Composite Model of Packetized VBR Source for Next Generation All-IP Mobile Networks

D. Moltchanov, Y. Koucheryavy, J. Harju
Institute of Communication Engineering,
Tampere University of Technology, Tampere, Finland,
{yk,moltchan,harju}@cs.tut.fi

Abstract— **We propose an integrated traffic model of single packetized variable bit rate traffic source for next generation (NG) All-IP networks. Basically, our model integrates two parts: mobility model and teletraffic model. The mobility part captures mobility behavior of a single user and modelled by Markov chain with finite state space. Teletraffic part captures teletraffic characteristics at both session and packet levels and represented by superposition of a Markov chain and a discrete-time batch Markovian arrival process (D-BMAP) respectively. Mobility and teletraffic parts are integrated, so that the whole model can be seen as triply stochastic process while does not lead outside D-BMAP framework. This is due to the special superposition of parts, which makes the teletraffic model to be probabilistic function of mobility one. Behavior of the proposed model indicates that even in presence of dependence between mobility of the user and its teletraffic demands it is still possible to estimate an actual value of session and packet level traffic offered by a single user to NG All-IP networks.**

Keywords: NG All-IP, 3G, teletraffic, mobility, cross-layering.

I. INTRODUCTION

Up to date the Internet and wireless communications have been regarded as separate technologies because of different types of traffic they were intended for. Nowadays the convergence of these technologies is getting clear. This process is stimulated by widespread grow of both the Internet and mobile communication systems along with users that wish to access Internet services 'anytime and anywhere'.

While next generation mobile systems are not clearly defined, there is a common agreement that these networks will rely on IP protocol as an end-to-end transport technology. The motivation is to introduce a common service platform and transport facilities for future composite 'mobile Internet' known as next generation (NG) All-IP networks.

In addition to broadband wireless access to the Internet, NG All-IP mobile systems must be able to provide quality of service (QoS) to their applications. This is an inherent problem for many service types even in fixed IP networks. Wireless and mobility issues add their own problems on top of this IP flaw. Inherent characteristics of mobile systems, like users' mobility, their teletraffic demands and unstable nature of the air interface have to be addressed before a required quality of user services will be achieved. These new challenges require development of new methods of teletraffic theory, optimization and design. Among others, the special attention should be paid to traffic modeling.

Due to the circuit switching technology utilized by second generation (2G) mobile systems, their performance evaluation have often been limited to call (session) level performance parameters like new call blocking, handover call blocking etc. NG All-IP networks are supposed to use packet-switching technology, and therefore, mobile users may not occupy constant bandwidth during the whole duration of a session. Therefore, these systems should potentially benefit from statistical multiplexing at the network layer. As a result, the focus of users' teletraffic modeling should be extended to include call and packet level models of single traffic sources. These models are useful to study the performance of the network and to predict the QoS expectations that a particular applications may experience at different levels of wireless link congestion.

A survey of research papers have shown that most traffic models designed for 2G mobile systems do not take into account mobility behavior of single user. Those models assume a stationary behavior of large population of users and try to capture call level parameters by distributions like exponential one or its mixtures (see [1], [2] and references therein). Once we have to consider user in isolation, the assumption of stationary characteristics of large population must be dropped. Indeed, stationary call level characteristics, while were shown to be fair predictors of large populations, are not appropriate for single user, where the mobility may significantly affect call level parameters, and as a result, packet level teletraffic characteristics.

To date the only example of traffic model, which simultaneously captures mobility and teletraffic demands of a single user, is a study presented by Antunes *et. al* [3]. Based on general assumptions regarding teletraffic and mobility of the user, they propose to use a combination of two Markovian arrival processes (MAP), one of which describes the mobility behavior while another one specifies call level teletraffic demands. Their model is quite complex, and therefore, its application in performance evaluation of 4G All-IP networks is rather limited.

In this paper we develop an integrated traffic model of packetized variable bit rate (VBR) traffic source for NG All-IP mobile networks. Our model consists of two major parts: mobility model and teletraffic model. Teletraffic model, in turn, consists of call level and packet level models. The mobility part captures mobility behavior of single user and modelled by a Markov chain with finite state space. Teletraffic part captures teletraffic characteristics at both session and packet levels and represented by the superposition of the Markov chain and the discrete-time batch Markovian arrival process (D-BMAP). Mobility and teletraffic parts are integrated, so that the whole model can be seen as triply stochastic process while does not lead outside the D-BMAP framework. This is due to the special superposition of parts, which makes the teletraffic model to be probabilistic function of mobility one, i.e. transition probabilities of Markov chain modeling call level teletraffic demands depend on the state of Markov chain that represents mobility behavior of the user.

Our paper is organized as follows. We make remarks on traffic modeling issues in NG All-IP networks in Section II. In Section III we consequently define mobility model, call level teletraffic model and packet level teletraffic model. We integrate them into one composite model in Section IV. Behavior of the model is considered in Section V. In Section VI we give numerical examples of the proposed model. Call and packet level performance evaluation are outlined in Section VII. Conclusions and further work are outlined in last section.

II. TRAFFIC MODELING ISSUES IN NG ALL-IP NETWORKS

A. *Mobility and call level teletraffic demands*

To illustrate the influence of user's mobility on its teletraffic demands let us define two classes of user's mobility behavior. In accordance with our classification a mobile user can be in one of the following states:

- Fixed: staying, sitting, etc.;
- Nomadic: walking, driving, etc.

Assume that when a mobile user is in the fixed state he uses the network services with probability p^F, $p^F \in (0,1)$, and with probability $(1-p^F)$ stays idle. Similarly, when a mobile user is on move he accesses services with probability p^N, $p^N \in (0,1)$, and does not use the network with complementary probability $(1-p^N)$. Indices F and N stand for fixed and nomadic states respectively. In what follows we explain why it may the case that the call level teletraffic demands depend on the mobility of the user which leading to $p^F \neq p^N$.

To date voice service has been a dominating application in 2G mobile systems. Due to physiological tests the aural information does not require too much attention to acquire it (up to 20%), and therefore, the moving user is aware of the situation around even when moving (driving, walking etc.). Therefore, to date one could assume $p^F = p^N$.

The situation changes with introduction of multimedia services. For example, the visual information requires up to 80% of human attention to acquire it. Moreover, the combination of aural and visual information dominates with overall percentage up to 97%, and therefore, may affect our attention significantly. When the aural information is presented in conjunction with other medias it is fair to assume that moving users will not use such services frequently. It leads to $p^F \neq p^N$ and dependence between mobility and call level teletraffic demands is evident. Additionally, in some countries there are already restrictions imposed on usage of mobile phones when driving. It also leads to $p^F \neq p^N$ even for conventional 2G systems.

We have to note that both p^F and p^N implicitly define user's traffic demands at the call level, and therefore, traffic demands at the packet level. Additionally, it is easy to see that the dependence of teletraffic demands on mobility behavior cannot be generally represented by deterministic functional relationship but has a probabilistic nature.

B. *Call level and packet level teletraffic demands*

3G mobile systems, which were recently given a lot of attention, are nowadays seen as an intermediate step between conventional 2G mobile systems and NG All-IP networks.

Due to the circuit-switched technology utilized by 2G systems, a user of the mobile system is assigned a constant bit rate channel during the cell dwell time (time, for which a particular user stays in a certain cell). When a free channel is not available, the user is forced to leave the system. Handover calls between cells are served almost similarly: the handover call to a new cell is assigned one of the free channels during the cell dwell time. If all channels in the new cell are busy, the handover call is terminated. Therefore, QoS provided to users by 2G mobile systems is typically limited to call level parameters e.g., new call blocking, handover call blocking etc. Correspondingly, traffic models designed for 2G systems are primarily call-oriented ones and capture call level parameters of traffic sources.

NG All-IP networks are supposed to primarily use packet-switching technology. As a result, one can expect that most calls, including voice ones, will be IP-based and require different, often variable rates. Thus, dealing with NG All-IP networks in addition to call level QoS parameters we have to provide IP level QoS guarantees e.g., packet loss, packet delays, delay jitter etc. To characterize performance of user services in NG All-IP networks, user's traffic models that simultaneously captures both call-level and packet level teletraffic characteristics are needed. An example of correspondence between the call level and packet level traffic demands for simple voice codec with silence suppression capabilities is shown in Fig. 1, where gray rectangles denote time when the session is on, black rectangles indicate the talkspurt periods. One can note that there is a strict correspondence between the calling activity of the user and its packetized traffic, i.e. when the session is off no packet are generated, when the session is on packets are generated according to a given codec. Additionally, we have to note that the parameters of packet source depends on the type of the source only and do not directly depend on mobility behavior of the user. Therefore, we assume that these parameters are constant every time when the session is on irrespective of the mobility state of a user.

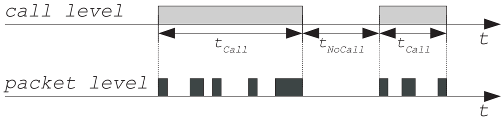

Fig. 1. Correspondence between the call level and packet level teletraffic.

III. MOBILITY AND TELETRAFFIC MODELS

A. Mobility model

Assume the same granularity of user's mobility as in Section II-B, and let us consider the case $p^F \neq p^N$. An extension to the case of more general mobility is straightforward.

Assume a discrete-time environment i.e. time axis is slotted with a certain granularity $\Delta = (t_{i+1} - t_i)$, $i = 0, 1, \dots$, and changes of the mobility state of a user can only occur at the slot boundaries. Considering the real mobility behavior of the user one can expect some type of positive correlation between state changes. Roughly speaking, if the user is in the nomadic state (fixed) at the slot i it is more likely he will stay in nomadic (fixed) state at the slot $(i+1)$. In this paper we propose to capture such type of autocorrelation by a simple Markov chain with two states as outline below.

Consider a discrete-time homogenous Markov chain $\{S^M(n), n = 0, 1, ..\}$ defined at the state space $S^M(n) \in \{1, 2\}$, where superscript M stands for mobility, state 1 corresponds to the fixed state and state 2 denotes the nomadic state. Let P^M be transition matrix of this Markov chain:

$$P^M = \begin{pmatrix} p^M_{11} & p^M_{12} \\ p^M_{21} & p^M_{22} \end{pmatrix}. \tag{1}$$

In accordance with proposed model time, user spends in both states (fixed and nomadic), are geometrically distributed with corresponding parameters (1). Therefore, the only parameters we have to estimate to parameterize the model are p_{11} and p_{22}. They depend on both environmental characteristics of the landscape and user preferences, and therefore, should be chosen with care. For example, if we consider a highway scenario it is wise to set $p_{22} = 1$ and our stochastic model degenerates to deterministic one.

For our model the autocorrelation of user's mobility is geometrically distributed with parameter given by the non-unit eigenvalue λ^M_2 of P^M. Note that λ^M_1 is always one since P^M is the one-step transition matrix of discrete-time homogenous ergodic Markov chain.

One can note that the model can be extended to capture more general distributions of sojourn times in both fixed and nomadic states including sum of geometrical, hypergeometrical and discrete-time phase distributions. This can be done by allowing more than one state of the Markov chain to denote each mobility state of the user. However, the minimum time the user stays in arbitrary state is still Δ. So that the careful choice of Δ is of high importance.

One may also define more general granularity of user's mobility behavior. For example three-states Markov chain with fixed, walking, driving states can be easily defined. However, while the correlation between two-states mobility behavior and usage of the network is evident, comprehensive studies are needed to judge whether it is necessary to use more general granularity.

B. Call level teletraffic model

So far we have defined model with two mobility states: fixed and nomadic. Let us consequently define call level teletraffic models for each of those states. According to our assumptions, call level teletraffic models must be probabilistic functions of user's mobility behavior.

Assume that the user is in the fixed state. Consider an arbitrary time slot n and assume that in the beginning of this slot a mobile user is in the fixed state and the call is active (session is on). Assume then that at the beginning of the next time slot a mobile user still remains in the calling mode with probability p^F_{11}, $p^F \in (0, 1)$, and with probability p^F_{12}, switches to the idle mode (session terminates), where $p^F_{12} = (1 - p^F_{11})$, i.e. time period for which a user stays in the calling mode is geometrically distributed. Similarly, assume that if a mobile user is in the fixed state and the call is not active in the arbitrary time slot n, at the beginning of the next time slot a call is still not active with probability p^F_{22}, $p^F_{22} \in (0, 1)$, and with complementary probability p^F_{21}, $p^F_{21} = (1 - p^F_{22})$, a mobile user switches to the calling mode. These probabilities constitute a simple Markov chain with two states as follows. Consider a discrete-time homogenous Markov chain $\{S^F(n), n = 0, 1, ..\}$ defined at the state space $S^F(n) \in \{1, 2\}$, where superscript F stands for fixed state, state 1 denotes the calling state and state 2 identifies the idle state. Let P^F be transition matrix of the Markov chain in the following form:

$$P^F = \begin{pmatrix} p^F_{11} & p^F_{12} \\ p^F_{21} & p^F_{22} \end{pmatrix}. \tag{2}$$

One can note that in accordance with our model session and idle times are geometrically distributed with corresponding parameters (2). Similarly to mobility model we can generalize these sojourn times to arbitrary mixture of geometric components.

Given the similar assumption of geometrically distributed sojourn times in the calling and idle states we define a discrete-time homogenous Markov chain $\{S^N(n), n = 0, 1, ..\}$ for nomadic state of the user governed by the transition probability matrix P^N in the following:

$$P^N = \begin{pmatrix} p^N_{11} & p^N_{12} \\ p^N_{21} & p^N_{22} \end{pmatrix}. \tag{3}$$

To parameterize both models we have to provide four parameters p^i_{11}, p^i_{22}, $i \in \{N, F\}$. According to our assumptions the value of p^N_{22} must never be less than p^F_{11}. Generally, it is not easy to provide accurate estimates of p^i_{11}, p^i_{22}, $i \in \{N, F\}$ since they depend on the user preference to make more calls in fixed state. However, the inequality $p^F_{22} > p^F_{11}$ clearly indicates that if we neglect dependence between mobility and teletraffic demands, the required resources of the mobile system may be overestimated leading to unwarranted investments to the system infrastructure. The choice of values p^i_{11}, p^i_{22}, $i \in \{N, F\}$ is up to networks operators and should be chosen as an estimators of user's behavior. For example, these estimators can be derived from users' questionnaires. Then, in order to determine the required capacity of NG All-IP mobile systems, these estimators should be used to parameterize the proposed model and determine actual values of traffic offered by mobile users.

C. Packet level model

There are a number of packet level traffic models developed for wired networks, each of which targeted on certain type of source. In our paper we assume that the packet arrival process from VBR source is represented by the discrete-time batch Markovian arrival process (D-BMAP) with two states of the underlying Markov chain.

The reason to use D-BMAP is twofold. Firstly, D-BMAP allows to model complex VBR traffic sources. It has been shown that voice, video and data traffic can be reasonably well represented by D-BMAP [4], [5], [6]. Secondly, the queuing of D-BMAP is well studied, and quite general results have already been obtained [7], [8], [9]. Therefore, the usage of D-BMAP provide a required versatility of the modeling environment and allows to preserve analytical tractability.

Consider a discrete-time homogenous Markov chain $\{S(n), n = 0, 1, \dots\}$ defined at the state space $S(n) \in \{1, 2\}$ and let P^P be its transition probability matrix. Let then $\{W(n), n = 0, 1, ..\}$ be the D-BMAP arrival process whose underlying Markov chain is $\{S(n), n = 0, 1, \dots\}$. We define D-BMAP as a sequence of matrices $D(k)$, $k = 0, 1, \dots$, each of which contains probabilities of transition from state to state with $k = 0, 1, \dots$ arrivals respectively:

$$d(k)_{ij} = \lim_{n \to \infty} Pr\{W(n) = k, S(n) = j | S(n-1) = i\}, \tag{4}$$

where $k = 0, 1, \dots, i, j \in \{1, 2\}$.

One can note that in accordance with definition, it is allowed for D-BMAP to have different conditional probability distribution functions $d(k)_{ij}$, $k = 0, 1, \dots$ for every different pair of states i and j. However, for our model we allow only those arrivals, when the state of modulating Markov chain changes from 1 to 1. Probabilities of other arrivals are set to zero. As a result, we obtain a simple process widely used for VBR traffic sources with ON-OFF nature. We assume that parameters of this process are stationary ones, depend on the type of the source only and do not directly depend on the mobility of the user. In what follows we assume that every time when the session is on packets are generated according to the proposed D-BMAP model, when the session is off no packet are generated.

IV. INTEGRATED MODEL

A. Call level integrated model

Let us firstly integrate the model of user's mobility behavior and model of the call level traffic demands. In fact, we make mobility and call level teletraffic models to be parts of the composite model in accordance with analytical paradigm proposed in [10].

According to our assumptions, call level teletraffic models must be probabilistic functions of user's mobility:

$$S^C(n) = f_{Pr}\left(S^M \in \{1, 2\}\right), \tag{5}$$

where $S^C(n)$ is the state of the composite model at the call level and index Pr denotes probabilistic relationship. Therefore, the choice of appropriate call level teletraffic model must depend on the current mobility state of the user:

$$\{S^I(n)\} = \begin{cases} \{S^F(n)\}, & S^M(n) = 1, \\ \{S^N(n)\}, & S^M(n) = 2. \end{cases} \tag{6}$$

Assume that the call level teletraffic demands of the user is generated according to Markov chains presented in Section III-B and choice of the model depends on mobility state of the model presented in Section III-A. Given that both Markov chains representing the call level teletraffic demands and associated with states of Markov chain of the mobility model have the same number of states, the Markov chain of integrated model can now be defined on the state space S^I as follows:

$$S^I \in \{(1, 1), (1, 2), (2, 1), (2, 2)\}, \tag{7}$$

where the first index in the state description denotes the state of the mobility model while the second one is the state of corresponding call level teletraffic model.

Note that in (7) we still deal with one dimensional Markov chain. To show it we can just re-enumerate the state space of resulting model. However, state description given by pairs (i, j), $i, j \in \{1, 2\}$ is simply more convenient and remains that the integrated model is just a special superposition of the mobility and teletraffic models.

It means that the choice of transition probabilities of Markov chains of integrated model must depend on the state of mobility model:

$$p^T_{i,j} = \begin{cases} p^F_{ij},\ i, j \in \{1, 2\}, & S^M(n) = 1, \\ p^N_{ij},\ i, j \in \{1, 2\}, & S^M(n) = 2. \end{cases} \tag{8}$$

Note that p^N_{ij}, p^F_{ij}, $i, j \in \{1, 2\}$ are different as far as they belong to different Markov chains. Therefore, session holding times as well as idle times are differently distributed in fixed and nomadic states.

Considering (7) and (8), one may note that an appropriate call level teletraffic model must only be associated with those states of the mobility model that correspond to the appropriate mobility state. Therefore, transition probability matrix of the of the integrated model is given by Kronecker product:

$$P^I = \left(\begin{array}{ccccc} & (1,1) & (1,2) & (2,1) & (2,2) \\ (1,1) & p^M_{11}p^F_{11} & p^M_{11}p^F_{12} & p^M_{12}p^N_{11} & p^M_{12}p^N_{12} \\ (1,2) & p^M_{11}p^F_{21} & p^M_{11}p^F_{22} & p^M_{12}p^N_{21} & p^M_{12}p^N_{22} \\ (2,1) & p^M_{21}p^F_{11} & p^M_{21}p^F_{12} & p^M_{22}p^N_{11} & p^M_{22}p^N_{12} \\ (2,2) & p^M_{21}p^F_{21} & p^M_{21}p^F_{22} & p^M_{22}p^N_{21} & p^M_{22}p^N_{22} \end{array}\right). \tag{9}$$

Our integrated traffic model, which is actually represented one-dimensional Markov chain, is simply probabilistic function of Markovian model of the user's mobility. Thus, the whole call level teletraffic model can be seen as a doubly stochastic process while does not lead outside one-dimensional Markov chain framework. It is also clear, that those cases when there are non-probabilistic functional relationships between the mobility behavior and teletraffic demands are the special cases of the model presented here.

Additionally, we have to recall that the definition of our integrated model can be extended by allowing more general models of both mobility behavior and call level teletraffic demands. However, when defining such models one have to note that due to the Kronecker product of transition probability matrices the state space of resulting Markov chain of integrated model grows fast.

B. Packet level integrated model

Assume that the user's packet teletraffic demands is modeled by the D-BMAP model presented in Section III-C. To integrate the packet level teletraffic model to already obtained composite call level teletraffic model we assume that each state of the latter model is the macrostate of the new one as shown in Fig. 2, where p^P_{ij}, $i,j \in \{1,2\}$ are transition probabilities of the packet level teletraffic model, $(l,m) \in \{S^I\}$ is the state of the call level integrated teletraffic model.

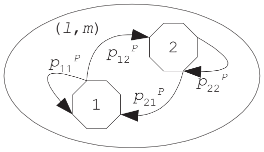

Fig. 2. Macrostate of packet level integrated traffic model.

Note that these additional states representing packetized traffic are only required in those states of integrated call level model in which the state of call level part is 1 (session is on). However, it is not possible to integrate these states without new parameterizing of the resulting model.

To decrease the complexity of parameterization we propose to superpose call level integrated model obtained earlier and D-BMAP model of packetized source. Indeed, it is possible since both models have Markovian structure. In accordance with superposition, states of packet level teletraffic model are integrated to *each state* of call level integrated model and the state space of new superposed model is:

$$S^I_{VBR} = ((1,1,1), \dots, (1,2,2), (2,1,1), \dots, (2,2,2)), \quad (10)$$

where the first index is the state of the mobility model, second one is the state of the call level teletraffic model and the last one – the state of the packet level teletraffic model.

To determine transition probabilities of new integrated packet level model we have to define Kronecker product of transition probability matrices of call level integrated model and underlying Markov chain of VBR packetized traffic model. At the last step we have to set all probabilities of arrivals when the Markov chain changes its state to and from other states, than those of the set $\{(i,j,1)\}$, $i,j \in \{1,2\}$, to zero. We have to note that despite of the triple description of states in (10) we are given a D-BMAP process.

Extension to the case of more complicated packetized traffic is also possible. However, since the complexity of the whole model depends on the complexity of the packetized traffic source, the limiting factor is the state space of underlying Markov chain of the resulting model.

V. BEHAVIOR OF THE PROPOSED MODEL

Sate transition diagram of the call level integrated teletraffic model is shown in Fig. 3. One can note that any state of the proposed model can be reached from any other state with probability, expressed through transition probabilities of initial mobility and call level teletraffic model. Additionally, there are no restrictions imposed by mobility of the user on its calling activity except for underlying stochastic modulation.

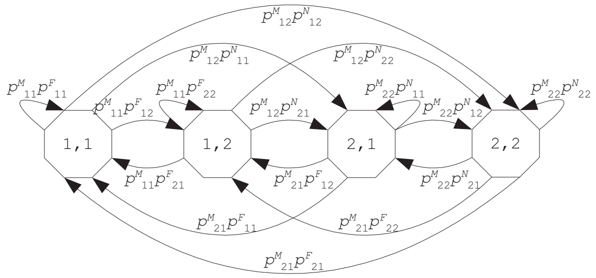

Fig. 3. Transition diagram of the call level integrated model.

To illustrate the behavior of the packet level teletraffic model let us consider its microscopic behavior presented in Fig. 4. Recall that our model is defined in discrete-time, slots of the teletraffic and mobility models are synchronized and equal to Δ. Assume that the model is in the state $(1,1,1)$ at the n^{th} time slot. It means that the user is in the fixed state, the call is on, and packet is generated. According to our model the session completion probability in this state is $p^F_{12} = (p^M_{11}p^F_{12} + p^M_{12}p^N_{12})$, while change to nomadic state occurs with probability $p^M_{12} = (p^M_{12}p^N_{11} + p^M_{12}p^N_{12})$. Suppose the state of the packet level integrated teletraffic model changes to $(2,1,1)$ at the slot boundary between n and $(n+1)^{th}$ slots. It means that the state of the mobility model changes from fixed to nomadic, while the states of the call and packet level teletraffic models remain the same. Therefore, the call is still on and one packet is emitted by the source as shown by the gray rectangle and the arrow respectively. However, the session completion probability is now given by $p^N_{12} = (p^M_{21}p^F_{12} + p^M_{22}p^N_{12})$ instead of $p^F_{12} = (p^M_{11}p^F_{12} + p^M_{12}p^N_{12})$. According to our model $p^N_{12} > p^F_{12}$, and therefore, the call is more likely to get off at the next slot boundary compared to the previous slot boundary when the mobility model was in the fixed state. It occurs and the model changes its state from $(2,1,1)$ to $(2,2,2)$. Since states of both the call level and the packet level teletraffic models change from 1 to 2, session is off and no packets are generated.

VI. NUMERICAL EXAMPLES OF THE PROPOSED MODEL

In this section we provide examples of the proposed teletraffic model. We have to note that our examples do not represent real calling behavior of the user. This is due to the fact that in real environment calling periods are significantly shorter compared to silence periods, and therefore, real models are not convenient for our illustrative purposes.

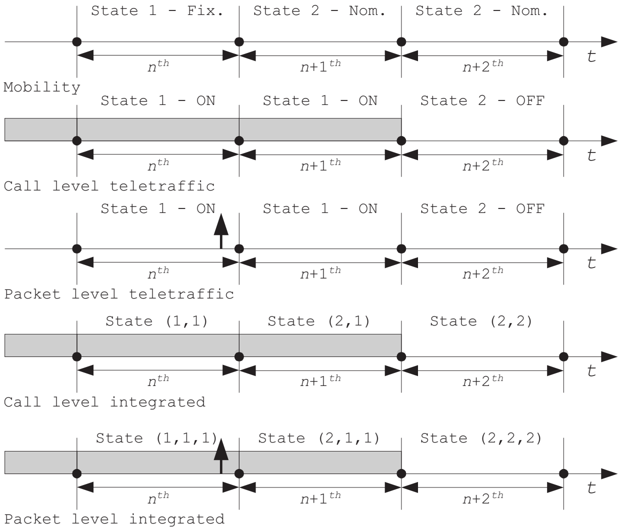

Fig. 4. Microscopic behavior of mobility, call level teletraffic, packet level teletraffic, call level integrated and packet level integrated models.

Our first example illustrates the influence of user's mobility behavior on its teletraffic demands. The transition probability matrices of the Markov chain representing the mobility behavior were chosen as follows:

$$P_1^M = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \qquad P_2^M = \begin{pmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{pmatrix}. \tag{11}$$

We choose the following transition probability matrices of the call level model for both fixed and nomadic states:

$$P^F = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}, \qquad P^N = \begin{pmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{pmatrix}. \tag{12}$$

We have to note that in order to visually illustrate the influence of mobility on teletraffic demands we chose very high session initiation rate in the fixed state (12).

Let A be the event that session is on and let I_A be its indicator i.e. $I_A \in \{0,1\}$. Obtained traces ($I_A = f(\Delta)$) are shown in the Fig. 5, where the upper figure illustrates the trace of the model with P_1^M, the figure below corresponds to the model with P_2^M. Note that in (11) p_{22}^M in P_1^M is nine times higher than p_{22}^M in P_2^M. It must lead to decrease of the mean sojourn time in the nomadic state and, as a result, to increase of the mean sojourn time in the fixed state. According to our model (12) the call initiation rate in the fixed state is three times higher than in the nomadic state. Therefore, the overall call initiation rate must be higher for model with P_1^M than for model with P_2^M. Observing the obtained traces one may note that the overall number of session is actually increasing. Chi-square test used with the level of significance set to 0.05 have shown that these traces are different.

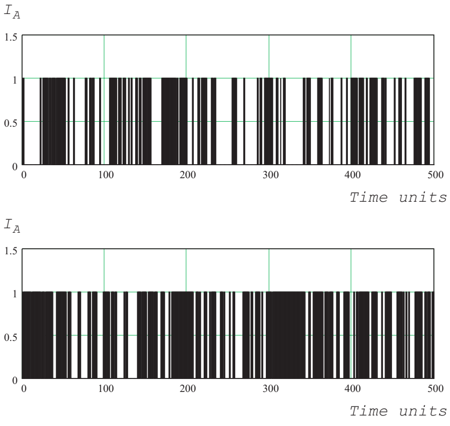

Fig. 5. Traces of the call level integrated model with different P^M.

Consider now how the teletraffic demands influence the behavior of the call level integrated model. To demonstrate it assume that the mobility of the user and its teletraffic demands in fixed state are restricted to Markov chains governed by the following transition probability matrices:

$$P^M = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}, \qquad P^F = \begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix}, \tag{13}$$

while the user's teletraffic demands in nomadic state differs according to the following transition probability matrices:

$$P_1^N = \begin{pmatrix} 0.7 & 0.3 \\ 0.1 & 0.9 \end{pmatrix}, \qquad P_2^N = \begin{pmatrix} 0.1 & 0.9 \\ 0.1 & 0.9 \end{pmatrix}. \tag{14}$$

Traces obtained with these parameters are shown in Fig. 6, where the upper figure illustrate the trace of the model with P_1^N, the figure below corresponds to the model with P_2^N. According to the chosen transition probability matrices the mean time a user spends in the fixed state equals to the mean time he spends in the nomadic state (13). Therefore, the only difference between models is in the session behavior of the user in these states. The session generation rate in both fixed and nomadic states are the same, while the session duration in the nomadic state is higher for model with P_1^N (15). Comparing traces presented in Fig. 6 one can note that the durations of sessions is smaller for model with P_2^N. Chi-square test used with the level of significance 0.05 have shown that these traces are different.

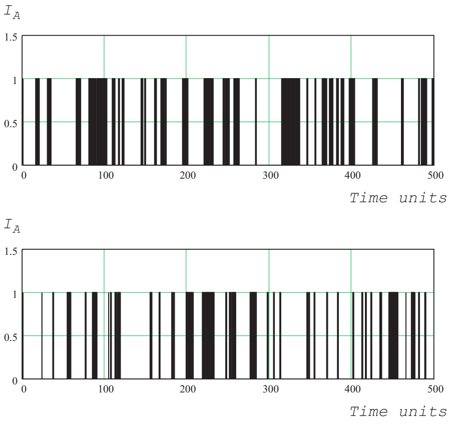

Fig. 6. Traces of the call level integrated model with different P^N.

Let us now consider numerical examples of the packet level integrated teletraffic model. Parameters of the mobility and call level models were set to:

$$P^M = \begin{pmatrix} 0.9 & 0.1 \\ 0.4 & 0.6 \end{pmatrix}, \qquad P^F = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix},$$
$$P^N = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix}. \qquad (15)$$

To represent the packet arrival process we chose a simple Bernoulli process, which is which is the special case of D-BMAP. The only parameters we have to define for Bernoulli process is the probability of one packet arrival $Pr\{k = 1\}$. Obtained traces of the packet level integrated teletraffic model are shown in the Fig. 7, where the upper figure illustrates the trace of the model with $Pr\{k = 1\} = 0.6$, the figure below correspondsto the model with $Pr\{k = 1\} = 1.0$ respectively. Boxes denote packet arrivals. One can note that in latter case we have a constant bit rate traffic model.

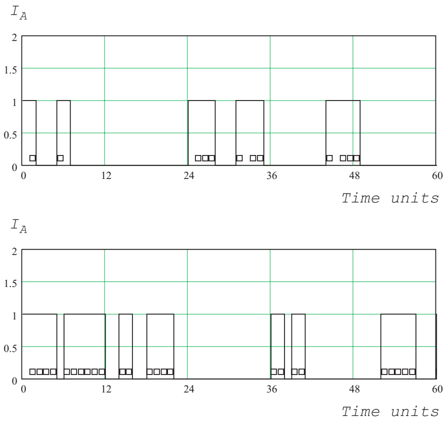

Fig. 7. Traces of the packet level integrated model with different $Pr\{k = 1\}$.

To illustrate the versatility of the model we then used a Batch Bernoulli Process (BBP) to emulate packet arrivals. Probability function (PF) of the process and corresponding trace of the model are shown in the the Fig. 8, where the upper figure illustrates the PF of the BBP, the figure below shows the trace of the model.

VII. PERFORMANCE EVALUATION

Since the proposed model captures both call and packet level characteristics, it can be used in performance evaluation studies at both call and packet levels.

A. Call level performance evaluation

There are a number of call level performance parameters that must be properly adjusted at the planning phase to guarantee an acceptable quality of service provided to the user by a mobile system. One of the most important parameters is the new call blocking probability. To determine the new call blocking probability experienced by a single user the proposed model call level integrated teletraffic model must be used together with model describing the number of busy channels in a given cell. The new call blocking probability can be derived as a joint probability as all available channel in the cell are busy and the proposed model make the transition to those states in which the call is on.

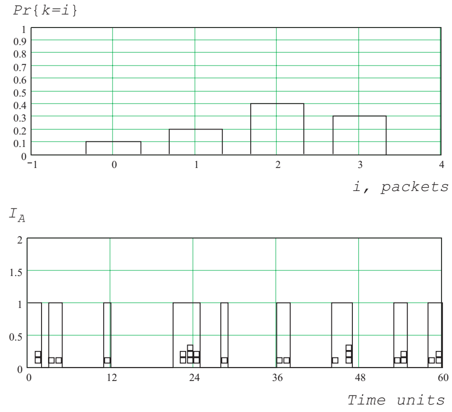

Fig. 8. PDF of the BBP and trace of the packet level integrated model.

B. Packet level performance evaluation

It is well-known that wireless channels have highly dynamic time-varying nature. Mobility of users is one of the major factors that affect the quality of the wireless channels. For example, 3G systems provide different effective transmission rates at the air interface for different mobility states of the user. Indeed, different propagation characteristics experienced by the receiver lead to the usage of error concealment techniques with different correction capabilities. As a result, the effective transmission rate is significantly lower in the nomadic state than in the fixed state. It leads to different requirements for buffer space of mobile terminals which is scare resource nowadays. Additionally, different performance parameters can be experienced by applications in fixed and nomadic states.

Let us assume that the user's teletraffic demands is represented by the model outlined in Section IV with the following parameters:

$$P^M = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}, \qquad P^F = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix},$$
$$P^N = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix}. \qquad (16)$$

To capture difference between transmission rates in the nomadic and fixed states we assume that the service rate of the packets is different in these states. Assume that the service rate of the packets in the nomadic state equals to $1/\Delta$. Then, the service rate of packets in the fixed state is x/Δ, $x = 1, 2, \dots$. To represent the packet level teletraffic demands we use BBP with the following parameters:

$$Pr\{W(n) = 0\} = 0.1, \qquad Pr\{W(n) = 1\} = 0.2,$$
$$Pr\{W(n) = 2\} = 0.4, \qquad Pr\{W(n) = 1\} = 0.3, \qquad (17)$$

where $\{W(n), n = 0, 1, \dots, \}$ is the packet arrival process. It should be noted that due to discrete-time nature of our model packet arrival process is the explicit function of the service rate in the fixed state Δ.

Let us firstly consider how the service rate in the fixed state affects the mean number of packets in the buffer $E[B]$ and the packet loss ratio L. The dependency of these parameters on the service rate in the fixed state are shown in the Fig. 9. The buffer space was set to 9 packets ($K = 9$). One may note that the increase of the service rate in the fixed state results in significant decrease of both $E[B]$ and L.

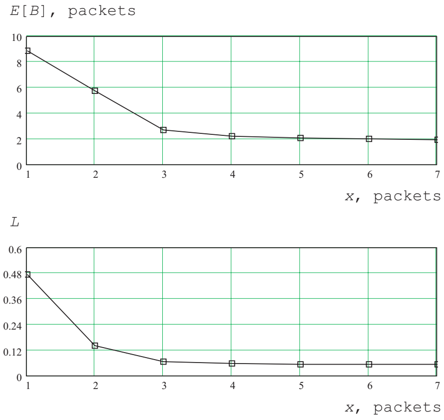

Fig. 9. Mean number of packets in the buffer and the loss ratio.

Consider now the influence of user's mobility on the packet level performance parameters. To illustrate it we set the service rate in the fixed state to $2/\Delta$, $K = 9$. Then we changed the parameters p^M_{11}, p^M_{22} of P^M from 0.1 to 0.9. $E[B]$, L as a functions of p^M_{11}, p^M_{22}, are shown in the Fig. 10. Observing these dependencies one may note that the mobility of the user strictly influence the packet level performance parameters experienced by applications. For example, the packet loss ratio corresponding to $p^M_{11} = 0.1$, $p^M_{22} = 0.9$, is more than twelve times higher than the loss ration corresponding to $p^M_{11} = 0.9$, $p^M_{22} = 0.1$. This leads to the clear conclusion that some applications may perform well in the fixed state and completely fail in the nomadic one.

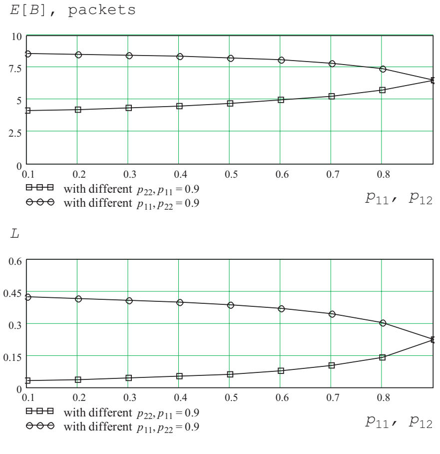

Fig. 10. Mean number of packets in the buffer and the loss ratio.

The loss ratio can be decreased by increasing the buffer size as outlined in the Fig. 11. This dependency were estimated for model with parameters given by (16) and the service time in the fixed state was set to $2/\Delta$.

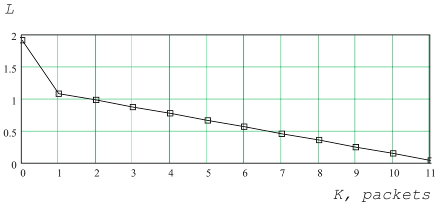

Fig. 11. Loss ratio as a function of the buffer size.

VIII. CONCLUSIONS

In this paper we presented an integrated model of a single packetized VBR traffic source for NG All-IP mobile networks. Our model consists of two parts: mobility model and teletraffic model. The mobility part captures mobility behavior of a single user and modelled by the Markov chain with finite state space. Teletraffic part captures teletraffic characteristics at both session and packet levels and represented by superposition of the Markov chain and the discrete-time batch Markovian arrival process (D-BMAP). Mobility and teletraffic parts are superposed, and the whole model is actually a triply stochastic process while does not lead outside the D-BMAP framework.

Our model retains the analytical tractability and can be used in call and packet level performance evaluation of NG All-IP networks as well as in simulation studies of these networks.

REFERENCES

[1] Y. Fang and I. Chlamtac. Teletraffic analysis and mobility modelling of PCS networks. *IEEE Trans. on Comm.*, 47(7):1062–1072, July 1999.

[2] P. Orlik and S. Rappaport. A model for teletraffic performance and channel holding time charachterization in wireless cellular communication with general session and dwell time distributions. *IEEE JSAC*, 16(5):788–803, June 1998.

[3] N. Antunes, A. Pacheco, and R. Rocha. An integrated traffic model for multimedia wireless networks. *Comp. Netw.*, 38:25–41, 2002.

[4] F. Beritelli, A. Lombardo, and G. Schembra. Performance analysis of an ATM multiplexer loaded with VBR traffic generated by multimode speech coders. *IEEE JSAC*, 17(1), January 1999.

[5] D. Moltchanov, Y. Koucheryavy, and J. Harju. The model of single smoothed MPEG traffic source based on the D-BMAP arrival process with limited state space. In *Proc. of ICACT*, pages 55–60, Phoenix Park, S. Korea, January 2003.

[6] A. Lombardo, G. Morabito, and G. Schembra. An accurate and treatable Markov model of MPEG video traffic. In *Proc. of IEEE INFOCOM*, pages 217–224, 1998.

[7] N. Kim, S. Chang, and K. Chae. On the relationships among queue length at arrival, departure, and rendom epochs in the discrete-time queue with D-BMAP arrivals. *Operations Research Letters*, 30:25–32, 2002.

[8] K. Kang and C. Kim. Performance analysis of statistical multiplexing of heterogeneous discrete-time Markovian arrival processes in an ATM network. *Computer Communications*, 20:970–978, 1997.

[9] D. Moltchanov, Y. Koucheryavy, and J. Harju. Superposed and perprocess analysis in $\sum$D-BMAP/D/1/*K* queuing system. In *Proc. of HET-NET'2003*, pages 42/1–42/11, UK, Ilkley, July 2003.

[10] A. Lombardo and S. Schembra. Characterization of intermedia synchronization and its effects on the performance of an ATM multiplexer loaded by multimedia traffic. *Perf. Eval.*, 35(1–2):75–95, March 1999.