Multi-User Queuing Analysis Considering AMC for Wireless VoIP Services

Howon Lee∗ and Dong-Ho Cho†
School of Electrical Engineering and Computer Science, KAIST
Email: *hwlee@comis.kaist.ac.kr, †dhcho@ee.kaist.ac.kr

Abstract—We propose a new framework to analyze performance considering finite-length queuing and adaptive modulation and coding (AMC) for multi-user VoIP services in wireless communication systems. We formulate an uplink VoIP system as two-dimensional discrete-time Markov chain (DTMC) based on a Markov modulated Poisson process (MMPP) traffic model and MCS (modulation and coding scheme)-level set transition reflecting users’ channel variations. We extend transition modeling of MCS-level for single-user to transition modeling of MCS-level set for multi-user. Throughout our DTMC formulation, we present various performance metrics, such as average queue-length, the average number of arrived and serviced packets, and packet dropping probability.

I. INTRODUCTION

Multi-user VoIP (Voice over IP) traffic can be generally modeled as a two-state Markov modulated Poisson process [1] [2]. In [3] and [4], the authors proposed a discrete-time Markov chain (DTMC) framework for uplink VoIP services based on the MMPP traffic model. Through the DTMC based on MMPP, they analyzed and demonstrated various performance results, such as average queue length, average throughput, and average packet dropping probability. These work only considered queuing effect to show the multi-user VoIP performance in the wireless environment. They did not take into account AMC together in the queuing analysis.

The adaptive modulation and coding (AMC) is one of the basic and key link adaptation techniques which exploit the time/frequency variations of a wireless channel by changing the basic transmission parameters [5] [6]. So, a lot of researchers are trying to adopt the AMC to the currently developing wireless communication systems, and most systems have already utilized the AMC technique, such as IEEE 802.11n and IEEE 802.16e-2005. In practical wireless systems, the queue in the data-link layer is not always full, because the data-traffic is not always fully generated from the application layer. The queue condition of the data-link layer can be different according to the property of the utilized data-traffic. But, most published papers considering the AMC assume that the queue status is always full. Also, most queuing works do not consider the AMC together in the formulation of a system [3] [4]. Actually, the simultaneous consideration of the queuing and AMC is very important in terms of the practical wireless communication system. In [7], although Qingwen Liu et al. proposed the framework for queuing with the AMC over wireless links, the authors analyzed the joint effects of the finite-length queuing and AMC just for the single-user case. They presented a general analytical procedure for their framework, and showed various results for several channel-condition and system-design parameters, such as dropping probability versus target packet error rate, average spectral efficiency versus average signal-to-noise ratio (SNR), and so on. In this paper, we introduce cross-layer approach considering multi-user uplink VoIP traffic and AMC.

The remainder of this paper is organized as follows: In Section II, we introduce VoIP traffic model based on MMPP, and MCS-level transition modeling in the case of a single-user. In this section, we extend the MCS-level transition modeling for the single-user to MCS-level set transition modeling for multi-user. In Section III, we propose a two-dimensional DTMC framework considering multi-user VoIP queuing and AMC. In Section IV, we demonstrate various numerical results. Finally, Section V concludes this paper.

II. VoIP TRAFFIC AND MCS-LEVEL SET TRANSITION MODELING

A. Assumptions

Channel quality information (CQI) does not vary during a MAC frame. It just varies from frame to frame. Also, Perfect CQI is available in the base station (BS). Transitions of MCS-levels just happen between the adjacent MCS levels. But, in case of MCS-level set transitions, relatively more transitions are possible, because the transition of each user can be different. It can make lots of MCS-level combinations. We assume that the users have the same MCS-level transition matrix. However, the transition of each user is independent. Moreover, there are no packet retransmissions. Unreceived and destroyed packets are not sent again.

B. VoIP Traffic Modeling

In general, we can formulate the VoIP traffic of a single user as a simple on-off model [8]. The probability that the status of the users is inactive (= off) in the simple on-off model can be obtained by \( P_{off} = \frac{1/\alpha}{1/\alpha + 1/\beta} \). Here, 1/\( \alpha \) and 1/\( \beta \) are the mean values of the on and off periods, which are distributed exponentially. Furthermore, all the traffic generated by the VoIP users in the cell can be modeled as a two-state MMPP model [1] [4], as shown in Fig. 1. The MMPP is a stochastic process in which the intensity of a Poisson process is defined by the states of a Markov chain. This MMPP model is very suitable for formulating the multi-user VoIP traffic, because the MMPP captures the interframe dependency.
between consecutive frames. The transition rate matrix (R) and the Poisson arrival rate matrix of the MMPP (Λ) can be expressed as

\[
R = \begin{bmatrix} -r_1 & r_1 \\ r_2 & -r_2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.
\]

In order to utilize the MMPP model, we should match the MMPP parameters \((r_1, r_2, \lambda_1, \lambda_2)\) in equation (1) with the parameters of the simple on-off model (\(\alpha \) and \(\beta\)). We here adopt the index of dispersion for counts (IDC) matching technique, because it yields adequate results for the matching of parameters and has appropriate computation complexity compared with other matching techniques [3]. Then, \(r_1, r_2, \lambda_1, \) and \(\lambda_2\) in the equation (1) can be calculated by

\[
r_1 = \frac{2(\lambda_2 - \lambda_1)}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_1)^2}, \quad r_2 = 2\frac{(\lambda_2 - \lambda_1)}{(\lambda_2 - \lambda_1)^2}, \quad \lambda_1 = \frac{A}{\sum_{i=1}^{N} \sum_{n=1}^{\infty} \frac{1}{\pi_i}}, \quad \lambda_2 = A \cdot \frac{1}{\sum_{i=1}^{N} \sum_{n=1}^{\infty} \frac{1}{\pi_i}}.
\]

Here, \(N\) is the total number of VoIP users in the system and \(A\) is the emission rate in the on-state (\(A = 1/T_{\text{basic}}\)). \(T_{\text{basic}}\) is a frame duration of voice codec, and the average arrival rate is \(\lambda_{\text{avg}} = N \times \frac{1}{p_{\text{on}}}.\) The probability that the status of the users is active can be obtained by \(p_{\text{on}} = \frac{\alpha-1}{\alpha+\beta}.\) The average number of active users is \(N_{\text{act,avg}} = \left[N \times p_{\text{on}}\right]\) and the steady-state probability of a one-dimensional Markov chain when considering \(N\) independent simple on-off voice users can be calculated by

\[
\pi_i = \left(\frac{N}{i}\right) p_{\text{on}}^i (1-p_{\text{on}})^{N-i}, \quad \text{IDC}(\infty) = 1 + \frac{2(\lambda_2 - \lambda_1)^2 r_2 r_5}{(r_1 + r_2)^2 (\lambda_2 - \lambda_2 r_2)}.
\]

C. MCS-Level Set Transition Modeling

1) MCS-level transition matrix for single-user: In the MCS-level transition matrix for the single-user, the state transitions just happen between the adjacent states. In [9], J. Razavilar et al. represented the state transition probabilities in case of the single-user \((P_{s, m \rightarrow m+1}, P_{s, m \rightarrow m-1}, P_{s, m})\) as follows.

\[
P_{s, m \rightarrow m+1} = \frac{N_{\text{act,avg}} T_{f}}{Pr(m)}, \quad \text{if } m = 1, \ldots, M - 1.
\]

\[
P_{s, m \rightarrow m-1} = \frac{N_{\text{act,avg}} T_{f}}{Pr(m)}, \quad \text{if } m = 2, \ldots, M.
\]

\[
P_{s, m} = 0, \quad \text{if } |m-n| \geq 2.
\]

Here, \(T_f\) is the duration of a MAC frame which is a basic transmission unit from the perspective of the BS in the MAC layer, and \(M\) is the total number of MCS-levels. \(N_{\text{m}}\) is the MCS-level crossing rate when the MCS-level of the user is \(m, \) and \(Pr(m)\) is the probability that the MCS-level of the user is \(m.\) From equation (2), the transition probability that the user will stay at the same MCS level \((P_{s, m})\) can be calculated as follows [10]

\[
P_{s, m} = \begin{cases} 1 - P_{s, m+1}, & \text{if } m = 1 \\ 1 - P_{s, m-1} - P_{s, m+1}, & \text{if } 1 < m < M \\ 1 - P_{s, M-1}, & \text{if } m = M.
\]

Thus, through equations (2) and (3), we can obtain the MCS-level transition matrix of single-user.

2) MCS-level transition rate for multi-user: When the MCS-level of the user is \(m:\) MCS-level crossing rate (either downward or upward) when the MCS-level of the user is \(m (N_m)\) can be described as [11]

\[
N_m = \frac{\sqrt{2 \pi m f_{\text{m}} \gamma_m}}{\gamma_{\text{avg}}} \Gamma\left(\frac{m f_{\text{m}} \gamma_m}{\gamma_{\text{avg}}} \right) \times \exp\left(-\frac{m f_{\text{m}} \gamma_m}{\gamma_{\text{avg}}} \right)
\]

where \(m f_{\text{m}}\) is a Nakagami fading parameter in the general Nakagami-m fading model [12].\(\gamma_m\) is a minimum SNR value which can support the MCS-level \(m,\) \(\gamma_{\text{avg}}\) is an average SNR value, and \(f_{\text{m}}\) denotes the mobility-induced Doppler spread. Given that the user’s SNR is \(\gamma_m,\) if we utilize the MCS-level \(m,\) we can barely satisfy the target PER \((P_{\text{th}})\) in the presence of additive white Gaussian noise (AWGN). \(\gamma_m\) can be obtained as \(\gamma_m = \frac{1}{g_{\text{m}} \Gamma\left(\frac{m f_{\text{m}} \gamma_m}{\gamma_{\text{avg}}} \right)}\) [13, eq. 6]. Here, \(g_{\text{m}}\) and \(\Gamma\) are fitting parameters, and Q. Liu et al. verified the accuracy of this PER approximation in [13]. Also, since Nakagami-m fading well represents urban multipath fading, we choose this fading model for performance analysis in this paper. This fading model could be better or more severe than Rayleigh fading by adjusting the parameter \(m f_{\text{m}}.\) Especially, if \(m f_{\text{m}} = 1,\) Nakagami-m fading is the same as Rayleigh fading. \(\Gamma(m f_{\text{m}})\) is a Gamma function, \(\Gamma(m f_{\text{m}}) = \int_{0}^{\infty} t^{m f_{\text{m}}-1} \exp(-t) \, dt\).

Thus, the probability that the MCS-level of the user is \(m:\) In this paper, to describe the received SNR \((\gamma)\) statistically, we utilize the Nakagami-m fading model [12]. In this fading model, the received SNR \((\gamma)\) can be represented as a random variable with a Gamma probability density function \((p_{\gamma}(\gamma))\) as follows:

\[
p_{\gamma}(\gamma) = \frac{1}{\Gamma(m f_{\text{m}})} \frac{m f_{\text{m}} \gamma_{\text{avg}}^{m f_{\text{m}}-1}}{\gamma_{\text{avg}}^m} \exp\left(-\frac{m f_{\text{m}} \gamma_{\text{avg}}}{\gamma_{\text{avg}}} \right)
\]

By using this equation, we can obtain the probability that the MCS-level of the user is \(m (Pr(m))\) [14].

\[
Pr(m) = \frac{\int_{0}^{\gamma_{\text{avg}}} p_{\gamma}(\gamma) d\gamma}{\Gamma(m f_{\text{m}})} = \frac{\frac{\gamma_{\text{avg}}}{\gamma_{\text{avg}}} - \Gamma(m f_{\text{m}}, m f_{\text{m}} \gamma_{\text{avg}})}{\Gamma(m f_{\text{m}})}
\]

Here, \(\Gamma(m f_{\text{m}}, x)\) is the complementary incomplete Gamma function, and \(\Gamma(m f_{\text{m}}, x) = \int_{x}^{\infty} t^{m f_{\text{m}}-1} \exp(-t) \, dt\)

2) MCS-level set transition matrix for multi-user: Given that the total number of users is \(N\) and the total number of MCS-levels is \(M,\) we can calculate the number of MCS-level sets (MCS-level combinations, \(K_{N,M}\))

\[
K_{N,M} = \sum_{m_N=1}^{M} \sum_{m_{N-1}=1}^{m_N} \sum_{m_{N-2}=1}^{m_{N-1}} \cdots \sum_{m_2=1}^{m_3} \sum_{m_1=1}^{m_2} 1.
\]
Thus, the size of MCS-level set transition matrix for multi-user becomes \( K_{N,M} \) by \( K_{N,M} \).

Based on the MCS-level transition probabilities of single-user, we can obtain elements of the MCS-level set transition matrix \((P_{ij}^M)\). First of all, we define the notation \( N^m_m \) and \( N^m_m \) which are the numbers of users whose MCS-levels are \( m \) in a current MAC frame and a next MAC frame, respectively. \( N^m_m \) and \( N^m_m \) can be described as follows.

\[
N^m_m = \begin{cases} 
N^m_{m,m+1} + N^m_{m,m}, & \text{if } m = 1 \\
N^m_{m,m+1} + N^m_{m,m-1} + N^m_{m,m}, & \text{if } 1 < m < M \\
N^m_{m,m-1} + N^m_{m,m}, & \text{if } m = M
\end{cases}
\]

where \( N^m_{m,m} \) means the number of users whose MCS-levels vary from \( m \) in the current frame to \( m' \) in the next frame. Also, from this equation, we can easily calculate \( N^m_m \).

\[
N^m_m = \begin{cases} 
N^m_{m,m} + N^m_{m+1,m}, & \text{if } m = 1 \\
N^m_{m-1,m} + N^m_{m,m} + N^m_{m+1,m}, & \text{if } 1 < m < M \\
N^m_{m-1,m} + N^m_{m,m}, & \text{if } m = M
\end{cases}
\]

Through equations (7) and (8), \( P_{ij}^M \) can be obtained as

\[
P_{ij}^M = \sum_{\psi \in \Psi_{S_i \rightarrow S_j}} \prod_{m=1}^{M} p^m_{\psi_m}.
\]

There are usually several transition-combinations for the \( S_i \rightarrow S_j \) transition, and \( \Psi_{S_i \rightarrow S_j} \) denotes the transition-combination set for the \( S_i \rightarrow S_j \) transition. Here, \( S_i (S_j) \) represents a MCS-level set in the current (next) frame. Then, the numbers of elements of \( S_i \) and \( S_j \) are \( n(S_i) = \sum_{m=1}^{M} N^m_m \) and \( n(S_j) = \sum_{m=1}^{M} N^m_m \), respectively. In equation (9), \( p^m_{\psi_m} \) can be calculated as equation (10). Consequently, from equations (9) and (10), we can obtain the MCS-level set transition matrix for multi-user.

### III. SYSTEM MODELING

We can formulate the multi-user queuing model for uplink VoIP services considering AMC as a two-dimensional DTMC, as shown in Fig. 2. In other words, we can analyze the behavior of queuing packets in a virtual uplink queue of the BS with this two-dimensional DTMC model. The transition matrix \((P)\) of our DTMC model can be defined as

\[
P = \begin{bmatrix}
A_{0,0} & A_{0,1} & \cdots & A_{0,L_{\max}} \\
A_{1,0} & A_{1,1} & \cdots & A_{1,L_{\max}} \\
\vdots & \vdots & \ddots & \vdots \\
A_{L_{\max},0} & A_{L_{\max},1} & \cdots & A_{L_{\max},L_{\max}}
\end{bmatrix}
\]

where \( L_{\max} \) is maximum queue length and the submatrix \((A_{i,j})\) is expressed as

\[
A_{i,j} = \begin{bmatrix}
B_{i,(S_1,0),(S_1,0)} & \cdots & B_{i,(S_1,0),(S_{K,M})} \\
B_{i,(S_2,0),(S_2,0)} & \cdots & B_{i,(S_2,0),(S_{K,M})} \\
\vdots & \vdots & \ddots & \vdots \\
B_{i,(S_{K,M},0),(S_{K,M},0)} & \cdots & B_{i,(S_{K,M},0),(S_{K,M})}
\end{bmatrix}
\]

In equation (12), \( A_{i,j} \) represents the variation of the number of queuing packets. That is, if \( i \) packets in the virtual uplink queue in the current frame, and it will be \( j \) packets in the next frame. In this submatrix \((A_{i,j})\), each element \((B_{i,(S_m,0),(S_m)})\) indicates the transitions between MCS-level sets \((S_m \rightarrow S_n)\) when the number of queuing packets is changed from \( i \) to \( j \). Also, \( B_{i,(S_m,0),(S_m)} \) is a 2-by-2 matrix, because our MMPP model has two phases, such as underloading and overloading. \( B_{i,(S_m,0),(S_m)} \) can be calculated in equation (13). In this equation, \( B_{\text{under-max}} \) is the number of uplink resources that are required to transmit a VoIP packet that is generated in the on-state when we use the highest MCS level, \( N(\psi_{on}) \) denotes the number of users whose packets are serviced in the current MAC frame, and \( \psi_{tot} \) is the whole user-set including all the users in the BS. Here, \( N(\psi_{on}) \) is just determined by the current MCS-level set \((S_m)\), and \( P_{m,m}^M \) denotes the MCS-level set transition probability.

In equation (13), the transition probability matrix \((U)\) and the diagonal probability matrix of the two-state MMPP model can be obtained as follows:

\[
U = (A - R)^{-1}A
\]

\[
D(m) = \begin{bmatrix}
\frac{(\lambda_k T_J)^m (-\lambda_k T_J)}{m} & 0 \\
0 & \frac{(\lambda_k T_J)^m (-\lambda_k T_J)}{m}
\end{bmatrix}
\]

Here, each element of \( D(m) \) is the probability that \( m \) VoIP packets arrive at the BS during the MAC frame duration \((T_J)\) in each phase of the two-state MMPP. In addition, \( P_{\text{Tr}}(N(\psi_{on}) = k, \psi_{tot} = S_m) = P_{m,m}^N(k) \) is the probability that the BS serves \( k \) VoIP packets from the virtual uplink queue, when the status of the MCS-level set is \( S_m \). This probability may vary according to the users’ MCS distribution in the MCS-level set, which can be obtained as

\[
P_{m,m}^N(k) = \frac{N(k)}{\sum_{k=0}^{L_{\max}} N(k)}
\]

Here, \( N(k) \) is the number of MCS combinations for the case that \( k \) VoIP packets are scheduled at a certain MAC frame from the virtual uplink queue, when the status of the current MCS-level set is \( S_m \).

Through the transition matrix \((P)\) in (11), we can obtain the steady-state probability matrix \((\pi_{M,A})\) for our two-dimensional DTMC model. This steady-state probability matrix can be calculated by solving equations \( \pi_{M,A} \cdot P = \pi_{M,A} \).
Thus, based on equation (21), we can define the VoIP capacity.

Here, $P$ is the threshold of the packet dropping probability that can arrive from each VoIP user during $\tau$. In the IEEE 802.16e-2005 system, we utilize an erTPS (extended realtime polling service) algorithm as a resource allocation method for a VoIP service [15].

A. Analytical Environment

The reference system is the IEEE 802.16e-2005 system, and we utilize an erTPS (extended realtime polling service) algorithm as a resource allocation method for a VoIP service [15]. In the IEEE 802.16e-2005 system, $T_f$ is 5 ms, and the total number of uplink resources is 144. We exploit half of the total uplink resources for VoIP services. So, $R_{tot}$ is 72 in this paper. Also, we assume the basic polling duration in the erTPS as $T_{basic}$. VoIP users do not use polling resources to transmit their VoIP packets. In order to obtain numerical results, we assume that the number of MCS-levels ($M$) is 2 (QPSK 1/2 and QPSK 3/4). Since the full size of the real-time protocol (RTP)/user datagram protocol (UDP)/internet protocol version 4 (IPv4) header is 40 bytes, we assume that the size of this header, when compressed by payload header suppression (PHS), is 16 bytes. The size of the general MAC header is 6 bytes, and $P_{lim}$ is 0.02. In addition, $f_d$ is 10 Hz, $L_{max}$ is 30, and $A_{max}$ is 200. We assume that $n_f$ is 1 to obtain numerical results.

We utilize the G.729B codec and assume that a voice activity detector operates perfectly. $T_{basic}$ is 10 ms, and this codec has two data rates (8 kbps and 0 kbps). The voice activity factor of G.729B is 0.4 in this paper.

B. VoIP Capacity

When the average SNR value increases, more VoIP users can utilize a high MCS-level. Then, the probabilities using the MCS-level sets which have better transmission efficiency could become larger. It makes resource utilization efficiency better compared with the case that the average SNR value is low. Following numerical results are obtained when the average SNR values are 3, 10, and 20 dB, and the target PER value ($P_0$) is 0.01.

Fig. 3 shows the average numbers of arrived packets and serviced packets, and the average uplink queue-length according to the increment of the number of users. Here, the average numbers of arrived packets are the same regardless of the average SNR value. In accordance with the increment of the number of VoIP users, the number of arrived packets grows linearly. On the other hand, the number of serviced packets increases against the increment of VoIP users up to a certain number. We here define this certain number as a threshold number. If the number of VoIP users becomes larger than the threshold number, the number of serviced packets does not increase against the increment of VoIP users up to a certain number.

\[ C_{VolP} = \arg \max N \in \{ N \mid P_{drop} \leq P_{lim} \} \]

Here, $P_{lim}$ is the threshold of the packet dropping probability for VoIP services.

IV. Numerical Results

A. Analytical Environment

The reference system is the IEEE 802.16e-2005 system, and we utilize an erTPS (extended realtime polling service) algorithm as a resource allocation method for a VoIP service [15]. In the IEEE 802.16e-2005 system, $T_f$ is 5 ms, and the total number of uplink resources is 144. We exploit half of the total uplink resources for VoIP services. So, $R_{tot}$ is 72 in this paper. Also, we assume the basic polling duration in the erTPS as $T_{basic}$. VoIP users do not use polling resources to transmit their VoIP packets. In order to obtain numerical results, we assume that the number of MCS-levels ($M$) is 2 (QPSK 1/2 and QPSK 3/4). Since the full size of the real-time protocol (RTP)/user datagram protocol (UDP)/internet protocol version 4 (IPv4) header is 40 bytes, we assume that the size of this header, when compressed by payload header suppression (PHS), is 16 bytes. The size of the general MAC header is 6 bytes, and $P_{lim}$ is 0.02. In addition, $f_d$ is 10 Hz, $L_{max}$ is 30, and $A_{max}$ is 200. We assume that $n_f$ is 1 to obtain numerical results.

We utilize the G.729B codec and assume that a voice activity detector operates perfectly. $T_{basic}$ is 10 ms, and this codec has two data rates (8 kbps and 0 kbps). The voice activity factor of G.729B is 0.4 in this paper.

\[ P_{drop} = 1 - N_{avg}(\psi_{on})/\rho \]

Thus, based on equation (21), we can define the VoIP capacity.

\[ C_{VolP} = \arg \max N \in \{ N \mid P_{drop} \leq P_{lim} \} \]

Here, $P_{lim}$ is the threshold of the packet dropping probability for VoIP services.
\[ N_{\text{avg}}(\psi_{\text{snr}}) = \sum_{i=0}^{\text{min}(i,j)} \sum_{j=0}^{S_i} \sum_{m=1}^{N_{\text{N,M}}} \min(i, j) \cdot \pi(j) \cdot \Pr(N(\psi_{\text{snr}}) = i, \psi_{\text{snr}} = S_m) \cdot \pi_M(m). \] (20)

V. CONCLUSIONS

In this paper, we presented a two-dimensional DTMC model considering MMPP based multi-user VoIP traffic and AMC. We extended the MCS-level transition matrix for single-user to the MCS-level set transition matrix for multi-user. By using this MCS-level set transition matrix and MMPP traffic model, we designed the uplink VoIP system as a two-dimensional DTMC model. We demonstrated various results, such as the average number of serviced and arrived packet, queue-length, and packet dropping probability. Namely, we obtained the VoIP capacity guaranteeing the threshold of the packet dropping probability. From the numerical results, when the average SNR values are 3, 10, and 20 dB, the VoIP capacities are 40, 50, and 60, respectively.

REFERENCES


