Performance Analysis of Wireless Networks Supporting High Speed Data Services

July 9, 2005

Abstract

This paper deals with the analytical modeling of the wireless networks supporting high speed data services. A call admission control scheme with fixed resource reservation based on quality of service (QoS) and priority is proposed. In this scheme, voice, video and data traffic are considered to be arriving according to Markov arrival process (MAP) as it allows for the correlation among the inter arrival times of the incoming calls. The channel holding times and the cell residential times are assumed to be phase type distributed. Each type of traffic is **distinguished** as handoff and new calls. Also, the high priority hard QoS (video1) and low priority soft QoS (video2) video calls are separated. Further, in this scheme, voice calls have the maximum priority followed by video1, video2 and data calls. The underlying stochastic process for the cell behavior is a quasi birth death process. The matrix analytic approach is applied to obtain the important performance measures such as call blocking and call dropping probabilities and total carried traffic. Our work can be used for dimensioning of the wireless networks with realistic traffic.

Keywords: Wireless networks, Markov arrival process, call blocking and call dropping probabilities, matrix analytic methods.

1 Introduction

The future wireless networks consisting of voice, data and video traffic promised to bring subscribers the integrated wireless optical and internet technologies for "always on" access to multimedia contents and services using portable and wireless devices with QoS guarantee. The wireless mobile networks are expected to support wideband data and video services in addition to voice service. To provide QoS guarantee in a wireless network environment, the bandwidth is required to be available wherever the user goes on. The bandwidth of wireless link is inherently limited and is generally much less than its wireline counterpart. Call admission control (CAC) schemes are generally employed in QoS provisioning for wireless communication systems. The arriving calls are granted/denied access to the network by the CAC schemes based on the predefined criteria taking the network loading conditions into considerations. CAC schemes plays a critical role in QoS provisioning in terms of signal quality, call blocking and call dropping probabilities, packet delay loss rate and transmission rate. A comprehensive survey on CAC in wireless networks is given in [13].

Many CAC schemes have been proposed for voice calls only, for voice and data calls, and voice, video and data calls and the various performance measures are determined. In the first and second generation wireless networks, CAC has been developed for a single service environment,that is, consisting of voice calls only. The analytical model for the CAC schemes in the wireless networks supporting only voice calls is discussed in [1, 6, 25]. Haring et al [1] discussed the CAC schemes for voice calls with the exponential inter arrival and channel holding times with fixed resource reservation for the handoff calls. Choi et. al. [6] have used the quasi birth and death(QBD) process to present the analytical model for the cut off priority scheme, with buffer for handoff and new voice calls. Dharmaraja et. al. [3] have obtained the performance measuresfor the wireless networks with generally distributed handoff inter arrival times. Jayasuriya et. al. [4] have carried out the simulations for the wireless networks with voice calls and found that the service time is best modeled using phase type distributions.

Alfa and Li [26] have given the analytical model for the wireless networks in which the voice calls are arriving according to Markov arrival process(MAP). Also, the call residence time and requested call holding time have been modeled as the general phase type distributions. The increasing number of customers and network complexity, customer behaviour in general and retrial phenomenon have all had significant impact on the network performance. **There is a large number of research papers dealing with customer retrials attempts [15, 16, 27, 18] in the queueing model aspects.** The literature on retrial queues has been summarized in [17]. Alfa and Li [9] have used the matrix analytic approach to describe the performance model with voice calls arriving according to MAP, with service and retrial times follow phase type distributions.

In the third generation wireless networks, multimedia services such as voice, video and data calls are offered with various QoS profiles. Hence, more sophisticated and efficient schemes are developed to cope with these changes. Enormous work has been done by the researchers on the performance evaluation of the CAC schemes for wireless networks with voice and data calls [7, 8]. Randhawa and Hardy [23] have presented an analytical model for the real and non real time traffic. They have assumed the exponentially distributed inter arrival times and channel holding times for the voice and data calls. The third generation wireless traffic is composed of voice and internet traffic. It has been considered in [19, 20, 21] that the internet traffic shows the property of self similarity and burstiness. **Thus, the future generation wireless traffic is expected to possess the properties of self similarity and burstiness. The traffic has to be modeled so that the self similar nature and burstiness are taken into consideration.** The self-similar nature of the wireless networks causes the correlation to exist in the inter arrival times of the voice, data and video calls, that makes the assumption of exponential inter arrival times and channel holding times invalid.

of phase

To measure the system performance in a realistic situation, it is important to model the traffic arrival process and channel holding times with the general distributions. In this paper, we propose the analytically tractable model of wireless networks with high speed data services consisting of voice, data and video calls. We have used Markov arrival process for the call arrivals as MAP has the favourable properties like the smoothening of the total traffic by the statistical aggregation of multiple MAP's which are individually burst in nature. Phase type distributions have been used to model channel holding times and retrial times of the calls.

Rest of the paper is organized as follows: A detailed description of the cellular networks with voice, data and video traffic with retrial phenomenon is given in section 2. In section 3, the CAC scheme has been proposed. In section 4, by transforming the stochastic behaviour of the cell into a quasi birth and death process, the explicit expression of the infinitesimal generator matrix governing the cell is obtained. Section 5 discusses the important performance measures which are essential to analyze the network efficiency. In section 6, the particular cases of this work are presented in detail. Finally, section 7 provides some concluding remarks.

2 Model Description

In the telecommunication systems, the service area is populated with base stations. The coverage area of each base station is called cell and hence the name cellular networks. We assume that cells in a network are statistically identically homogeneous with uniform traffic within the cells. In this performance model, the incoming traffic has been classified as voice(ve), video1(vo1), video2(vo2) and data(dt) calls. For each traffic, two types of connections (new and handoff) may appear at the base station. Fixed guard channel scheme has been proposed so as to provide the continuous connectivity to mobile users. The retrials of the blocked calls have the negative influence on the fresh and handoff calls being connected as the offered load becomes higher. Here, the calls arriving to find at least g channels busy (g varies for different traffic) goes to "orbit" and keep on retrying for free channels to get connected.

A general phase type distributed cell residence times and call holding times are assumed. The call arrival process for different traffic is modeled as follows:

The voice calls, new and handoff, both are modeled using a continuous time MAP as it allows for the correlation of the inter arrival times of new and handoff calls. As discussed in the previous section, the bursty and self similar nature of the arriving traffic allows for the existence of correlation among the inter arrival times of new and handoff calls.

Data calls are assumed to arrive according to three state MAP as mobile station can be in any one of the three different modes that are sleep, wake and active.

Finally, we consider the video calls. Video calls has been distinguished as hard QoS video calls referred as video1 calls, e.g., video conferences, live programs etc. and soft QoS video calls, e.g., video on demand, referred as video2 calls. Like data and voice calls, video1 and video2 calls have also been classified as new and handoff calls. Video calls arrival also show the self similar nature and a need to model the traffic according to a distribution that can capture the correlation among the call arrivals arises. We assume that the arrival processes of video1 and video2 traffics are MAPs with different parameters.

Some other detailed assumptions are as follows:

1. **The vehicular mobility is characterized by the cell residence time R_s of a vehicle in a cell. The random variable R_s is** assumed to have a general phase type distribution with representation (α, T) and dimension r. The residual residence time $\overline{R}_S$ is also phase type distributed with representation $(\overline{\alpha}, T)$ and dimension r[11].

2. The requested call holding times $H_i^N (i = ve, vo, d)$ of new calls are the duration of the requested new call connection to the network, are independent and identically distributed random variables with phase type distribution and having the representation (β_i^N, S_i^N) and dimension h_i^N. **The residual call holding times $\overline{H_i^N}$ are also phase type distributed with representation $(\overline{\beta_i^N}, S_i^N)$ and and dimension h_i^N. Note that video1 and video2 calls are not distinguished once connected.**

3. The actual service time of new calls of type i, defined by $S_i^N = min(H_i^N, \overline{R}_s)$ is the minimum of the call holding time and residual residence time. In light of the results in [11], S_i^N is phase type distributed with representation $(\delta_{N_i}, L_{N_i}) = (\overline{\alpha} \otimes \beta_i^N, T \oplus S_i^N)$ and dimension rh_i. Similarly, actual service time of handoff calls of type i defined by $S_i^H = min(\overline{H}_i^H, R_s)$ as the minimum of the residual call holding time and cell residence time is phase type distributed with representation $(\delta_{H_i}, L_{H_i}) = (\alpha \otimes \overline{\beta}_i^H, T \oplus S_i^H$ and dimension rh_i^H. To avoid the computational complexity, we assume that S_i^N, S_i^H are identical and is taken to be S_i for each traffic class i. Note that $\otimes$ denotes the kroneckar product and $\oplus$ is the kroneckar sum.

4. Calls namely new voice(nve), handoff video2(hvo2), new video2(nvo2), new and handoff data(nd, hd), when on arriving find insufficient bandwidth to serve them joins the orbit. Let $\lambda_i(i = nve, hvo2, nvo2, hd, nd)$ be the arrival rate of calls of type i (to be discussed later in this section) and p_i be the probability with which calls of type i retries from the orbit. Then, p_i can be computed as $\frac{\lambda_i}{\Sigma_i \lambda_i}$. The calls in the orbit retry after a random interval R_t which is again having phase type distribution with representation (δ, L) and dimension n.

5. The calls in the orbit will keep on retrying until it is successful in obtaining a channel or it departs from the cell. Consider a random variable $R = min(R_t, \overline{R}_s)$ also follows the phase type distribution represented as $(a_r, A_r) = [\overline{\alpha} \otimes \delta, T \oplus L]$. Here R represents the interval to the next retrial or departure from the cell, which ever happens first. There are two absorbing states in this distribution- one absorbing state corresponds to the state of departure from orbit as cell residence time completes and the other absorbing state corresponds to the departure from the orbit due to successful connectivity. Absorption to the states of connectivity has rates given by the vector $A_r^o(2) = e \otimes L^o$ and the rates of absorption to the states of departure from the cell $A_r^o(1) = T^0 \otimes e$. Hence $A_r e + A_r^o(1) + A_r^o(2) = 0$.

6. Total voice call arrival process is modeled by continuous time Markov arrival process which we call MAP1. Then using the methodology in [9], let C^{ve} be the infinitesimal generator matrix of the K state Markov process and the sojourn time in state i is exponentially distributed with parameter $\lambda_i, i = 1, 2, \ldots, K$. The new voice and handoff call arrival rates are given by $\lambda_{nve} = \pi C_n^{ve} e$ and $\lambda_{hve} = \pi C_h^{ve} e$ respectively, with, $C^{ve} = C_0^{ve} + C_n^{ve} + C_h^{ve}$. Similarly, we model the video1, video2 and data calls total arrival process as MAP2, MAP3, MAP4 respectively, with the corresponding arrival rates denoted as $\lambda_{hvo1}, \lambda_{nvo1}, \lambda_{hvo2}, \lambda_{nvo2}, \lambda_{hd}, \lambda_{nd}$.

3 Call Admission Control Scheme

In the proposed CAC scheme, the voice traffic is accorded the highest priority followed by video and then data calls. The total traffic arriving in the network is classified, in decreasing order of priorities as handoff voice, new voice, handoff video1, new video1, handoff video2, new video2, handoff data and new data with notations hve, nve, hvo1, nvo1, hvo2, nvo2, hd and nd respectively. We assume that voice and data calls are assigned only one channel whereas the video calls are assigned four channels.
Fixed number of channels h_1, h_2, h_3 and h_4 are reserved for handoff calls of voice, video1 video2 and data calls respectively, whereas n_1, n_2 and n_3 channels for new voice, new video1 and new video2 type of calls. Since the new data calls are of least priority and can tolerate the delay, there is no need to reserve the channels for them.

Let M be the total channels available in a cell. We define the following notations:

$$\begin{aligned} M_1 &= M - \Sigma_{i=1}^{3}(h_i + n_i) - h_4 \qquad M_5 = M - h_1 - n_1 - h_2 \\ M_2 &= M - \Sigma_{i=1}^{3}(h_i + n_i) \qquad M_6 = M - h_1 - n_1 \\ M_3 &= M - \Sigma_{i=1}^{2}(h_i + n_i) - h_3 \qquad M_7 = M - h_1 \\ M_4 &= M - \Sigma_{i=1}^{2}(h_i + n_i) \end{aligned}$$

The acceptance of the traffic in a cell is made according to the Table 1.

When the handoff call(voice or video1) arrivals find no free channels to serve them according to their guard channel policy they are dropped. New call arrivals are blocked in accordance with their guard channel policy when no free channels are available.

Busy channels i	Connectivity to call type	Calls going to orbit (size of orbit = J) calls in orbit < J
$0 \leq i < M_1$	all calls allowed	-
$M_1 \leq i < M_2$	hve, nve, hvo1, nvo1, hvo2, nvo2, hd	nd
$M_2 \leq i < M_3$	hve, nve, hvo1, nvo1, hvo2, nvo2	hd, nd
$M_3 \leq i < M_4$	hve, nve, hvo1, nvo1, hvo2	nvo2, hd, nd
$M_4 \leq i < M_5$	hve, nve, hvo1,nvo1	hvo2, nvo2, hd, nd
$M_5 \leq i < M_6$	hve, nve, hvo1	hvo2, nvo2, hd, nd
$M_6 \leq i < M_7$	hve, nve	hvo2, nvo2, hd, nd
$M_7 \leq i < M$	hve	nve, hvo2, nvo2, hd, nd
$i = M$	No calls	nve, hvo2, nvo2, hd, nd

Table 1: Traffic arrivals

4 Analytical Model

The wireless network considered in this paper is a homogeneous network, every cell in this network is considered to be statistically identical and independent of each other. Thus, by analyzing the performance of a single cell, the performance of the whole network can be approximately characterized. Here, we considered the general phase type distributions for cell residence times, channel holding times and retrial times and the Markov arrival process for the traffic arrival of voice, video and data. The traffic arrival process, channel holding times and retrial times are all mutually independent. In terms of the results of the section 2, by transforming the cell's state into a quasi birth and death process, we find the explicit iterative expressions for the infinitesimal generator matrix in this section and then the steady state probabilities and performance measures are obtained in section 5.

Let $\{X(t), t \geq 0\}$ be a stochastic process for a given cell with M and J as the total number of channels and buffer size in the orbit respectively. First, define the set for traffic $m(m = ve, vo, d)$, $\mathbf{S}_i^m = (\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_i)$ with $\mathbf{s}_v = 1, 2, \ldots, hr,\ 1 \leq v \leq i$ and $\mathbf{r}_i = (r_1, r_2, \ldots, r_i)$ with $r_v = 1, 2, \ldots, nr\ 1 \leq v \leq i$. Let i be the number of calls receiving service out of which j and k be the number of ongoing video(1 and 2) and data calls respectively and l be the total number of new voice, handoff video2, new video2, handoff data and new data calls in the orbit. Let $u_{ve}, u_{vo1}, u_{vo2}, u_d$ be the phase of $MAP1, MAP2, MAP3$ and $MAP4$ of voice, video1, video2 and data calls arrival process respectively. $\mathbf{S}^{ve}_{i-j-k}, \mathbf{S}^{vo}_j, \mathbf{S}^d_k$ be the set of phase of service for voice, video(1 and 2) and data calls respectively and $\mathbf{r}_l$ be the set of retrial phase of l calls in the orbit. The state space can be denoted as $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ where the sub-state space are $\Omega_1 = \{0, 0, 0, 0, u_{ve}, u_{vo1}, u_{vo2}, u_d\}$; this corresponds to no call in the system that is neither in the service nor in the orbit, $\Omega_2 = \{0, 0, 0, l, u_{ve}, u_{vo1}, u_{vo2}, u_d, \mathbf{r}_l\}$; this corresponds to no calls receiving service and l calls in the orbit, $\Omega_3 = \{i, j, k, l, u_{ve}, u_{vo1}, u_{vo2}, u_d, \mathbf{S}^{ve}_{i-j-k}, \mathbf{S}^{vo}_j, \mathbf{S}^d_k, \mathbf{r}_l\}$; this corresponds to the state when i calls are ongoing in the system and l calls are waiting in the orbit. Here, $0 \leq i \leq M$, $0 \leq j \leq min\{i, [\frac{M_6}{4}]\}$ say j_m, $0 \leq k \leq min\{i - j, M_2\}$, say k_m and $0 \leq l \leq J$. $[x]$ means the greatest integer function of x. Suppose that the underlying Markov chain of the arrival process of voice, video1, video2 and data calls respectively has K_1, K_2, K_3 and 3 states so that $1 \leq u_m \leq K_m, m = (ve, vo1, vo2)$ and $1 \leq u_d \leq 3$ for data traffic.

The generator matrix for the stochastic process is given below:

$$Q = \begin{bmatrix} Q_{01} & Q_{00} & & & & \\ \ddots & \ddots & \ddots & & & \\ & Q_{i2} & Q_{i1} & Q_{i0} & & \\ & & \ddots & \ddots & \ddots & \\ & & & & Q_{M2} & Q_{M1} \end{bmatrix}$$

where the elements of Q are explained in the following.

$$Q_{i0} = \begin{bmatrix} A^i_{01} & A^i_{00} & & & \\ \ddots & \ddots & \ddots & & \\ & 0 & A^i_{j1} & A^i_{j0} & \\ & & \ddots & \ddots & \\ & & & 0 & A^i_{j_m 1} \end{bmatrix} \quad Q_{i2} = \begin{bmatrix} C^i_{01} & 0 & & & \\ \ddots & \ddots & \ddots & & \\ & 0 & C^i_{j2} & C^i_{j1} & 0 \\ & & \ddots & \ddots & \\ & & & C^i_{j_m 2} & C^i_{j_m 1} \end{bmatrix}$$

and $Q_{i1} = diag(B^i_0, \ldots, B^i_j, \ldots, B^i_{j_m})$.

The blocks Q_{i0} represents the transition in the Markov process when i number of calls in service increases by one. Similarly, Q_{i2} gives the transitions when one of the i calls in service gets completed. Q_{i1} represents the no change in the service domain of the system, that is, the number of ongoing calls do not change due to retrials or call arrivals.

In order to proceed with the analysis conveniently, we introduce the following notations as defined in [9].

Definitions:

- I_w: Identity matrix of dimension w and O_w: Square matrix of zeros of dimension w.
- $I(w,s) = \underbrace{I_w \otimes, \ldots \otimes, I_w}_{s}$ and $O(w,s) = \underbrace{O_w \oplus, \ldots \oplus, O_w}_{s}$
- $W_k(i) = \underbrace{L_k \oplus, \ldots \oplus, L_k}_{i}$. It represents no service completion by any of the i calls of type $k, (k = ve, vo, d)$ in service.
- $W_r(i) = \underbrace{A_r \oplus, \ldots \oplus, A_r}_{i}$. It represents no retrial attempts by any of the i calls in the orbit.
- $\widehat{W_r}(i) = \underbrace{(A^0_r(2)a_r) \oplus, \ldots \oplus, (A^0_r(2)a_r)}_{i}$. It represents unsuccessful retrial by any of the i calls in the orbit.
- $V_k(i) = \sum_{j=0}^{i-1} I(rh_i, j) \otimes L^0_k \otimes I(rh_i, i-j-1)$. It represents a service completion by one of the i calls of type $k, (k = ve, vo, d)$ in service.
- $V_r(i,1) = \sum_{j=0}^{i-1} I(nr, j) \otimes A^0_r(1) \otimes I(nr, i-j-1)$. It represents departure from the cell and orbit by one of the i calls in the orbit without connecting.

- $\widehat{V_r}(i,1) = \sum_{j=0}^{i-1} I(nr,j) \otimes (A_r^0(2) \otimes \delta) \otimes I(nr, i-j-1)$. It represents successful retrial by one of the i calls in the orbit.

First, we consider the blocks Q_{i0}.

1. The block matrix elements A^i_{j1} and A^i_{j0} represents an increase by one in the number of voice or data calls and video calls respectively either due to fresh arrivals of traffic or successful retrials by the calls in the orbit.

   - A^i_{j1} is an upper diagonal block matrix with main diagonal elements A^{ij1}_{k1} and A^{ij1}_{k0} as the upper diagonal block matrix. A^{ij1}_{k1} represents an increase by one in the number of calls in service by one due to fresh arrival or successful retrials by the voice calls when the system is having i calls in service and out of those j are video calls and k are data calls. Likewise, increase in the number of calls receiving service due to fresh arrivals or successful retrials by data calls is represented by A^{ij1}_{k0}. Each of the block matrix A^{ij1}_{k1} and A^{ij1}_{k0} is a lower diagonal matrix with the main diagonal elements as A^{ij1k1}_{l1} and A^{ij1k0}_{l1} respectively, representing the transitions corresponding to the fresh arrivals of the voice and data calls. The transition taking place due to connectivity by the voice and data calls from the orbit are respectively, given by A^{ij1k1}_{l2} and A^{ij1k0}_{l2}. They are the lower diagonal elements of the matrices A^{ij1}_{k1} and A^{ij1}_{k0}.

   The expressions are given as:

   For $i = 0, 1, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ and $l = 0, 1, \ldots, J$

$$A^{ij1k1}_{l1} = \lambda^{i,j}_{VE} \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus W_r(l) \quad (1)$$

$$A^{ij1k0}_{l1} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus \lambda^{i,j}_D \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus W_r(l) \quad (2)$$

   For $i = 0, 1, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ and $l = 1, \ldots, J$

$$A^{ij1k1}_{l2} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus g^{i,j}\widehat{V}_r(l) \quad (3)$$

$$A^{ij1k0}_{l2} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus h^{i,j}\widehat{V}_r(l) \quad (4)$$

   where

$$\lambda^{i,j}_{VE} = \begin{cases} C^{ve}_{hv} + C^{ve}_{nv}, & 0 \le i+3j < M_7 \\ C^{ve}_{hv}, & M_7 \le i+3j < M \\ 0 & \text{otherwise} \end{cases} ; \quad g^{i,j} = \begin{cases} p_{nve}, & 0 \le i+3j < M_7 \\ 0, & \text{otherwise} \end{cases}$$

   and

$$\lambda^{i,j}_{D} = \begin{cases} C^{d}_{hd} + C^{d}_{nd}, & 0 \le i+3j < M_1 \\ C^{d}_{hd}, & M_1 \le i+3j < M_2 \\ 0, & \text{otherwise} \end{cases} ; \quad h^{i,j} = \begin{cases} p_{hd} + p_{nd}, & 0 \le i+3j < M_1 \\ p_{hd}, & M_1 \le i+3j < M_2 \\ 0, & \text{otherwise} \end{cases}$$

2. The upper diagonal element A^i_{j0} of the matrix Q_{i0} represents an increase by one in the number of calls receiving service when video(1 and 2) calls from the orbit or the fresh arrival gets connected. For each i and j, A^i_{j0} is a diagonal matrix, A^{ij0}_{k1} being the diagonal elements. For each i, j and k, A^{ij0}_{k1} is a lower diagonal matrix. The main diagonal element of this matrix are A^{ij0k1}_{l1} represents an

increase in the video calls due to fresh arrivals of the video calls, whereas increment in the number of video calls receiving service due to successful retrials from the orbit are represented by the A_{l2}^{ij0k1}, the system is having i calls ongoing out of which j are video calls, k are data calls and l calls are in the orbit. The calls in the orbit keep on retrying for the free channel until connected or departs from the cell due to cell residence times.

For $i = 0, 1, \ldots, M-1$, $j = 0, 1, \ldots, j_m - 1$, $k = 0, 1, \ldots, k_m$ and $l = 0, \ldots, J$

$$A_{l1}^{ij0k1} = C_0^{ve} \oplus \lambda_{VO}^{i,j} \oplus C_0^d \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus W_r(l) \tag{5}$$

For $i = 0, 1, \ldots, M-1$, $j = 0, 1, \ldots, j_m - 1$, $k = 0, 1, \ldots, k_m$ and $l = 1, \ldots, J$

$$A_{l2}^{ij0k1} = C_0^{ve} \oplus C_0^{vo1} \oplus C_0^{vo2} \oplus C_0^d \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus m^{i,j}\widehat{V}_r(l) \tag{6}$$

where

$$\lambda_{VO}^{i,j} = \begin{cases} (C_{hvo1}^{vo1} + C_{nvo1}^{vo1}) \oplus (C_{hvo2}^{vo2} + C_{nvo2}^{vo2}), & 0 \leq i+3j+4 < M_3 \\ (C_{hvo1}^{vo1} + C_{nvo1}^{vo1}) \oplus C_{hvo2}^{vo2}, & M_3 \leq i+3j+4 < M_4 \\ (C_{hvo1}^{vo1} + C_{nvo1}^{vo1}) \oplus C_0^{vo2}, & M_4 \leq i+3j+4 < M_5 \\ C_{hvo1}^{vo1} \oplus C_0^{vo2}, & M_5 \leq i+3j+4 < M_6 \\ 0, & \text{otherwise} \end{cases}$$

and

$$m^{i,j} = \begin{cases} p_{hhvo2} + p_{nvo2}, & 0 \leq i+3j+4 < M_3 \\ p_{hvo2}, & M_3 \leq i+3j+4 < M_4 \\ 0, & \text{otherwise} \end{cases}$$

Now, we will consider the lower diagonal matrix blocks Q_{i2} of the matrix Q. For each $i = 0, 1, \ldots, M$, Q_{i2} is a lower diagonal block matrix with C_{j1}^i as the diagonal elements and C_{j2}^i as the lower diagonal elements.

1. C_{j1}^i represents a reduction by one in the number of voice or data calls receiving service through the call completion. It can be written as lower diagonal matrix with main diagonal elements $C_{01}^{ij1}, \ldots, C_{k1}^{ij1}, \ldots, C_{k_m1}^{ij1}$ and lower diagonal elements $C_{12}^{ij1}, \ldots, C_{k2}^{ij1}, \ldots, C_{k_m2}^{ij1}$.

   - C_{k1}^{ij1} represents reduction by one in the number of ongoing voice calls and is a diagonal matrix with diagonal elements C_{l1}^{ij1k1}. For $i = 1, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ $l = 0, 1, \ldots, J$.

$$C_{l1}^{ij1k1} = C_0^{ve} \oplus C_0^{vo1} \oplus C_0^{vo2} \oplus C_0^d \oplus V_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus W_r(l) \tag{7}$$

   - C_{k2}^{ij1} represents reduction by one in the number of ongoing calls due to service completion by one of the data calls when the system is having i calls ongoing with j video and k data calls. C_{k2}^{ij1} is also a diagonal matrix represented as

$$C_{k2}^{ij1} = diag(C_{01}^{ij1k2}, \ldots, C_{l1}^{ij1k2}, \ldots, C_{J1}^{ij1k2})$$

   For $i = 1, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 1, \ldots, k_m$ $l = 0, 1, \ldots, J$

$$C_{l1}^{ij1k2} = C_0^{ve} \oplus C_0^{vo1} \oplus C_0^{vo2} \oplus C_0^d \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus V_d(k) \oplus W_r(l) \tag{8}$$

2. The lower diagonal elements C^i_{j2} of the matrix Q_{i2} is a diagonal block matrix that represents service completion by one of the number of ongoing video calls. The diagonal elements of this matrix C^i_{j2} are given by C^{ij2}_{k1} for $k = 0, 1, \ldots, k_m$. Each of C^{ij2}_{k1} is $diag(C^{ij2k1}_{01}, \ldots, C^{ij2k1}_{l1}, \ldots, C^{ij2k1}_{J1})$. For $i = 1, \ldots, M$, $j = 1, \ldots, j_m$, $k = 0, \ldots, k_m$ $l = 0, 1, \ldots, J$

$$C^{ij2k1}_{l1} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus V_{vo}(j) \oplus W_d(k) \oplus W_r(l) \qquad (9)$$

Lastly, we consider the diagonal elements $Q_{i1}, i = 0, 1, \ldots, M$, of the generator matrix Q. Q_{i1} corresponds to no change in the number of calls receiving service at the same time the number of calls in the orbit may increase when an arriving calls finds insufficient channels and joins the orbit or decrease due to cell residence time of the call gets over.

1. The diagonal elements of Q_{i1} are B^i_j that represents no change when i calls are ongoing in the system out of which j are video calls. It is possible through no service completion by any of the ongoing calls as well as no fresh arrival of calls or through successful retrial of the calls from the orbit. $B^i_j = diag(B^{ij}_0, \ldots, B^{ij}_k, \ldots, B^{ij}_{k_m})$. The matrix B^{ij}_k represents no change in the calls receiving service when k are data calls. Each of B^{ij}_k is a tridiagonal matrix.

   - The main diagonal element B^{ijk}_{l1} represents no change in the calls in service as well as no change in the number of calls in the orbit, that is, no call joins the orbit or departs the orbit. For $i = 0, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, \ldots, k_m$ $l = 0, 1, \ldots, J$

   $$B^{ijk}_{l1} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus \widehat{W}_r(l) \qquad (10)$$

   - The lower diagonal elements B^{ijk}_{l2} represents a call departure from the orbit and no change in the service domain of the system. For $i = 0, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, \ldots, k_m$ $l = 1, \ldots, J$

   $$B^{ijk}_{l2} = C^{ve}_0 \oplus C^{vo1}_0 \oplus C^{vo2}_0 \oplus C^d_0 \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus V_r(l,1) \qquad (11)$$

   - The upper diagonal elements B^{ijk}_{l1} represents no change in the number of calls receiving service but an increase in the number of calls in the orbit. New voice, handoff video2, new video2, handoff data and new data calls when on arriving do not find free channels(according to the CAC scheme) joins the orbit. For $i = 0, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, \ldots, k_m$ $l = 0, 1, \ldots, J-1$,

   $$B^{ijk}_{l1} = C^{vo1}_0 \oplus n^{i,j,k} \oplus W_{ve}(i-j-k) \oplus W_{vo}(j) \oplus W_d(k) \oplus Wr(l,1) \qquad (12)$$

   where

   $$n^{i,j,k} = \begin{cases} C^{nd}_0 \oplus C^{ve}_0 \oplus C^{vo2}_0, & M_1 \le i+3j < M_2 \\ (C^d_{hd} + C^d_{nd}) \oplus C^{ve}_0 \oplus C^{vo2}_0, & M_2 \le i+3j < M_3 \\ C^{vo2}_{nvo2} \oplus (C^d_{hd} + C^d_{nd}) \oplus C^{ve}_0, & M_3 \le i+3j < M_4 \\ (C^{vo2}_{nvo2} + C^{vo2}_{hvo2}) \oplus (C^d_{hd} + C^d_{nd}) \oplus C^{ve}_0, & M_4 \le i+3j < M_7 \\ C^{ve}_{nv} \oplus (C^{vo2}_{nvo2} + C^{vo2}_{hvo2}) \oplus (C^d_{hd} + C^d_{nd}) \oplus C^{ve}_0, & M_7 \le i+3j < M \\ 0, & \text{otherwise} \end{cases}$$

5 System Performance Measures

We will first consider the stationary probability of the Markov chains and then list some important measures along with their formulas. These measures are used to bring out the quantitative behavior of the queueing models under study.

Let $\mathbf{x}$ partitioned as $\mathbf{x} = \mathbf{x}(i,j,k,l)$ with $0 \leq i \leq M, \quad 0 \leq j \leq j_m, \quad 0 \leq k \leq k_m, \quad 0 \leq l \leq J$ where j_m, k_m is as defined in the previous section, denote the steady state probability vector for the generator matrix Q. That is, $\mathbf{x}$ satisfies $\mathbf{xQ} = 0, \quad \mathbf{xe} = 1$.

Each $\mathbf{x}(i,j,k,l)$ is a vector ordered in the lexicographical order based on the system state $(i, j, k, l, u_{ve}, u_{vo1}, u_{vo2}, u_d, \mathbf{S}^{ve}_{i-j-k}, \mathbf{S}^{vo}_{j}, \mathbf{S}^{d}_{k}, \mathbf{r}_l)$.

From these results we obtain the following performance measures:

1. *The probability mass function of the number of calls in orbit.* The probability that there are l calls in the orbit is given by

$$P_r(l) = \sum_{i=0}^{M}\sum_{j=0}^{j_m}\sum_{k=0}^{k_m} \mathbf{x}(i,j,k,l)\mathbf{e}, \qquad 0 \leq l \leq J$$

2. *The probability mass function of the number of voice calls in the system.* The probability that there are v voice calls is given by

$$P_{ve}(v) = \sum_{i=v}^{M}\sum_{j=0}^{j_m}\sum_{k=0}^{k_m}\sum_{l=0}^{J} \mathbf{x}(i,j,d,l)\mathbf{e}, \qquad \text{with } i+3j \leq M$$

$$\text{and } M-(j+k) = v \quad \text{where } 0 \leq v \leq M$$

3. *The probability mass function of the number of video calls in the system.* The probability that there are j video calls is given by

$$P_{vo}(j) = \sum_{i=j}^{M}\sum_{k=0}^{k_m}\sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}, \qquad 0 \leq j \leq j_m$$

4. *The probability mass function of the number of data calls in the system.* The probability that there are k data calls is given by

$$P_d(k) = \sum_{i=k}^{M}\sum_{j=i-k}^{j_m}\sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}, \qquad 0 \leq k \leq k_m$$

5. The dropping probabilities of handoff voice and video1 respectively are given as

$$\begin{aligned} P_{hv} &= \sum_{i}\sum_{j}\sum_{k=0}^{k_m}\sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}, \qquad i+3j \leq M, \text{and } 0 \leq i \leq M, \; 0 \leq j \leq j_m. \\ P_{hvo1} &= \sum_{i}\sum_{j}\sum_{k=0}^{k_m}\sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}, \qquad M_6 \leq i+3j \leq M, \text{ and } 0 \leq i \leq M, \; 0 \leq j \leq j_m. \end{aligned}$$

6. The blocking probabilities of new video1 calls is given by

$$P_{nvo1} = \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}, \quad M_6 \le i+3j \le M \quad \text{and } 0 \le i \le M,\ 0 \le j \le j_m.$$

7. The probability that new voice, handoff video2, handoff data and new data calls respectively has to join the orbit is given as follows for $0 \le i \le M,\ 0 \le j \le j_m$

$$\begin{aligned}
P_{nv} &= (1-P_r(J)) \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}. \ \ M_7 \le i+3j \le M. \\
P_{hvo2} &= (1-P_r(J)) \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}. \ \ M_4 \le i+3j \le M. \\
P_{nvo2} &= (1-P_r(J)) \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}. \ \ M_3 \le i+3j \le M. \\
P_{hd} &= (1-P_r(J)) \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}. \ \ M_2 \le i+3j \le M. \\
P_{nd} &= (1-P_r(J)) \sum_i \sum_j \sum_{k=0}^{k_m} \sum_{l=0}^{J} \mathbf{x}(i,j,k,l)\mathbf{e}. \ \ M_1 \le i+3j \le M.
\end{aligned}$$

Note that $(1-P_r(J))$ is the probability that the orbit is not full.

8. The Total Carried Traffic defined as the average number of channels occupied is given by

$$TCT = \sum_{i=0}^{M} \sum_{j=0}^{j_m} \sum_{k=0}^{k_m} \sum_{l=0}^{J} (i+3j)\mathbf{x}(i,j,k,l)\mathbf{e}$$

9. The average length of orbit is given by

$$L = \sum_{i=0}^{M} \sum_{j=0}^{j_m} \sum_{k=0}^{k_m} \sum_{l=1}^{J} l\mathbf{x}(i,j,k,l)\mathbf{e}$$

10. The average value of waiting time of calls in the retrial queue is given by

$$E(T) = L/[\lambda_{nv}(1-P_{nv}) + \lambda_{hvo2}(1-P_{hvo2}) + \lambda_{nvo2}(1-P_{nvo2}) + \lambda_{hd}(1-P_{hd}) + \lambda_{nd}(1-P_{nd})]$$

6 Particular Cases

In this section, we illustrate the usefulness of our retrial queueing system by describing a variety of queueing phenomenon which are obtained as particular cases of our model description.

6.1 Poisson arrivals, exponential service and retrial times, four type of calls

We consider that the arrival of handoff and new voice, video1, video 2 and data calls are modeled by a Poisson process with rates $\lambda_{hv}, \lambda_{nv}, \lambda_{hvo1},\ \lambda_{nvo1},\ \lambda_{hvo2}, \lambda_{nvo2},\ \lambda_{hd}$ and λ_{nd} respectively. We also assume that, for voice and data calls, actual service times are exponentially distributed with rates μ_{ve} and μ_d respectively, and for video1 and video2 calls, actual service times are exponentially distributed with rate μ_{vo}. The time to quit from the orbit is exponentially distributed with rate μ_{rq} and time for any call to get connection is also exponential distributed with rate μ_{rc}. Let $\{X(t), t \geq 0\}$ be a stochastic process in a given cell on the following state space:

$$\Omega = \{i, j, k, l;\ 0 \leq i \leq M;\ 0 \leq j \leq j_m;\ 0 \leq k \leq k_m;\ 0 \leq l \leq J\}$$

The memoryless property of the exponential inter arrival times of calls, channel holding times and retrial times in the orbit assures that the problem can be formulated as a continuous time Markov chain.

This stochastic process can be studied as a QBD process and the corresponding generator matrix Q is as given in the section 4. Now, we obtain the elements of matrix Q for this particular case.

First, we give the expression of $A_{l1}^{i,j,1,k,1}$ and $A_{l2}^{i,j,1,k,1}$ (refer equations (1) and (3)). When $i = 0$,

$$A_{l1}^{0,0,1,0,1} = \lambda_{VE}^{0,0}, \quad l = 0, \ldots, J; \quad A_{l2}^{0,0,1,0,1} = g^{0,0} * l * \mu_{rc}, \quad l = 1, \ldots, J;$$

for $i = 1, 2, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ and $l = 0, 1, \ldots, J$,

$$A_{l1}^{i,j,1,k,1} = \lambda_{VE}^{i,j}; \quad A_{l2}^{i,j,1,k,1} = g^{i,j} * l * \mu_{rc}$$

where

$$\lambda_{VE}^{i,j} = \begin{cases} \lambda_{hv} + \lambda_{nv}, & 0 \leq i + 3j < M_7 \\ \lambda_{hv}, & M_7 \leq i + 3j < M \\ 0, & \text{otherwise} \end{cases} \quad \text{and } g^{i,j} = \begin{cases} p_{nv}, & 0 \leq i + 3j < M_7 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, we have the expression of $A_{l1}^{i,j,1,k,0}$ and $A_{l2}^{i,j,1,k,0}$ (refer equations (2) and 4)). When $i = 0$,

$$A_{l1}^{0,0,1,0,0} = \lambda_{D}^{0,0}, \quad l = 0, \ldots, J; \quad A_{l2}^{0,0,1,0,0} = h^{0,0} * l * \mu_{rc}, \quad l = 1, \ldots, J;$$

for $i = 1, 2, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ and $l = 0, 1, \ldots, J$,

$$A_{l1}^{i,j,1,k,0} = \lambda_{D}^{i,j}; \quad A_{l2}^{i,j,1,k,0} = h^{i,j} * l * \mu_{rc}$$

where

$$\lambda_{D}^{i,j} = \begin{cases} \lambda_{hd} + \lambda_{nd}, & 0 \leq i + 3j < M_1 \\ \lambda_{hd}, & M_1 \leq i + 3j < M_2 \\ 0, & \text{otherwise} \end{cases} \quad \text{and } h^{i,j} = \begin{cases} p_{hd} + p_{nd}, & 0 \leq i + 3j < M_1 \\ p_{hd}, & M_1 \leq i + 3j < M_2 \\ 0, & \text{otherwise} \end{cases}$$

Now, we give the expression of $A_{l1}^{i,j,0,k,1}$ and $A_{l2}^{i,j,0,k,1}$ (refer equations (??) and (??)). When $i = 0$,

$$A_{l1}^{0,0,0,0,1} = \lambda_{VO}^{0,0}, \quad l = 0, \ldots, J; \quad A_{l2}^{0,0,0,0,1} = m^{0,0} * l * \mu_{rc}, \quad l = 1, \ldots, J;$$

for $i = 1, 2, \ldots, M$, $j = 0, 1, \ldots, j_m$, $k = 0, 1, \ldots, k_m$ and $l = 0, 1, \ldots, J$,

$$A_{l1}^{i,j,0,k,1} = \lambda_{VO}^{i,j}; \quad A_{l2}^{i,j,0,k,1} = m^{i,j} * l * \mu_{rc}$$

where

$$\lambda_{VO}^{i,j} = \begin{cases} \lambda_{hvo1} + \lambda_{nvo1} + \lambda_{hvo2} + \lambda_{nvo2}, & 0 \leq i+3j+4 < M_3 \\ \lambda_{hvo1} + \lambda_{nvo1} + \lambda_{hvo2}, & M_3 \leq i+3j+4 < M_4 \\ \lambda_{hvo1} + \lambda_{nvo1}, & M_4 \leq i+3j+4 < M_5 \\ \lambda_{hvo1}, & M_5 \leq i+3j+4 < M_6 \\ 0, & \text{otherwise} \end{cases}$$

$$m^{i,j} = \begin{cases} p_{hvo2} + p_{nvo2}, & 0 \leq i+3j+4 < M_3 \\ p_{hvo2}, & M_3 \leq i+3j+4 < M_4 \\ 0, & \text{otherwise} \end{cases}$$

Now, we consider the block elements of Q_{i2} (refer equations (7) and (9)). We have, for $i = 1, 2, \dots, M$, $j = 0, 1, \dots, j_m$, $k = 0, 1, \dots, k_m$ and $l = 0, 1, \dots, J$,

$$C_{l1}^{i,j,1,k,1} = (i-j-k) * \mu_{ve}; \ C_{l1}^{i,j,2,k,1} = j * \mu_{vo}; C_{l1}^{i,j,1,k,2} = k * \mu_d$$

Finally, we consider the blocks of Q_{i1} (refer equations (10) and (12)). When $i = 0$, we have

$$B_{l1}^{0,0,0} = -(\lambda_{VE}^{0,0} + \lambda_{VO}^{0,0} + \lambda_D^{0,0}), \ l = 0, 1, \dots, J; \quad B_{l2}^{0,0,0} = l * \mu_{rq}, \quad l = 1, \dots, J.$$

For $i = 1, 2, \dots, M$, $j = 0, 1, \dots, j_m$, $k = 0, 1, \dots, k_m$ and $l = 0, 1, \dots, J$,

$$\begin{aligned} B_{l1}^{i,j,k} &= -(\lambda_{VE}^{i,j} + \lambda_{VO}^{i,j} + \lambda_D^{i,j}) - (i-j-k) * \mu_{ve} - j * \mu_{ve} - k * \mu_d, \\ B_{l0}^{i,j,k} &= n^{i,j,k}; \quad B_{l2}^{i,j,k} = l * \mu_{rq} \end{aligned}$$

where $l = 0, 1, \dots, J-1$,

$$n^{i,j,k} = \begin{cases} \lambda_{nd}, & M_1 \leq i+3j < M_2 \\ \lambda_{hd} + \lambda_{nd}, & M_2 \leq i+3j < M_3 \\ \lambda_{nvo2} + \lambda_{hd} + \lambda_{nd}, & M_3 \leq i+3j < M_4 \\ \lambda_{hvo2} + \lambda_{nvo2} + \lambda_{hd} + \lambda_{nd}, & M_4 \leq i+3j < M_7 \\ \lambda_{nv} + \lambda_{hvo2} + \lambda_{nvo2} + \lambda_{hd} + \lambda_{nd}, & M_7 \leq i+3j < M \\ 0, & \text{otherwise} \end{cases}$$

Finally,

6.2 MAP arrivals, exponential service and retrial times, one type of call [31]

Our model description can also be used to describe the analytical framework of the system with one type of calls arriving according to Markov arrival process, exponential service and retrial times. Each call require only one channel for its connectivity. Below we discuss the modeling of such a system from our model description.

Let C be the irreducible infinitesimal generator matrix of a K-state Markov process. Let C_0 and C_N represents the transition corresponding to no arrivals and transitions corresponding to arrivals such that infinitesimal generator matrix corresponding to the underlying CTMC is $C = C_0 + C_N$. Let π be the stationary probability vector for C which can be calculated using $\pi C = 0, \ \pi e = 1$. The call arrival rate is given by $\lambda = \pi C_N e$. The calls in the orbit keep on retrying for free channels after a random interval that is exponentially distributed with rate θ. Finite number of calls are allowed to go to orbit as it is of fixed size J. The channel holding times of the calls is exponentially distributed with parameter μ.

The state space is given by $\{(i,j,u), 0 \le i \le M, 0 \le j \le J, 1 \le u \le K\}$, where i is the number of busy channels which also represent the number of ongoing calls, j is the number of calls in the orbit, and u give the transition phase of the underlying Markov chain of the calls arrival process. Note that even if there are calls waiting in the orbit and the channels are free, only the calls that retry from the orbit and fresh arriving calls get connected.

We see that, in this case the matrix Q get simplified by two levels. The matrix Q is given as

$$Q = \begin{pmatrix} Q_{01} & Q_{00} & & & & \\ Q_{12} & Q_{11} & Q_{10} & & & \\ & \ddots & \ddots & \ddots & & \\ & & Q_{i2} & Q_{i1} & Q_{i0} & \\ & & & \ddots & \ddots & \ddots \\ & & & & Q_{M2} & Q_{M1} \end{pmatrix}$$

The blocks Q_{i0} represents an increase in the number of calls receiving service, Q_{i1} represents no change in the number of calls in the service and Q_{i2} gives the reduction in the number of calls receiving service due to call completion when i calls in service for $i = 0, 1, \ldots, M$.

First we consider the blocks Q_{i0}. Increase in the number of calls can take place either due to fresh arrivals of calls or due to successful retrial by the calls from the orbit. It is written as a lower diagonal matrix.

$$Q_{i0} = \begin{pmatrix} A^i_{01} & & & & \\ A^i_{12} & A^i_{11} & & & \\ & \ddots & \ddots & & \\ & & A^i_{(J-1)2} & A^i_{(J-1)1} & \\ & & & A^i_{J2} & A^i_{J1} \end{pmatrix}$$

The matrices A^i_{j1} and A^i_{j2} corresponds to the increase in the number of calls in service due to fresh arrivals of calls and due to successful retrials of calls respectively. The expressions are given as:
$A^i_{j1} = C_N, i = 0, 1, \ldots, M-1, j = 0, 1, \ldots, J$
$A^i_{j2} = j\theta I, i = 0, 1, \ldots, M-1, j = 1, \ldots, J.$
Next,we consider the blocks Q_{i1}. No change in the number of calls is possible when no service completion takes place or no fresh arrival of calls occur or no successful retrials of calls from the orbit happens; at the same time the number of calls in orbit may vary. It is written as a tri-diagonal matrix with main diagonal elements $(B^i_{01}, B^i_{11}, \ldots, B^i_{j1}, \ldots, B^i_{J1})$, where B^i_{j1} represents no change in the number of calls of orbit also. The lower diagonal elements $(B^i_{12}, B^i_{22}, \ldots, B^i_{j2}, \ldots, B^i_{J2})$ represents no change in the calls receiving service but the calls in orbit decreases by one. This happens only if the call from the orbit is successful in occupying a free channel. The upper diagonal elements $(B^i_{01}, B^i_{10}, \ldots, B^i_{j0}, \ldots, B^i_{J-10})$ represents no change in the ongoing calls but an increase in the calls in the orbit occurs. This happens when $i = M$. The expressions are given as

$$B^i_{j1} = \begin{cases} -i\mu I - j\theta \otimes C_0, & i = 0, \ldots, M-1, \ j = 0, 1, \ldots, J \\ -i\mu I + j\theta I \otimes C_0, & i = M, \ j = 0, 1, \ldots, J-1. \\ -i\mu I + j\theta I \otimes (C_0 + C_N), & i = M, \ j = J \end{cases}$$

The matrix B^i_{j0} represents an increase in the number of calls in orbit. It is possible only when on arriving the calls do not find free channels and it joins the orbit.

$$B^i_{j0} = \begin{cases} 0 & 0 \le i \le M-1 \\ C_N, & i = M, 0 \le j \le J-1 \end{cases}$$

Decrement in the number of calls in the orbit is represented by B^i_{j2}. For $i = 0, 1, \ldots, M-1$, the transition when the calls in the orbit do not retry are considered, as retrial by calls from the orbit would be successful thereby increasing the number of calls in service. When $i = M$, the calls in the orbit make unsuccessful retrials. Therefore,

$$B^i_{j2} = \begin{cases} 0, & i = 0, 1, \ldots, M-1 \\ j\theta I, & i = M \end{cases}$$

Lastly, we consider the lower diagonal blocks Q_{i2}. It is a diagonal matrix for each i, the diagonal elements give by C^i_{j2} for each $j = 0, 1, \ldots, J$. C^i_{j2} represents reduction in the number of ongoing calls due to call completion when there are i calls in service and j calls in the orbit. The corresponding expression is given by $-j\theta I + i\mu I \otimes C_N$, for $j = 0, 1.\ldots, J$ and $i = 1, \ldots, M$.

6.3 Poisson arrivals, phase type service and exponential retrial times, one type of call

Our model can also be simplified to present the analytical model of the system consisting of one type of calls arriving in accordance with Poisson distribution with rate λ and having phase type service times. Also, when the channels are busy, the calls go to orbit from where they keep on retrying for the free channel. We assume that the calls in the orbit keep on retrying until connected. The retrial times are exponentially distributed with rate θ. The channel holding times of the calls is phase type distributed with representation (β, S) of dimension N. The exit rate vector is $s_0 = -Se$.

The state space is given by $\{(i, j, \mathbf{S}_i), 0 \le i \le M, 0 \le j \le J\}$, where i is the number of busy channels which also represent the number of ongoing calls, j is the number of calls in the orbit, $j = 0, 1, \ldots, J$, $\mathbf{S}_i = (s_1, s_2, \ldots, s_i), s_v = 1, 2, \ldots, N \ 1 \le v \le i$ represents the set of phases of i calls that are ongoing. Note that even if there are calls waiting in the orbit and the channels are free, only the calls that retry from the orbit and fresh arriving calls get connected.

Notations:

$$W(i) = S \oplus S \oplus S, \ldots, \oplus S, \ i \text{ times}$$

It represents no call completion by any of the i calls in service.

$$V(i) = \sum_{j=0}^{i-1} I(N, j) \otimes s_0 \otimes I(N, i-j-1)$$

represents the service completion by one of the i calls.

The infinitesimal generator matrix is same as given in case2. We give the corresponding elements below: The upper diagonal block elements of Q matrix are given as $A^i_{j1} = \lambda I, i = 0, 1, \ldots, M-1, j = 0, 1, \ldots, J$ $A^i_{j2} = j\theta I, i = 0, 1, \ldots, M-1, j = 1, \ldots, J$.
Similarly, the diagonal blocks elements are described as

$$B^i_{j1} = \begin{cases} -\lambda - j\theta, & i = 0, j = 0, 1, \ldots, J \\ S_1(i, j), & i = 1, \ldots, M-1, \ j = 0, 1, \ldots, J \\ W(i) + j\theta - \lambda I, & i = M, \ j = 0, 1, \ldots, J. \\ W(i) + J\theta + \lambda I, & i = M, \ j = J \end{cases}$$

where $S_1(i, j) = W(i) - j\theta I - \lambda I$. It represents no service completion by any of the i calls, no retrial attempts, and no call arrivals to the system.

$$B^i_{j0} = \begin{cases} 0 & 0 \le i \le M-1 \\ W(i) + \lambda I + j\theta I, & i = M, 0 \le j \le J-1 \end{cases}$$

$$B^i_{j2} = \begin{cases} 0, & i = 0, 1, \ldots, M-1 \\ j\theta I, & i = M \end{cases}$$

The lower diagonal block of the Q matrix represents the reduction in the number of ongoing calls due to service completion, which is a diagonal block matrix for each i and j . The corresponding expression is given by $C^i_{j2} = -j\theta I - \lambda I \oplus V(i)$, for $j = 0, 1 \ldots, J$ and $i = 1, \ldots, M$.

Thus, we see that our model can be considered as a generalization of many retrial queueing systems. Also further assumptions can be made to get the analytic model for queueing systems without retrials.

7 Conclusions and Future Work

In this paper, we have presented a performance model to obtain call blocking and dropping probabilities and other important measures for the wireless networks supporting high speed data rate. In this model, we have considered the Markov arrival process of voice, data and video traffic and phase type distributions of call completion times and retrial times. A new CAC scheme with fixed resource reservation based on the priority of traffic is proposed. We believe, that, based on our CAC scheme and the performance analysis, the service provider will be able to enhance wireless networks designs and protocols. Also, they can guarantee QoS requirements to vendors and give new directions in adding further features in future wireless networks. Our work can be used for more accurate dimensioning of wireless networks with realistic traffic. We propose to work on the problem of determining optimal number of guard channels that should be reserved for voice, data and video calls so as to maximize the service providers revenue at the same time meeting the desired QoS. We also propose to analyze our CAC scheme with dynamic resource reservation for different type of traffics.

References

[1] G.Haring, R.Marie, R. Puigjaner and K.S.Trivedi, "Loss Formulas and Their Applications to Optimization for Cellular Networks", *IEEE Trans. on Vehicular Technology*, 50(3)(2001), 664-673.

[2] D.P.Gaver, P.A.Jacobs and G.Latouche, "Finite Birth and Death Models in Randomly Changing Environments", *Advances in Applied Probability*,16(1984), 715-731.

[3] S.Dharmaraja, K.S.Trivedi and D.Logothetis, "Performance Modelling of Wireless Networks With Generally Distributed Handoff Interarrival Times", *Computer Communications*, 26(2003), 1747-1755.

[4] Aruna Jayasuriya, David Green and John Asenstorfer, "Modelling Service Time Distributions in Cellular Networks Using Phase Type Service Distributions", *IEEE International Conference, Helsinki, Finland*, 2(2001), 440-444.

[5] Bong Due Choi, Sung Ho Choi, BaraKim and DanKen Sung, "Analysis of Priority Queueing Systems Based on Thresholds and its Applications to Signalling System No.7 with Congestion Control", *Computer Networks*, 32(2)(2000), 149-170.

[6] Sung Ho Choi and K. Sohraby, "Analysis of a Mobile Cellular System with Handoff Priority and Hystereis Control", *INFOCOM 2000*, 2000, 217-224.

[7] Wenbin Gong and Zhongmin Gan, "Performance Analysis for Call Admission Control on Voice/Data Integration Mobile Communication System", *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communication*,(2003), 50-54.

[8] Bin Li, Lizhong Li, Krishna M. Sivalingam and Xi-Ren Cao, "Call Admission Control for Voice/Data Integrated Cellular Networks: Performance Analysis and Comparative Study", *IEEE Journal on Selected Areas in Communication*, 22(4)(2004), 706-718.

[9] A. S.ALfa and Wei Li, "PCS Networks with Correlated Arrival Process and Retrial Phenomenon",*IEEE Trans. on Wireless Communications*, 1(4)(2002), 630-637.

[10] S.Dharmaraja and Attahiru Sule Alfa, "Performance Analysis of 3G Wireless Networks Supporting Voice, Data and Video Traffic", *Proc. of International Conference on Wireless Networks (ICWN 2003)*, Las Vegas, U.S.A. 2003, 495-498.

[11] M.F.Neuts, *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach*, John Hopkins University Press, Baltimore, 1981.

[12] Srinivas R. Chakraborthy and Alexander N. Dudin, "A Multi Server Retrial Queue with BMAP arrivals and Group Services", *Queueing Systems*, 42(2002), 5-31.

[13] Mohamed Hossam Ahmed, "Call Admission Control in Wireless Networks: A Comprehensive Survey", *IEEE Comm. Surveys and Tutorials*, 7(1), Jan. 2005.

[14] Mainak Chatterjee, S.K. Das and Giridhar D. Mandyam, "Performance Evaluation of Voice-Data Integration for Wireless Data Networking", *Int. Conf. on Networking, (ICN)*, LNCS, July 2001, 157-166.

[15] Vladimir V. Anisimov and Jesus R. Artalejo, "Analysis of Markov Multiserver Retrial Queues with Negative Arrivals", *Queueing Systems*, 39(2001), 157-182.

[16] M.F. Neuts and B.M. Rao, "Numerical Investigation of a Multi Server Retrial Model", *Queueing Systems*, 7(1990), 169-190.

[17] G. Falin, "A Survey of Retrial Queues",*Queueing Systems*, 7(1990), 127-168.

[18] Qi-Ming HE, Hui Li, Yiqiang and Q. Zhao, "Ergodicity of the BMAP/Ph/s/s+K Retrial Queue with PH-Retrial Times", *Queueing Systems*, 35(1-4)(2000), 323-347.

[19] K.Mitchell, K.S. Sohraby, A.van de Licfvoort and J.P Lace,"Approximation Models of Wireless Cellular Networks Using Moment Matching." *IEEE Journal on Selected Areas of Communication*, 19(11)2001.

[20] Alexander K;emm, Christoph Lindemann and Marco Lohmann, "Traffic Modeling of IP Networks Using the Batch Markovian Arrival Process", *Computer Performance Evaluation/TOOLS* 2002, 92-110.

[21] Ioannis D. Moscholios, Michael D. Logothetis and George K. Kokkirekes, "Call Burst Blocking of On-Off Traffic Sources with Retrials Under the Complete Sharing Policy", *Performance Evaluation*, 59(2005), 279-312.

[22] N.G. Bean and D.A. Green, "When is a MAP Poisson?" *Mathematical and Computer Modelling*, 31(2000), 31-46.

[23] T.S. Randhawa and R.H.S. Hardy, "Performance Evaluation of Priority based Call Admission Control in Multi Service DS-CDMA Cellular Networks Supporting Video Traffic" *in Proc. IEEE Inter. Conf on Communications*, 15(2002), 3096-3105.

[24] I.V. Martin F., Juan J. Alins, Monica Aguilar-Igartua and Jorge Mata Diaz,"Modeling an Adaptive-Rate Video Streaming Service Using Markov Reward Models", *IEEE Inter. Conf. on QoS in Heterogeneous Wired/Wireless Networks*, 2004.

[25] K. S. Trivedi, S.Dharmaraja and X.Ma, "Analytic Modeling of Handoffs in Wireless Cellular Networks", *Information Sciences*, 148(2002), 155-166.

[26] A. S. Alfa and W. Li, "A Homogeneous PCS Network with Markov Call Arrival Process and Phase Type Cell Residence Time", *Wireless Networks*, 8(2002), 597-605.

[27] B.D. Choi and Y. Chang, "$MAP_1, MAP_2/M/c$ Retrial Queue with the Retrial Group of Finite Capacity and Geometric Loss" , *Mathematical and Computer Modeling*, 30(1999), 99-113.

[28] S. R. Chakraborthy and Alexander N. Dudin, "A Multiserver Retrial Queue with BMAP Arrival and Group Services" , *Queueing Systems*, 42(2002), 5-31.

[29] Fabien Houeto and Samuel Pierre, "Quality of Service and Performance Issues in Multiservice Networks subject to Voice and Video Traffics", *Computer Communications*, 28(2005), 393-404.

[30] Haw-Yun Shin and Jean-Lien C. Wu, "The Study of Dynamic Multi-Channel Scheme with Channel de-allocation in Wireless Networks", *Computer Networks*, 45(2004), 463-482.

[31] A. S. Alfa and K.P. Sapna Isotupa, "An $M/Ph/k$ Retrial Queue with Finite Number of Sources", *Computers and Operations Research*, 31(2004), 1455-1464.