A Distributed Hand-over Management and Pattern Prediction Algorithm for Wireless Networks With Mobile Hosts

Peppino Fazio, Mauro Tropea, Salvatore Marano
D.E.I.S. Department, University of Calabria
87036, Rende, Italy
e-mail: {pfazio, mtropea, marano}@deis.unical.it

Abstract— **In last years, wireless networking is becoming very popular because it is able to satisfy user requests in terms of Quality of Service (QoS); when mobility is present, perhaps, hand-over issues are relevant when hosts change coverage areas during their active sessions. It is very important to mitigate mobility effects, employing an appropriate bandwidth management policy. In our work, we propose two integrated schemes: the first one is based on Markov theory and is aimed at the prediction of mobile hosts movements (in terms of future cells), while the second one is based on statistical theory and is aimed at the minimization of the wasted bandwidth used for passive reservations. So, the proposed Pattern Prediction and Passive Bandwidth Management Algorithm (3P-BMA) is the result of the integration of the Markov predictor and the statistical bandwidth management scheme. 3P-BMA is completely independent on the considered technology, mobility model and vehicular environment. We do not care if the coverage is made by UMTS or WLAN technologies, if hosts are pedestrians or mobile users, etc. Some campaigns of simulation have been led-out in order to confirm the effectiveness of the proposed idea in terms of prediction accuracy, Call Dropping/Blocking probabilities and system utilization.**

Keywords- Hand-over, CDP, CBP, Mobility, Prediction, Pattern, Survey, Markov, Citymob, Passive, Resource, Reservation, Time, Multiplexing.

I. Introduction

In wireless networking, available communication protocols give only one way to ensure QoS and service continuity to mobile users: making a bandwidth reservation over all the cells that a MH will visit during its active connection. For example, the Mobile ReSerVation Protocol (MRSVP) can be used to make passive requests [1], but a prediction scheme is mandatory, in order to know which coverage cells a user will probably visit during its call Call Holding Time (CHT). In this work, we are considering infrastructured networks, where a communication is not possible in a direct way among nodes, unlike ad-hoc architectures (Vehicular, Delay Tolerant and Mobile networks) where, for example, energy consumption issues maybe not trivial [2], [3]. Generally, Mobile Hosts (MH), when dealing with wireless networks, experience some service disruptions for the suffered congestion level, that is variable from a coverage area to another one: each coverage cell manages its connections independently of neighbour conditions. Mobility prediction and early bandwidth reservations are used to guarantee service continuity when mobile hosts are moving among different coverage cells, but passive reservations lead the system to waste bandwidth resources since they are not used until the mobile host enters the considered cell. In this work a new mobility prediction scheme based on Distributed Markov Chains (DMCs) is proposed in order to handle passive reservation and a statistical algorithm, based on the analysis of the Cell Stay Time (CST) distribution, is introduced for reducing bandwidth wastage. The Pattern Prediction and Passive Bandwidth Management Algorithm (PPP-BMA or 3P-BMA) is the result of the integration of the predictor and the bandwidth management scheme. While the cell prediction algorithm is based on the Markov theory and it is aimed at discovering the probably visited future coverage cells, the MRSVP pre-reservation phase is enhanced introducing the time prediction of passive reservations, in order to avoid leaving it unused until the MH enters the considered cell. We employed the Citymob for Roadmaps (C4R) mobility generator [4], in order to appreciate prediction performance when mobility traces are extracted from real roadmaps. 3P-BMA is completely general and does not depend neither on the specific coverage technology, nor on the adopted signaling protocol: the cellular system does not care if MHs are using GSM or WLAN, or if they are in a free-space or urban environment. 3P-BMA is distributed and it has been tested through some deep campaigns of simulations. In this paper, section II gives a good overview of the existing related work, section III proposes the new prediction scheme and section IV shows simulations results. Conclusions are summarized in Section V.

II. State Of The Art And Contributions

Many efforts have been published in last years about prediction schemes and mobility analysis for QoS networks; passive resource management is critical for providing service guarantees in wireless networking. Two passive reservation techniques are proposed in [6], exploiting Wiener prediction and time series theory, making in-advance reservations under non-Poisson and/or non stationary arrival processes, arbitrary distributed call and channel holding time and arbitrary per-call resource demands. In [5] authors optimize system parameters in terms of Call Dropping Probabilities (CDP) and Call Blocking Probabilities (CBP) introducing a prediction algorithm based on data mining approaches, in order to implement a distributed Call Admission Control (CAC) scheme, considering also the throttle flag as indication of the usage of each cell. The authors of [7] propose a new framework to estimate service patterns and to track mobile users, basing the decisions on historical records and predictive patterns of mobile users allowing the estimation of next cells into which a mobile user will possibly move. In [8] the authors give a contribution in WLAN infrastructure planning, basing their decisions on mobility prediction: they propose a new method for feature extraction with a novel neural network classifier based on a hidden genetic algorithm, reaching an acceptable prediction accuracy. In this paper we also demonstrate how the introduction of a certain grade of bandwidth reusability leads to an increasing of system performance, especially in terms of channels utilization. In our previous works, like [9] and [10], a prediction technique based on the CST evaluation of a mobile user is proposed. A formula that relates cell coverage radius and speed is calculated and resource reservation techniques have been proposed, so it is possible to evaluate the number of coverage cells that users will visit during their CHT.

The majority of prediction schemes are aimed at the prediction of a single cell only, without introducing the concept of bandwidth reuse. In this work, instead, a distributed prediction approach is proposed, with an integrated time multiplexing scheme; in particular, the main contributions are:

a) Reservation protocol integration with Markov chains, in order to realize a complete prediction scheme;
b) Distributed model employment (each coverage cell uses a particular Markov chain in order to describe and predict local host movements), in order to improve system utilization;
c) Multiplexing of passive reservations, by which the in-advance reserved amount of bandwidth in a cell can be reused by active MHs, leading to a reduction of the total resource wastage;
d) Model training by taking into account local trajectories (each predictor is specialized for the specific coverage area, with different traffic densities).

III. The Pattern Prediction and Passive Bandwidth Management Algorithm

This section gives a complete description of the proposed idea. Firstly, the way the signaling protocol is used to make CAC and passive reservations is illustrated, then the derivation of the complete model is shown, as well as the integration with DMCs. It must be outlined that the proposed idea does not depend on the employed protocol: for example, it can be one of those described in [1], [11], [20]. We chosen the MRSVP [1], with which one reservation is made by a user on the current coverage cell (active reservation), while passive ones are made on the predicted remote cells.

A) *The distributed markovian prediction scheme*

We considered a cellular system that is composed by a set of coverage cells $C=\{c_l\}$, with $||C||=c$, $1<l<c$ and a coverage radius r_l. Each of them has an associated Distributed Bandwidth Predictor DBP_l, composed by an appropriate semi-Markovian model, as described in the following. As shown in our previous works ([10], [13]), a set S_{ho} of n possible movement directions can be obtained if the coverage area is approximated with a n-edge regular polygon (fig.1). The directions can be indicated with $d_1 ... d_n$, so $S_{ho}=\{d_1, ..., d_n\}$ and $||S_{ho}||=n$. In the classical approaches on cellular networks n is set to 6. The MRSVP session starts with the active service request performed by a MH u on its active cell c_l; if there are no free channels in c_l, the call is refused (with a RESV_NACK message), else the cell c_l applies the results obtained in [1],[9] to evaluate the number of predicted hand-over events NHO. If no hand-over events are predicted (the CST>>CHT), then the call is accepted (u will visit only the current cell c_l), else the DBP_l is used to predict the neighbor cell $nc \in Adj(c_l)$, where $Adj(c_l)$ is the set of neighbors of cell $c_l \in C$ and $||Adj(c_l)||=n$, where n is the number of possible hand-over directions.

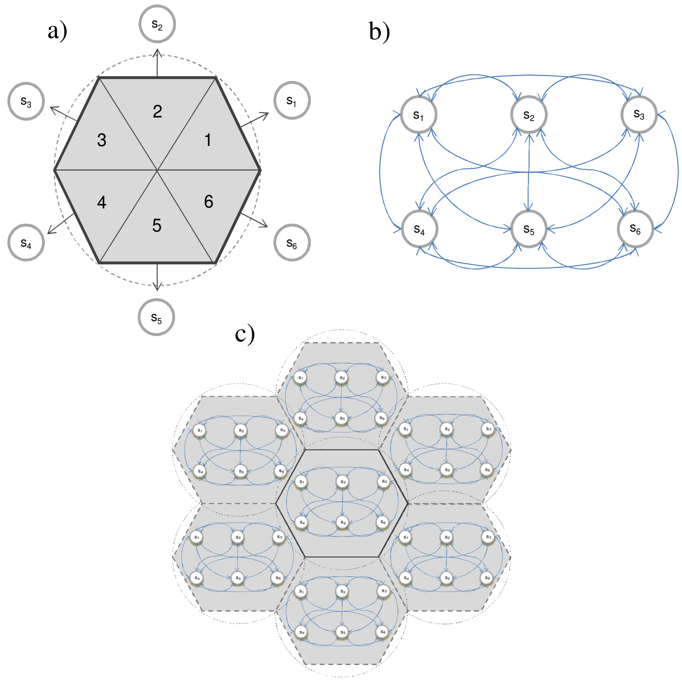

Fig. 1. Coverage cell approximation (a), transition diagram of the related Markov chain (b) and (c) system modeling with (c=7 and n=6).

The DBP_l is represented by a finite set of states and transitions among the states are governed by a set of transition probabilities. The proposed model, illustrated in fig. 1, associates one state of the chain to one hand-over direction. So, once the set of possible states $S=\{s_1,s_2,...,s_n\}$ and a finite state sequence $Q=q_1, q_2, ..., q_m$, with $||S||=n$ and $||Q||=m$ (m is the length of the observation sequence) are defined, three key elements can be defined:

a) A set of state transition probabilities $\boldsymbol{\Lambda}$:

$$\Lambda = \{\lambda_{ij}\}, \lambda_{ij} = p(q_{t+1} = s_j / q_t = s_i) \tag{1}$$

where q_t represents the current state and:

$$\lambda_{ij} \geq 0, \ \ 1 \leq i, j \leq n \qquad \sum_{j=1}^{N} \lambda_{ij} = 1, \ \ 1 \leq i \leq n \tag{2}$$

b) The initial probability array $\boldsymbol{\pi}$:

$$\pi = \{\pi_i\}, \pi_i = P(q_1 = s_i), \quad 1 \le i \le n \tag{3}$$

c) The state sojourn times array $\boldsymbol{ST}$:

$$ST = \{t_i \,/\, P(t_i < \bar{t}) = \int_{-\infty}^{\bar{t}} f_i(t)dt, \quad 1 \le i \le n\} \tag{4}$$

where $f_i(t)$ is the pdf associated to state s_i sojourn time.

Once the parameters in equations (1), (3) and (4) are defined, we can write that the *l-th* DBP can be completely described as follows by a triplet:

$$DBP_l = (\Lambda_l, \mu_l, \pi_l) \tag{5}$$

At this point, the obtained graph can be considered as the state transition map of the HFSMC. Figure 1c illustrates how a distributed set of DBPs $\boldsymbol{D}=\{DBP_l,\ 1\le l\le c\}$ can be used to model the whole cellular system. Supervised training can be approached because observations of MHs movements are possible (in our case by a detailed system simulator, as explained in next sections) and $\mathrm{DBP_l}$ inputs and desired outputs are known. Training observations consist in a set of hand-over direction sequences.

The Maximum Likelihood Estimates (MLE) can be used for evaluating of $\boldsymbol{\Lambda_l}$, $\boldsymbol{\pi_l}$ and $\boldsymbol{ST_l}$ as follows:

$$\lambda_{l_{ij}} = \frac{TR_l(s_i, s_j)}{N_l(s_i)}, \quad \pi_{l_i} = \frac{FIRST_l(q_1 = s_i)}{N_l(q_1)}, \quad pdf_{t_{l_i}} = pdf_{CST_{l_i}} \tag{6}$$

where $1 \le i, j \le n,\ 1 \le l \le c$, $TR_l(s_i,s_j)$ is the number of observed transitions from state i to state j in cell c_l (a transition from s_i to s_j occurs when, in the training data, a MH hands-in a cell from direction d_i and hands-out to direction d_j), $N_l(s_i)$ is the number of transitions from state s_i to any other state in cell t, π_{ti} represents the probability that state s_i (hand-over direction d_i) is the first observed state (q_1) in the training observations for cell t and it is calculated as the ratio between the number of occurrences of s_i being the first observed state $FIRST_t(q_1=s_i)$ and the number of total observations of first states $N_t(q_1)$. For the state sojourn times array $\boldsymbol{ST}$, it must be noticed that, in the proposed model, the time elapsed from the hand-in on direction d_i to the hand-out on direction d_j in a given cell c_l matches with the CST of the same cell, so we can write that the pdfs of CST_{li} and t_{li} are the same. In addition, CST_{li} is independent on the hand-in and hand-out directions, so $CST_{li}=t_{li}=CST_l=t_l\quad \forall c_l \in C$. Details about learning and evaluation can be found in [15], [16]. It is clear that, before the prediction algorithm takes place, each DBP_l belonging to cell $c_l \in C$ needs to be trained, so the terms expressed in eq.(6) can be evaluated by observing MHs movements (in our case we carried out a campaign of simulations, observing MH behaviors from the traces generated by [4]).

B) The statistical bandwidth management

Another important contribution of the proposed work is the passive resource multiplexing: when a MH pre-reserves a certain amount of passive bandwidth in the remote coverage cells, it may be considered as available resource when other incoming traffic makes service request into the system. As discussed in early works ([1], [9]), the pdf of the average time spent by a user in the coverage cell $c_l \in C$ can be approximated by a Gaussian function $pdf_{CSTl}=N_l(\mu_l,\sigma_l)$. Let us indicate with $ho_{in}(h)$ and $ho_{out}(h)$ the predicted hand-in and hand-out times to/from a cell respectively for the *h-th* hand-off event. It can be written that:

$$ho_{in}(h+1) = ho_{out}(h) = ho_{in}(k) + \overline{ho_l}(h), \tag{7}$$

where l is the predicted identifier of the cell that will be probably visited on the *(h+1)-th* hand-off and $\overline{ho_l}$ is a realization of $N_l(\mu_l,\sigma_l)$. Generalizing eq.(7), it can be written that:

$$ho_{in}(h+1) = \sum_{m=0}^{h}\left[ho_{in}(m) + \overline{ho_{lm}}(m)\right], \tag{8}$$

where ho_{lm} is a pdf realization for the predicted cell c_{lm} for the *m-th* hand-off and $\tau_{in}(0)$ is assumed to be the time at which the call has originated. As stated in [17], $ho_{in}(.)$ is still a random variable, because if Γ is a random variable then $pdf(\Gamma)=pdf(a+\Gamma)$ with $a \in R$, then $pdf_{ho(h+1)}=pdf_{ho(h)}=\ldots=pdf_{ho(1)}$, so:

$$pdf\left(\sum_{m=1}^{h} ho_{in}(m)\right) = N(\mu_{TOT}, \sigma_{TOT}), \tag{9}$$

$$where\ \ \mu_{TOT} = \sum_{m=1}^{h}\mu_{cst_m} \quad and \quad \sigma_{TOT} = \frac{\sqrt{\sum_{m=1}^{h}\sigma^2_{cst_m}}}{h}.$$

Let us hypothesize that each coverage cell $c_l \in C$ has a total channel capacity B_l. Without loss of generality, we can write that $B_l=B,\ \forall c_l \in C$. Each passive request has a predicted hand-in time and a predicted hand-out time, as derived in eq. (8), indicated with $ho_{in\text{-}m}(h)$ and $ho_{out\text{-}m}(h)$ respectively, where h indicates that the bandwidth reservation of duration $ho_{out\text{-}m}(h)$ - $ho_{in\text{-}m}(h)$ is made for the *h-th* hand-off of the *m-th* call. If *x-th* and *y-th* requests have no time intersections between their passive reservations in the considered cell, then the same channel can be used, so $req_x \cap req_y = \varnothing$ if:

$$(ho_{out\text{-}x}(j) < ho_{in\text{-}y}(k)) \,||\, (ho_{in\text{-}x}(j) > ho_{out\text{-}y}(k)), \tag{10}$$

where j,k are the hand-off indexes for x and y respectively. In other words, $req_x \cap req_y = \varnothing$ if call_x's reservation ends before *req_y*'s one starts or after *req_y* has finished in the current cell.

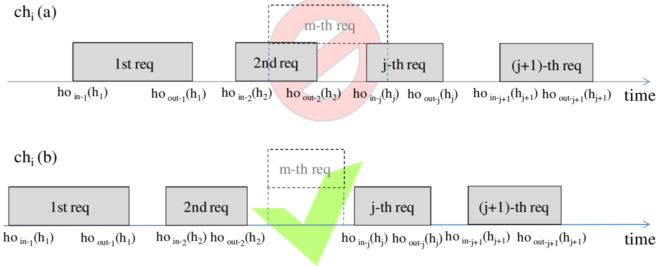

Fig. 3. An example of a refused passive request (a) and an accepted one (b).

Figure 3 illustrates an example of time multiplexing: in the upper case (case a), passive req_m cannot be accommodated in the channel ch_i because $req_m \cap req_2 \neq \varnothing$ and $req_m \cap req_j \neq \varnothing$; in the lower case $req_m \cap req_k \neq \varnothing$ with $k=1,2,j,j+1$, so the passive request can be accepted on channel i. The multiplexing scheme has to optimize the passive bandwidth utilization and the fairness criterion should be respected among the available channels. An appropriate allocation policy must be considered, so the following index, called *bandwidth_fairness_index,* (*bfi*), can be defined:

$$bfi(ch_i) = \frac{1}{T_{ch_i}} \cdot \sum_x \Delta ho_{x\_ch_i}, \tag{11}$$

where $\Delta ho_{x\_chi} = ho_{out\text{-}x}(.)\text{-}ho_{in\text{-}x}(.)$ is a passive reservation of req_x belonging to channel i (ch_i) and $T_{chi}=max_x(ho_{out\text{-}x}(.))\text{-}min_x(ho_{in\text{-}x}(.))$ is the total predicted period of reservation for channel ch_i. It gives an idea of the percentage of time that the considered channel will be occupied. The proposed algorithm tries to obtain $bfi(ch_i) \cong bfi(ch_j)$, introducing the needed fairness condition. When a new passive request req_x arrives to the cell c_l, candidate channels are sorted in increasing order of mux-gap, then the multiplexing algorithm tries to insert req_x into the first "available" channel, in increasing order of *bfi*. If the condition illustrated in eq.(10) is not satisfied for any channel, req_x will be rejected.

IV. PERFORMANCE EVALUATION

Different campaigns of simulations have been led out in order to evaluate 3P-BMA performance in terms of average prediction error, channel assignments, call dropping probability and call blocking probability. The considered scenario consists of a set of c=35 coverage cells, depicted in fig. 4 where, for example, real paths of two down-towns Toulon (France) and Barcelona (Spain) have been considered through the C4R [4] with a 950m x 950m map.

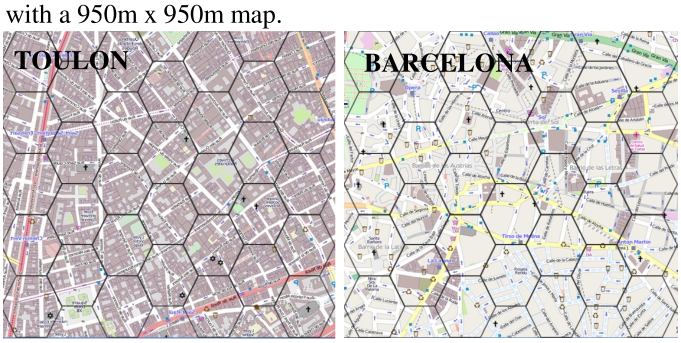

Fig. 4. Simulation maps with c=35, r=150m and B=30.

Curves are shown for different values of training set dimension (from 50 to 300) and r has been set to 150m for space limitations. All the cells have the same coverage radius $r_l=r$, $\forall c_l \in C$ and an exponentially distributed CHT with mean λ=180s has been considered. In the simulation scenario, each coverage cell offers B=30. Border effects on mobility are neglected by ignoring mobile trajectories with paths outside the coverage set. Simulation time has been set to 2500s for each run. First of all, the DBP_l models have to be trained, so a first dedicated campaign of simulations has been carried out in order to obtain the appropriate training data. In particular, the proper size of the training set has been investigated and fig. 5 shows the trend of the prediction accuracy for different number of observations. The parameter depicted in fig. 5 is evaluated as the ratio between the number of correctly predicted hand-overs observations and the number of total observations: accuracy is evaluated on all the observed hand-out directions. From fig. 5 it is evident how good results are reached (from 75% to 95%), because each coverage cell has its own HFSMC and the training is made only on the roads that belong to the specific cell.

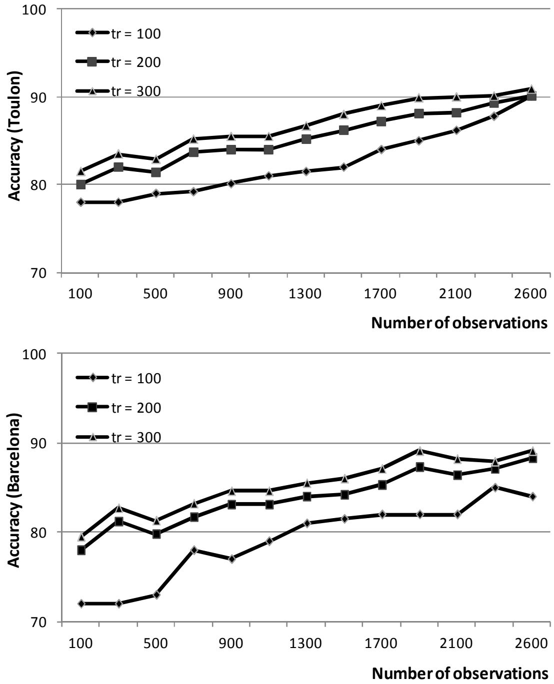

Fig. 5. Prediction accuracy or different training set dimensions (100, 200 and 300) for Toulon (a) and Barcelona scenarios (b).

Training set dimension have to be carefully chosen, in order to avoid over-fitting phenomena [12],[19]. In our case, a training set of 200 items brings the predictor to acceptable performance in terms of prediction accuracy.

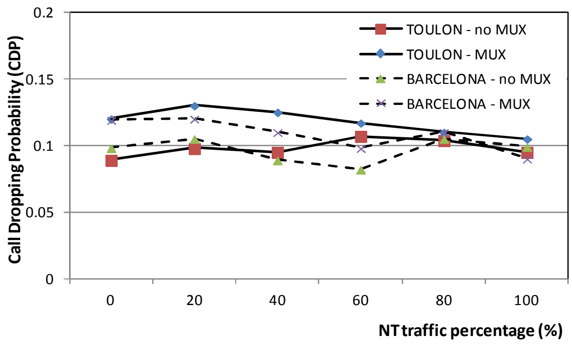

Fig. 6. Call Dropping Probability for Toulon and Barcelona scenarios.

Figure 6 depicts the average trend of the Call Dropping Probability (CDP) for different percentages of Non-Tolerant (NT) traffic (the ratio between the number of users that request QoS and the number of total requests). In addition, NT traffic have been considered with a best-effort complementary traffic (no passive reservations are made for this kind of traffic). As it can be seen, it does not strictly depend on NT traffic percentage and its values belongs to the range [0.076, 0.137]. In both cases (with or without multiplexing) the values are acceptable and no big differences can be observed among the considered scenarios. In fig. 7, the Call Blocking Probability (CBP) is illustrated: in all cases, for higher percentages of MIP traffic the system have to manage more passive reservation requests with the same channels availability, so the call admission control denies the access more frequently. The difference is evident when the multiplexing scheme is used: there is a gain in terms of admitted flows, because the system can accommodate more users, since the bandwidth availability for passive reservations is heavily increased (by channels time multiplexing).

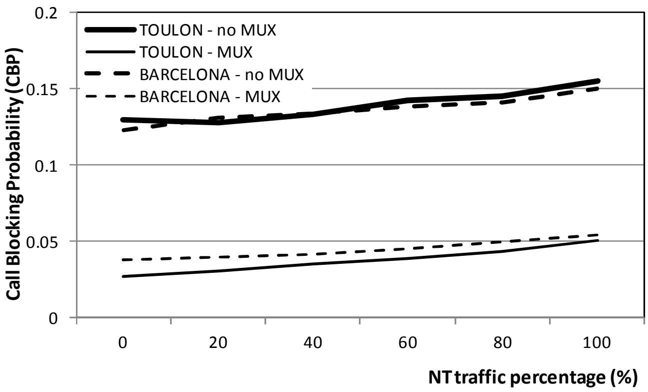

Fig. 7. Call Blocking Probability for Toulon and Barcelona scenarios.

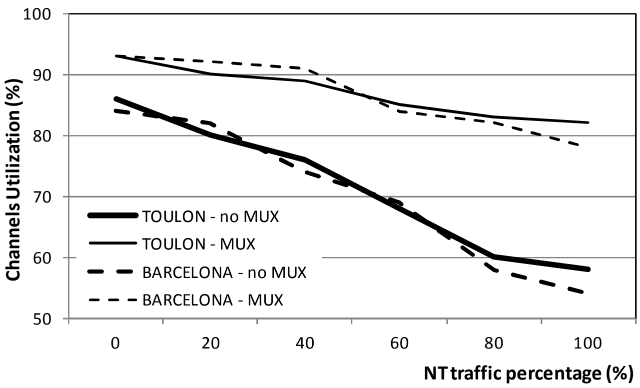

Fig. 8. Average channels utilization percentage for Toulon and Barcelona scenarios.

Figure 8 illustrates the course of the average channels utilization of the whole cellular network: system channels utilization is evaluated as $(B_{la} / B)\ \forall c_l \in C$, where B_{la} is the number of channels used for active connections in cell c_l. All the values are then averaged on the number of cells of the system (c=35 in the considered case). In both no-mux cases, where the multiplexing scheme is not employed), when NT traffic increases, more passive reservations are made into the system, so a higher number of channels are, in-advance, reserved for the arriving MHs. In this way a bandwidth wastage is introduced and channels utilization falls below 60%. When the available channels are dedicated to multiple passive reservations, system utilization increases drastically, obtaining a heavy enhancement of about 24%.

V. CONCLUSIONS

This paper aims at proposing a new Markovian prediction scheme for wireless cellular networks, with the benefits of an integrated policy for passive reservations multiplexing. It is based on Hidden Finite State Markov Chains (HFSMC) processes and guarantees service continuity in QoS networks, without disrupting system utilization performance. The idea is independent on the considered coverage technology (UMTS, WLAN, GSM) and it is of general application. The strength of the proposed idea resides in the integration of MRSVP, Markov predictor and Time Multiplexing, that leads to the MMP, which offers very good performance in terms of accuracy, utilization of the system and service continuity.

REFERENCES

[1] De Rango F., Fazio P., Marano S. , "Utility-based predictive services for adaptive wireless networks with mobile hosts", IEEE Transactions on Vehicular Technology, 2009.

[2] Socievole A., De Rango F., “Evaluation of routing schemes in opportunistic networks considering energy consumption”, Proceedings of the 2012 International Symposium on Performance Evaluation of Computer and Telecommunication Systems, SPECTS 2012.

[3] Socievole A., Marano S., “Evaluating the impact of energy consumption on routing performance in delay tolerant networks”, 8th International Wireless Communications and Mobile Computing Conference, IWCMC 2012, pp. 481-486.

[4] Martinez, F.J., Cano, J.-C., Calafate, C.T., Manzoni, P., “CityMob: A Mobility Model Pattern Generator for VANETs”, ICC Workshops 2008, pp. 370 – 374, Beijing.

[5] T.Zhang et al., “Local Predictive Resource Reservation for Handoff in Multimedia Wireless IP Networks,” IEEE Journal on Selected Area in Communications, vol.19, no.10, Oct.2001,pp.1931-1941.

[6] Chen-Feng Wu, Liang-T. Lee, Hung-Y. Chang, and Der-F. Tao “A Novel Call Admission Control Policy Using Mobility Prediction and Throttle Mechanism for Supporting QoS in Wireless Cellular Networks”, Journal of Control Science and Engineering, Volume 2011, 11 pages.

[7] I.F.Akyildiz, W.Wang, “The Predictive User Mobility Framework for Wireless Multimedia Networks,” IEEE/ACM Transactions on Networking, vol.12, no.6, Dec.2004, pp.1021-1035.

[8] Velmurugan, L. and P. Thangaraj, “A Hidden Genetic Layer Based Neural Network for Mobility Prediction”, American Journal of Applied Sciences 9 (4): 526-530, © 2012 Science Publications.

[9] F. De Rango, P.Fazio, S.Marano, “Cell Stay Time Analysis under Random Way Point Mobility Model in WLAN Networks”, IEEE Communication Letters, Vol.10, Issue 11, pp.763-765, Nov. 2006.

[10] F. De Rango, P.Fazio, S.Marano, “Mobility Prediction and Resource Reservation in WLAN Networks under a 2D Mobility Models,” 63rd Vehicular Technology Conference (VTC Fall, Canada, Sept..25-28, 2006.

[11] I. Jawhar, Jie Wu, “A dynamic range resource reservation protocol for QoS support in wireless networks”, AICCSA '05 Proceedings of the ACS/IEEE 2005 International Conference on Computer Systems and Applications, Pages 65-I, IEEE Computer Society, DC, USA 2005.

[12] Ming Xu, Zhijiao Zhang and Yingwen Chen, “A Resource Reservation Protocol for Mobile Cellular Networks”, Parallel and distributed processing and applications, Lecture Notes in Computer Science, 2005, Volume 3358/2005, 534-543.

[13] P.Fazio, F.De Rango, S.Marano, “2D Movement Direction-Based Reservation Scheme for WLAN Clusters with Passive Advanced Reservations,” in IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2008), Niagara Falls, Canada, May 4-7, 2008.

[14] F. De Rango, P.Fazio, S.Marano, “Cell Stay Time Prediction for Mobility Independent Predictive Services in Wireless Networks,” IEEE Wireless Communications and Networking Conference (WCNC2005), New Orleans, Los Angeles, USA, 13-17 March 2005.

[15] Sherif Akoush and Ahmed Sameh, “Mobile User Movement Prediction Using Bayesian Learning for Neural Networks”, Proceeding of the 2007 International Conference on Wireless Communications and Mobile Computing, pp.191-196, 2007.

[16] Micheal I. Jordan, “Graphical Models”, Statistical Science, vol.19, pp.140-155, 2004.

[17] Montgomery, C., “Applied Statistics and Probability for Engineers”. 3rd ed. Hoboken, NJ: Wiley, 2003.