The proposed framework along with the envisioned mobile network architecture is described in Section III. Section IV evaluates the performance of the proposed framework and showcases its potential in achieving its design objectives. The paper concludes in Section V.

II. RELATED WORK

Many frameworks for mobility prediction and network resources management have been proposed in the recent literature [4-8]. In [4], the authors proposed a framework that integrates user mobility prediction models with resource availability prediction models to keep a constant or less fluctuating streaming rate and to ultimately ensure steady QoE (Quality of Experience). The proposed framework in [4] consists of two entities, namely (1) Rate

Recommender (RR) and (2) User Equipment (UE). However, the framework does not propose users' mobility features estimation scheme and network resource availability estimation scheme. Also, the framework does not take into account users' aggregation in order to reduce the processing cost of users' mobility features estimation. Wang *et al.* [5] explored an open framework perspective whereby a distributed context management architecture and a communication model based on standard protocols are proposed. In their framework, UbiPaPaGo and UbiHandoff are located in the application service provider; so, all the operations are performed by the network system; this option causes an overload of work. Rathnayake *et al.* [6] presented an architecture which enables client side decision making, taking into consideration network and bandwidth availability predictions, user preferences and the application requirements to optimally schedule data transfers while taking into account the uncertainty in predictions. Their proposed architecture consists of an API (Application Interface), a transport service and two main functional units (i.e., a prediction engine and a scheduling engine). Unfortunately, their proposed architecture provides support for only non- or near-real-time applications. Also, they do not take into account the users' mobility. Prasad and Agrawal [7] proposed a architecture for predictive mobility management that does not estimate the handoff times and is limited to the next cell prediction. Chenn-Jung *et al.*[8] proposed an architecture of resource management scheme that is limited to the next cell. Jun, et al. [11] proposed a CAC and bandwidth reservation scheme which is cell-oriented, distributed and supports prioritized handoff. The current available bandwidth is estimated making use of historical available bandwidth data. Indeed, by mapping the average value of historical bandwidth observations, the scheme estimates the available bandwidth in each cell of the network. Thus, a new call is accepted if the estimated available bandwidth is sufficient to accommodate the call, along the path to destination (it is assumed that the path is known in advance); otherwise, it is blocked. The main limitation of this scheme is the fact that using historical network bandwidth observations (cell behavior/state) does not provide an accurate estimation of available bandwidth compared to using individual users' behaviors.

All in all, existing frameworks have one or more of the following limitations: (1) they do not take into account users' aggregation in users' mobility prediction model [4-8]; (2) all the operations are located in the network system [5]; which adds high burden to the network workload; (3) they are limited to only non- or near-real-time applications [6]; (4) they are limited to only next cell prediction [5-8]; (5) they do not identify the equipments (e.g., users' equipments or network system) which perform prediction schemes and maintain the databases [7, 8]. In this paper, we propose a framework that proposes solutions to these limitations.

III. PROPOSED SCHEME

The objective of the proposed MPBRF framework is to satisfy the requirements, in terms of bandwidth, of each mobiles user along his movement path across different cells towards to the destination. For this purpose, the framework predicts (1) the mobile user path to destination; (2) the entry/exit times of the mobile user to/from cells along the path to destination; and (3) the available bandwidth in each cell along the path to destination. It then accepts the user request, if there is sufficient available bandwidth along the path to accommodate the request; otherwise, it rejects the user request. In the following section, we present the envisioned network and MPBRF architectures. We then present the MPBRF process to accept or reject new calls. Finally, we present the group-based path prediction model (i.e., GB-PPM).

A. Architecture

The network architecture and its components are conceptually depicted in Fig. 1. The figure portrays a network consisting of two parts, fixed and mobile operator networks, which are inter-connected via adequate gateways.

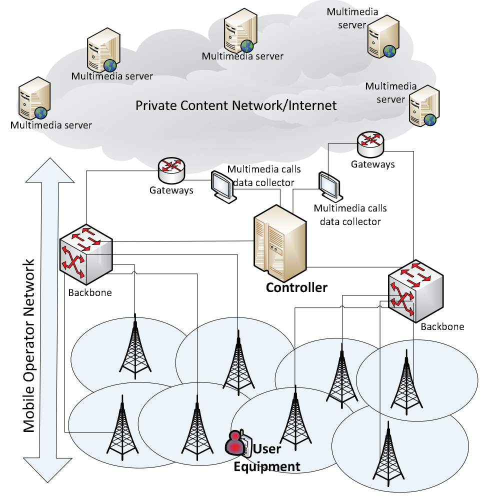

Fig. 1: Envisioned network architecture.

The wired part can be the Internet comprising a number of multimedia servers. The mobile operator administrates a new entity, called Controller, that performs bandwidth management and calls admission control, upon new call request, in order to improve the network performance (i.e., reduce handoff dropping rates while maintaining acceptable new call blocking rates). A number of multimedia calls data collectors are deployed over the entire mobile operator network to collect data about calls and forward them to the Controller which inserts them in its appropriate databases.

As depicted in Fig. 2, MPBRF consists of two blocks: UE and the Controller (CTL). UE is responsible for predicting the mobile user path to destination when the navigation zone is lightly dense while CTL is responsible for predicting the path of the group of users when the navigation zone is highly dense. CTL is also responsible for predicting the entry/exit times of the mobile user to/from cells along the path to destination and available bandwidth in each cell along the path to destination. As briefly described in the introduction section, UE consists of two modules, namely DP and PP-U. As the name infers, DP (resp. PP-U) predicts the user's destination (resp. the user's path to the predicted destination). UE also maintains a number of databases: (1) User Contextual (UC) represents information about user such as his road profile, his tasks, his Interests, his Frequently Visited Locations (FVLs); (2) Navigation Map (NM) contains static data about geographic areas (road topology) including identifiers of nodes (e.g., road junction and FVL) and road segments; we assume that UE has data about road topology and user's road profile; UC may be built by having the user fill in a questionnaire and explicitly express his preferences with regard to different types of roads and his interests with regard to different places within his living area; (3) User Movement Trace (UMT) contains user ID, date, time and node ID (a FVL or a road junction) that represents user location at that specific date and time; (4) User's Frequently Visited Location Trace (UFVLT) contains user ID, date, arrival time, departure time and node ID (a FVL) that represents user location at a specific date from a specific arrival time to a specific departure time; (5) User Movement Characteristics (UMC) contains time, acceleration, speed and road segment ID (a road between two nodes) that represents the user location at a specific time and moving at a specific acceleration and speed; and (6) User Stop Duration (USD) contains time, stop duration and road junction ID (a node ID) that represents the user stop location at a specific time during a specific stop duration. UE will have the responsibility to collect "continuously" data of UMT (whenever current user reaches any type of node), UFVLT (whenever current user reaches a FVL node), USD (whenever current user reaches a road junction node) and UMC (whenever current user transits a road segment).

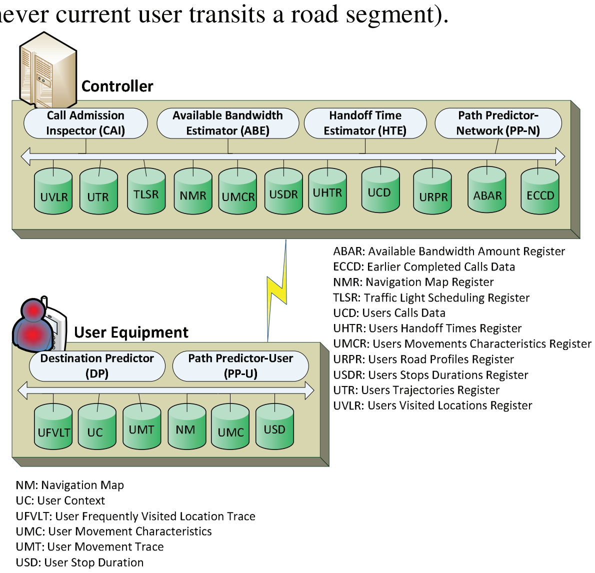

Fig. 2: MPBRF Architecture.

The key component of MPBRF is CTL. As depicted in Fig. 2, the design of CTL consists of four main modules, namely PP-N, HTE, ABE and CAI. PP-N predicts the path of a group of users from their common source to their common destination; HTE predicts the entry/exit times of the mobile users into/from cells along their path to destination; ABE predicts the available bandwidth in each cell along the path to destination; and CAI makes decision on accepting or rejecting new calls. In addition to NMR (resp., UMCR, USDR and UTR) (see Fig. 2) which regroups the stored records in NM (resp. UMC, USD and UMT) of all the users, CTL maintains the following databases: (1) Users Visited Location Register (UVLR) contains users ID and their current locations (i.e., latitude and longitude); each user location is recorded in UVLR one time and updated whenever user changes his location; (2) Traffic Light Scheduling Register (TLSR) contains color ID, action and duration that represent the duration of the action to do when the color is on; (3) Users Handoffs Times Register (UHTR) contains cell ID, t^l and t^u that represent the lower and upper bound values of the estimated time when the user will reach the cell; (4) Users Calls Data (UCD) contains call ID, allocated bandwidth, start time of call t_s, call time t_c and user ID that represents user performing the call; each stored record in UCD is deleted when the call is completed; (5) Available Bandwidth Amount Register (ABAR) contains cell ID, time and amount that represents the predicted amount of bandwidth into the cell at that specific time; (6) Earlier Completed Calls Data (ECCD) contains call ID, date, allocated bandwidth, completed time and expected completion time; each entry/record in ECCD is extracted from UCD before its deletion when the call is completed before its expected time; and (7) Users' Road Profile Register (URPR) contains profile, subzone source, subzone destination and path that represents profile-based path from subzone source to subzone destination associated to profile; data in URPR are provided by the GB-PPM procedure. Static databases (UC, NM, NMR and TLSR) are updated every 6 months; they can be set according to frequency of change of prediction zone and user's context. Each stored record in UMT, UFVLT, URPR, UTR, ECCD and ABAR is deleted after 30 days which we consider a sufficient period to learn users' habits . Each stored record in UMC, UMCR, USD, USDR and UHTR is deleted when the user reaches his trip destination.

B. New call request acceptance/rejection process

The role of MPBRF is to decide on whether to accept or reject a new call request based on predicted available bandwidth in each cell along the path to destination. Fig. 3 illustrates the process of new call acceptance/rejection.

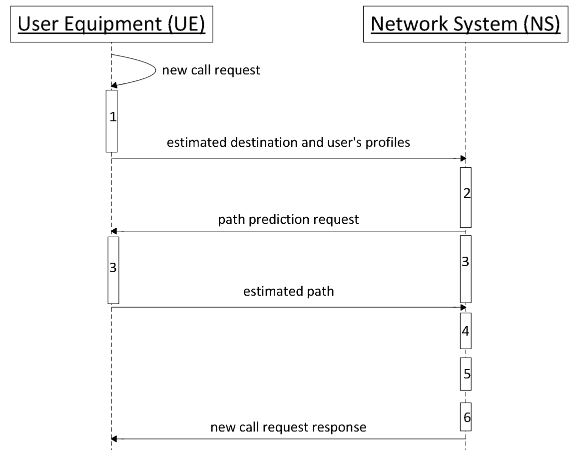

Fig. 3: MPBRF process for new call acceptance/rejection.

The operations are: (1) destination prediction, (2) navigation zone density computation, (3) path prediction according to the navigation zone density, (4) handoff times' estimation, (5) available bandwidth estimation and (6) new call admission control. More specifically, when a user requests a new call, MPBRF determines his destination using DPM [1]; this operation is performed by the DP of the UE using the databases UC, UMT and UFVLT. Then, UE notifies the estimated destination to CTL. In this same message, UE sends the reference of user's new call request (i.e., information which allows identifying the multimedia application/service in order to get the required bandwidth and runtime duration), user's road characteristic profile and his motion data (i.e., data stored into UMC and USD). It shall be recalled that CTL stores data from UMC (resp. USD) into UMCR (resp. USDR). Based on the considered user's current location (stored into UVLR), his estimated destination (forwarded by his UE) and the information about other users' locations (stored into UVLR), CTL computes the density of the navigation zone where the considered user is currently located; the navigation zone represents the rectangle area whose diagonal is the straight line from the user's current location to the estimated destination. For the sake of better understanding, we consider the example shown in Fig. 4; the grey zone represents the navigation zone from A to B.

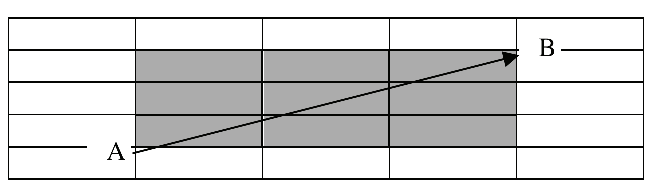

Fig. 4: Navigation zone illustration.

If the navigation zone is lightly dense, CTL notifies UE to perform the prediction/estimation of the path to destination using PPM [2] and to return its result to CTL. Otherwise, CTL predicts the path using GB-PPM, described in details in the next subsection. The key difference between PPM and GB-PPM is that PPM predicts path to destination for each mobile user while GB-PPM predicts path to destination for a group of users and that is for the sake of better scalability. Notice that, it is PP-U of UE (resp. PP-N of CTL) which implements PPM (resp. GP-PPM) using information available at databases UMT and NM (resp. URPR). The predicted path is stored into database UTR of CTL. Knowing the predicted path, CTL performs the prediction/estimation of the entry/exist times to/from cells using HTEMOD [3]; this operation is carried out by HTE of CTL using databases UTR, UMCR, USDR and TLSR; HTE output is stored into UHTR of CTL. Then, using databases UHTR and UCD (collected thanks to the multimedia data collectors), ABE of CTL estimates the available bandwidth in each cell along the path to destination. The results of ABE are stored into ABAR of CTL. Finally, using databases UHTR, ABAR and ECCD (collected thanks to the multimedia data collectors), CAI of CTL checks if there is sufficient available bandwidth along the path to accommodate the user's new call; otherwise, it rejects the user's new call. Fig 5 illustrates the MPBRF activities diagram for new call acceptance/rejection. The remainder of this section presents in details GB-PPM.

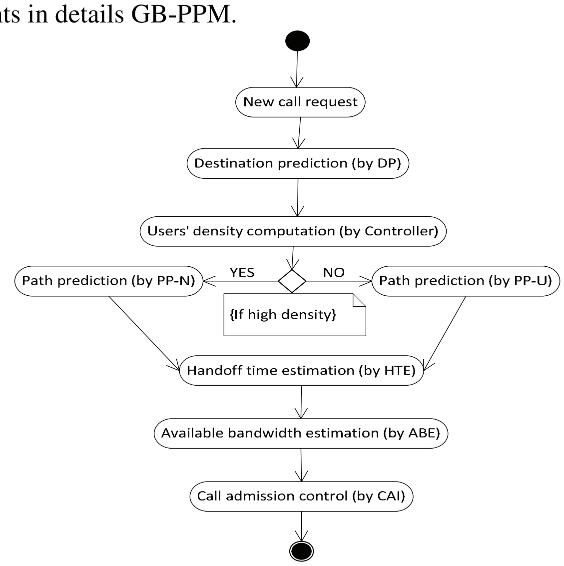

Fig. 5: MPBRF activity diagram.

C. Group-based path prediction model

In this section, we present the GB-PPM scheme that aims to determine the paths, named profile-based path, between two subzones (e.g., from a source subzone to a destination subzone) using spatial conceptual maps and road profiles; whereby a road profile is a combination of preferences in terms of road characteristics; the information about road characteristics, used to generate the road profiles, is extracted from navigation maps. For example, based on the road characteristics highway, multi-lane, one-way, without traffic light and without stop sign, we obtain 32 road profiles, as illustrated in Table I. More specifically, GB-PPM divides the spatial conceptual maps into subzones. Then, for each road profile P*i*, GB-PPM computes the profile-based path from a source subzone to a destination subzone that best meets road profile P*i*. Indeed, the operation of GB-PPM consists of choosing a road segment (among one or more segments) at each road junction towards the destination subzone; a road segment is the road portion between two road junctions or between a road junction and destination; notice that the selected road segment must be located in the current selection area; the current selection area represents the rectangle whose diagonal is the straight line from the current road junction to the destination subzone.

Table I: Road profiles

Road characteristics					Road profiles
high-way	multi-lane	one-way	traffic light	stop sign	
yes	*yes*	*yes*	*yes*	*yes*	P1
yes	*yes*	*yes*	*yes*	*no*	P2
yes	*yes*	*yes*	*no*	*yes*	P3
yes	*yes*	*yes*	*no*	*no*	P4
…	…	…	…	…	…
no	*no*	*no*	*no*	*no*	P32

For the sake of better understanding, and using the example shown in Fig. 4, the selection process starts from the source subzone and is repeated until the destination subzone is reached. At each road junction, GB-PPM starts by a pre-selection process choosing a set Ψ of road segments, among the adjacent road segments to the current road junction, which are located in the current selection area. The GB-PPM selection process is performed based on this pre-selected set. The selection of an adjacent road segment from the pre-selected set Ψ as a next segment is performed using the considered road profile..The adjacent road segment which meets more the considered road profile is selected. Indeed, GB-PPM computes the probability that adjacent road segment *i* of Ψ is the next segment towards destination subzone given road profile P_i. The expression of this probability is defined as follows:

$$p(i,P_j) = \frac{Num(i,P_j)}{\sum_{a=1}^{N_a} Num(a,P_j)} \quad (1)$$

where $Num(i,P_j)$ is the number of times the transition from current junction *S* to road junction *i* is performed, in the past, by users, with road profile P_i, and N_a is the cardinality of Ψ ; $Num(i,P_j)$ can be obtained using UMT and UC. In case of equality between some road segments, the road segment whose direction is more oriented towards the destination subzone is selected; GB-PPM uses the deviation function d to measure the deviation rate of an angle, whereby d is defined as follows:

$$d : [0,180] \rightarrow [0,1]$$

$$d(\theta) = 1 - \frac{\theta}{180}$$

For the sake of illustration, we consider the example shown in Fig. 6. Let *i* be an element of the set of the selected road segments which are in equality. Let S and θ_{SD}^{i} be the angle formed by *SD* (i.e., vector from current road junction S to destination subzone D) and *Si* (i.e., vector from current road junction S to adjacent road junction i). GB-PPM measures the deviation rate of each road segment *i* by computing $d(\theta_{SD}^{i})$ and selects the road segment with the greatest value as the next road segment. The GB-PPM selection process is repeated, making use of the selected road junction (i.e., the end of the selected road segment) as current road junction (i.e., source of selection process) until the destination subzone is reached. Thus, the list of selected road segments constitutes the profile-based path from source subzone to destination subzone for the considered road profile.

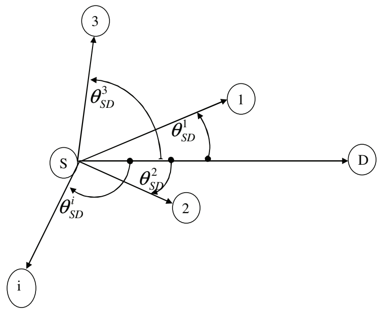

Fig. 6: An example of road segment selection.

IV. Performance Evaluation

To evaluate MPBRF, we used generated mobile user traces; these traces were generated by the Generic Mobility Simulation Framework (GMSF) project [9]. A user trace consists of a user ID, time, acceleration, speed, road segment ID, cell ID, Cartesian coordinates (X and Y) and event that represents the user action (i.e., move, perform handoff, stop or change road segment) at a specific time on a particular road segment. However, we add some programs to take into account the network cells, handoff point (intersection of a road and a cell border) and traffic light management. Also, when the distance to the front user is less than 6 meters, the considered user's speed is decreased; otherwise, when the distance to the front user is less than 3 meters and the current speed of considered user is higher than the speed of the front user, the considered user's speed is limited to the front user's speed. The mobility model used for simulation is Manhattan Model: the road topology is a grid consisting of horizontal and vertical roads. For the sake of simplicity, we assume that (1) the cell coverage is formed by nine blocks (i.e., rectangular area formed by three road segments per side); and (2) only one of the two ends of any road segment has a traffic light. For better understanding, let us consider the example shown in Fig. 7; the grey zone and the black points represent cell coverage and the traffic light locations, respectively.

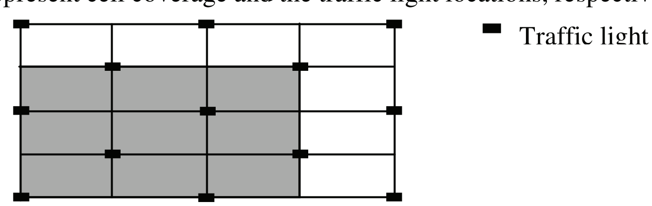

Fig. 7: Illustration of cell coverage and traffic light locations

Table II shows the values of the parameters used in the simulations. Similar to [10, 11], new call requests are generated according to a Poisson distribution with rate λ (calls/second/user) and the minimum bandwidth granularity that may be allocated to any call is 1 *bandwidth unit* (BU). The call duration is assumed to be exponentially distributed with a mean of 300 sec. As comparison terms, we use the scheme described in [11], referred to as AP1. We selected AP1 because, to the best of our knowledge, it represents the most recent work related to CAC and bandwidth reservation in wireless mobile networks that outperforms existing approaches. The new call blocking rate (*Rb*) is computed as follows:

$$Rb = \frac{n_b}{m_b} \quad (2)$$

where n_b and m_b denote the number of new call requests blocked and the total number of new call requests (accepted and blocked). The handoff call dropping rate (*Rd*) is computed as follows:

$$Rd = \frac{n_d}{m_d} \quad (3)$$

where n_d and m_d denote the number of handoff calls dropped and the number of call requests accepted. The bandwidth utilization rate (*Rbw*) is computed as follows:

$$Rbw = \frac{bw_{alloc}}{bw} \quad (4)$$

where bw_{alloc} and bw denote the average amount of allocated bandwidth per time unit (T_k-T_{k-1}) and the overall cell capacity.

Table II: Simulation parameters.

Parameter	Value
GMSF-Mobility model	Manhattan Model (MN): simulation area size=100km2 (10000 m/dimension), number of blocks in one dimension =25, number of users=1500, maximum speed=14 m/s, simulation time =1200sec (20mn).
Low density	< 22 users/cell
acceleration	uniformly distributed between -0.2 m/s2 and 0.2 m/s2
Path prediction accuracy	0.80
$T_k - T_{k-1}$	5 sec
new call requirement	chosen from the set {1, 2, 3, 4} BUs with equal probability

Fig. 8 shows (a) the average new call blocking rate, (b) the average handoff call dropping rate; and (c) the average bandwidth utilization rate for varying cell capacities. In this experiment, the call arrival rate λ is set to 0.03call/second/user.

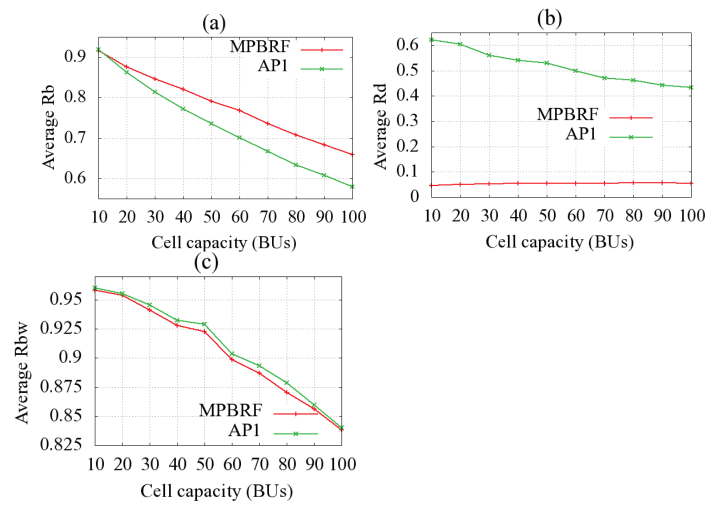

Fig. 8: Performance metrics (Rb, Rd, and Rbw) vs. cell capacity variation.

Fig. 8a shows that AP1 slightly outperforms MPBRF. Indeed, AP1 provides an average of 0.72 per cell while MPBRF provides an average of 0.77 per cell; the average relative improvement (defined as [average *Rb* of MPBRF - average *Rb* of AP1]) of AP1 compared to MPBRF is about 0.05 (5%); this average relative improvement is negligible. We also observe that, for the two schemes, the average new call blocking rate decreases when the cell capacity increases. This is expected since when the cell capacity increases, the number of accepted new call request increases and thus the new call blocking rate decreases. Fig. 8b clearly shows that MPBRF outperforms AP1; MPBRF provides an average of 0.05 per cell while AP1 provides an average of 0.51 per cell; overall, the average relative improvement (defined as [average *Rd* of AP1 - average *Rd* of MPBRF]) of MPBRF compared to AP1 is about 0.46 (46%). Fig. 8c shows that AP1 slightly outperforms MPBRF; AP1 provides an average of 0.91 per cell while MPBRF provides an average of 0.90; the average relative improvement (defined as [average *Rbw* of MPBRF - average *Rbw* of AP1]) of AP1 compared to the proposed framework is about 0.01 (1%); the average relative improvement is negligible too. All in all, Fig. 8 shows that MPBRF allows high bandwidth utilization and considerably reduces the handoff call dropping while maintaining acceptable new call blocking rate, regardless cell capacity.

Fig. 9 shows (a) the average new call blocking rate; (b) the average handoff call dropping rate; and (c) the average bandwidth utilization rate for varying call arrival rates. In this scenario, the cell capacity is 100 BUs.

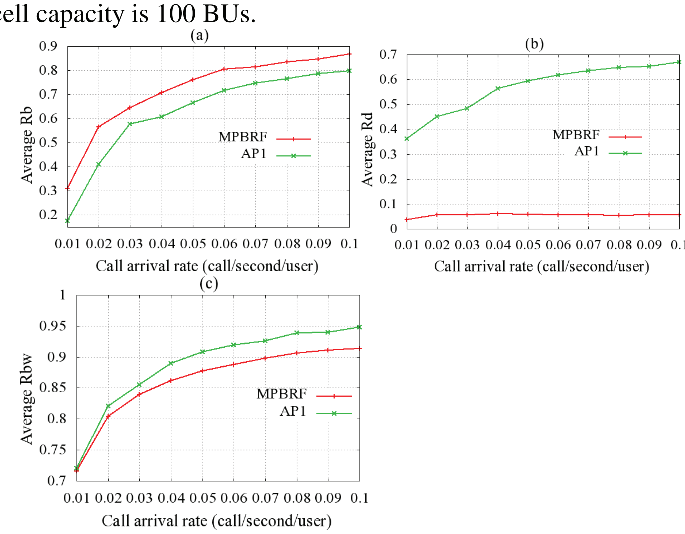

Fig.9: Performance metrics (Rb, Rd and Rbw) vs. call arrival rate variation.

Fig. 9a shows that, for the two schemes, the average new call blocking rate increases with the call arrival rate. This is expected since when the call arrival rate increases, the number of new call requests increases and thus the new call blocking rate increases. We also observe that AP1 slightly outperform MPBRF; AP1 provides an average of 0.62 per call arrival rate while MPBRF provides an average of 0.71 per call arrival rate; the average relative improvement (defined as [average *Rb* of MPBRF - average *Rb* of AP1]) of AP1 compared to MPBRF is about 0.09 (9%). However, Fig. 9b shows that MPBRF handily outperforms AP1. Indeed, MPBRF provides an average of 0.06 while AP1 provides an average of 0.56; the average relative improvement (defined as [average *Rd* of AP1 - average *Rd* of MPBRF]) of MPBRF compared to AP1 is about 0.50 (50%); which is significant. In other words, compared to AP1, MPBRF provides a reduction of 50% in handoff call dropping rate and an increase of 9% in new call blocking rate.

Fig. 9c shows that, for the two schemes, the average bandwidth utilization rate increases with call arrival rate. This is expected since when the call arrival rate increases, the amount of allocated bandwidth increases and thus the bandwidth utilization rate increases. Fig. 9c also shows that AP1 slightly outperforms MPBRF; AP1 provides an average of 0.88 per call arrival rate while MPBRF provides an average of 0.86 per call arrival rate; the average relative improvement (defined as [average *Rbw* of MPBRF - average *Rbw* of AP1]) of AP1 compared to MPBRF is about 0.02 (2%); the average relative improvement is negligible and we can claim that MPBRF allows high bandwidth utilization regardless the call arrival rate. To conclude, Fig. 9 confirms the results of Fig. 8: MPBRF proposes an efficient call admission control scheme regardless call arrival rate.

V. Concluding Remarks

In this paper, we proposed a new framework, called mobility prediction and bandwidth reservation framework (MPBRF), that bases its decision on whether to accept or reject a new call request on predicted available bandwidth in each cell along the path to destination in order to reduce call dropping rate while maintaining an acceptable new call blocking rate and a high bandwidth utilization rate to ultimately ensure acceptable QoS. To make MPBRF scalable (with the number of users), we also proposed a group-based path prediction model, called GB-PPM, that estimates path to destination for a group of users (not only for a single user). We evaluated MPBRF via simulations and compared its performance against that of another approach proposed in [11], referred to as AP1. The simulation results show that MPBRF outperforms (i.e., provides the lowest handoff call dropping rate and a negligible gap in average relative improvement concerning new call blocking rate and bandwidth utilization rate) AP1 for varying cell capacities and call arrival rates. The simulation results confirm the good performance of the framework in reducing call dropping rate, maintaining an acceptable new call blocking rate and a high bandwidth utilization rate.

REFERENCES

[1] A. Nadembega, T. Taleb and A. Hafid, "A Destination Prediction Model based on Historical Data, Contextual Knowledge and Spatial Conceptual Maps," *in Proc. of IEEE ICC'12*, Ottawa, Ontario, CANADA, Jun. 2012.

[2] A. Nadembega, A. Hafid, and T. Taleb, "A Path Prediction Model to Support Mobile Multimedia Streaming," *in Proc. of IEEE ICC'12)*, Ottawa, Ontario, CANADA, Jun. 2012.

[3] A. Nadembega, A. Hafid, and T. Taleb, "Handoff Time Estimation Model for Vehicular Communications," *in Proc. of IEEE ICC'13*, Budapest , Hungary, Jun. 2013.

[4] T. Taleb, A. Hafid, and A. Nadembega, "Mobility-Aware Streaming Rate Recommendation System," *in Proc. of IEEE GLOBECOM'11*, Houston, Texas, USA, Dec. 2011.

[5] C.-Y. Wang, H.-Y. Huang. and R.-H. Hwang , *et al.*, "Mobility management in ubiquitous environments," *in Personal Ubiquitous Comput.,* Vol. 15, No. 3, pp. 235-251, Mar. 2011.

[6] U. Rathnayake, *et al.*, "EMUNE: Architecture for Mobile Data Transfer Scheduling with Network Availability Predictions," *in Mob. Netw. Appl.,* Vol. 17, No. 2, pp. 216-233, Apr. 2012.

[7] P. S. Prasad and P. Agrawal, "A generic framework for mobility prediction and resource utilization in wireless networks," in *Proc. of 2nd Int'l Conf. on Commun. Sys. and Networks (COMSNETS)*, Bangalore, India, Jan. 2010.

[8] H. Chenn-Jung, *et al.*, "A Probabilistic Mobility Prediction Based Resource Management Scheme for WiMAX Femtocells," in *Proc. of 2010 Int'l Conf. on Measuring Technology and Mechatronics Automation (ICMTMA)*, Changsha, China, Mar. 2010.

[9] R. Baumann, F. Legendre and P. Sommer, "Generic mobility simulation framework (GMSF)," *in Proc. of 1st ACM SIGMOBILE Workshop on Mobility models*, Hong Kong, China, May 2008.

[10] K. Madhavi, R. K. Sandhya and R. P. Chandrasekhar, "Optimal Channel Allocation Algorithm with Efficient Bandwidth Reservation for Cellular Networks," in *Int'l J. of Computer Applications,* Vol. 25, No. 5, pp. 40-44, Jul. 2011.

[11] Y. Jun, S. S. Kanhere and M. Hassan, "Improving QoS in High-Speed Mobility Using Bandwidth Maps," *in IEEE Trans. on Mobile Computing,* Vol. 11, No. 4, pp. 603-617, Apr. 2012.