Erlang Capacity of Multi-class TDMA Systems with Adaptive Modulation and Coding

Wang, Hua; Iversen, Villy Bæk

Published in:
Proceedings of IEEE ICC 2008

Link to article, DOI:
10.1109/ICC.2008.30

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Erlang Capacity of Multi-class TDMA Systems with Adaptive Modulation and Coding

Hua Wang and Villy B. Iversen
Department of Communications, Optics & Materials
Technical University of Denmark, Lyngby, Denmark
Email: {huw, vbi}@com.dtu.dk

Abstract—Erlang capacity is traditionally defined as the maximum value of offered traffic among different service classes that the system can support when the blocking probabilities at the call admission control (CAC) level do not exceed certain thresholds. That is valid when a fixed amount of bandwidth is allocated to each user in each frame during the whole service time. However, with the introduction of adaptive modulation and coding (AMC) scheme employed at the physical layer, outage might occur due to the fact that the allocation of bandwidth is dynamic based on the time-varying wireless channel conditions. In this paper, we evaluate the Erlang capacity of a TDMA system with AMC supporting voice and data traffics, by taking both the blocking and the outage probabilities into account. The analytical models for calculating the blocking and the outage probabilities are developed separately, and a joint algorithm for determining the Erlang capacity of the system is proposed with some numerical examples.

I. INTRODUCTION

Future generation wireless communication systems are evolving to provide a wide range of services, including voice, data and multi-media applications, with different quality of service (QoS) requirements, such as delay and throughput.

The economical usefulness of a system is effectively measured by the Erlang capacity, which is generally defined as the maximum traffic load that the system can support with a certain blocking probability requirement. Many models have been proposed at separate layers, such as the Rayleigh, Rician and Nakagami fading models at the physical layer [5][7], and queuing models at the data link layer [4]. In traditional channelized multiple access systems, e.g., TDMA and FDMA, each user is assigned a fixed amount of bandwidth during the whole service time, and the Erlang capacity can be easily obtained by using the well-known Erlang-B formula. However, traditional queuing models do not consider the time-varying nature of wireless channels due to multipath fading and Doppler shift. Unlike wired networks, even if large bandwidth is allocated to a certain wireless connection, the QoS requirements may not be satisfied when the channel experiences deep fades.

In order to enhance the spectrum efficiency while maintaining a target packet error rate (PER) over wireless links, adaptive modulation and coding (AMC) scheme has been widely adopted to match the transmission rate to time-varying channel conditions. With AMC, the allocation of bandwidth to each user is no longer deterministic (i.e., a fixed amount of bandwidth), but in a dynamic behavior. Therefore, the calculation of the Erlang capacity should take both the call admission control (CAC) module at the data link layer as well as the AMC module at the physical layer into considerations.

An analytical model to investigate the performance of transmissions over wireless links is developed in [1], where a finite-length queuing is coupled with AMC. However, the author only concentrates on a single-user case. Reference [2] calculates the Erlang capacity of WiMAX systems with fixed modulation scheme, where two traffic streams, streaming and elastic flows, are considered. In this paper, we investigate the Erlang capacity of multi-user multi-class TDMA Systems under the joint effect of CAC and AMC, by coupling the analysis of the blocking probability and the outage probability. We first present analytical models to calculate the blocking and the outage probabilities respectively. Based on that, a joint algorithm to determine the Erlang capacity of the system is proposed. The Erlang capacity in this paper is defined as the maximum traffic load among different service classes that the system can support when both the blocking and the outage probabilities do not exceed certain thresholds.

The rest of the paper is organized as follows. In Section II, we introduce the system model and the call admission control policy. In Section III, analytical models for calculating the blocking and the outage probability with AMC are presented together with the joint algorithm. Numerical results are shown in Section IV. Finally, a conclusion is drawn in Section V.

II. SYSTEM MODEL

We consider an infrastructure-based wireless access network, where connections are established between base station (BS) and mobile stations (MSs). Two types of services, voice service and data service, are supported by the system. Voice service is not tolerant to packet delay, thus requires a constant bit rate. Data service is more elastic in terms of being able to vary the transmission rate according to the channel conditions, but also requires a minimum throughput.

At the physical layer, the data is transmitted frame by frame, where each frame is comprised of a fixed number of time slots, each of which contains a fixed number of symbols. In the downlink, the transmission to different users is scheduled on a time-division multiplexing (TDM) fashion, while time-division multiple access (TDMA) is applied in the uplink. Adaptive modulation and coding scheme is employed at the physical layer, where multiple transmission modes are available, with each mode representing a pair of specific modulation format
traffic stream

the Erlang capacity of the system is proposed.

A. Analysis of Blocking Probability

We use the BPP (Binomial, Poisson & Pascal) traffic model in our analysis [4]. This model is insensitive to the service time distributions, thus is very robust for applications. Each traffic stream $i$ is characterized by the offered traffic $A_i$, the mean holding time $\mu_i$, and the number of channels $c_i$ needed for establishing one connection. The offered traffic $A_i$ is usually defined as the average number of call attempts per unit time.

Peakedness $Z_i$ is the variance/mean ratio of the state probabilities when the system capacity is infinite, and it characterizes the arrival process. For $Z_i = 1$, we have a Poisson arrival process, whereas for $Z_i < 1$, we have a finite number of users and more smooth traffic (Engset case). For $Z_i > 1$, it corresponds to a more bursty Pascal arrival process.

In this paper, we assume that voice users arrive following a fixed amount of bandwidth dedicated for voice users, and a forward error correcting (FEC) code. The transmission mode is determined by the instantaneous signal-to-noise ratio (SNR). We assume that the channel is frequency flat, and remains constant within a frame, but may vary from frame to frame. Furthermore, we assume that the BS has perfect knowledge of channel state information (CSI) of each user. Based on these assumptions, the AMC scheme adjusts its transmission mode on a frame-by-frame basis.

Voice and data users arrive at the cell in a random order. The CAC module decides whether an incoming call should be admitted or dropped. To prioritize voice service, CAC reserves a fixed amount of bandwidth dedicated for voice users, and allows both voice and data users to compete for the remaining bandwidths. Let $K$ denote the total number of time slots in a frame for data transmissions, of which $\Delta K$ time slots are reserved for voice users only. Let $\{n_v, n_d\}$ denote the number of admitted voice users and data users in the cell, respectively. Each voice user is allocated $\tau_v$ time slots per frame on average to ensure a constant bit rate, and each data user is allocated at least $\tau_d$ time slots per frame on average to guarantee a minimum throughput. The remaining time slots can be allocated to data users on demand to increase their throughput. Under the above conditions, the state space of the system is given by:

$$S := \{n_v, n_d\} \in \mathbb{N} \times \mathbb{N}$$

subject to:

$$n_v \cdot \tau_v + n_d \cdot \tau_d \leq K$$

$$n_d \cdot \tau_d \leq K - \Delta K$$

III. ANALYSIS MODELS

In this section, we first develop an analytical model based on multi-class loss systems [4] to calculate the state space and the blocking probability of each traffic stream. Then, given the distribution of the number of voice and data users admitted in the system, the outage probability caused by AMC can be derived by analyzing the stochastic model of different transmission modes. Finally, a joint algorithm for determining the Erlang capacity of the system is proposed.

The probability of having $a$ voice users and $b$ data users in the system can be represented as:

$$P(n_v = a) = \frac{q(n_v = a)}{\sum_{i \in \Omega_v} q(n_v = i)} \quad a \in \Omega_v$$

$$P(n_d = b) = \frac{q(n_d = b)}{\sum_{i \in \Omega_d} q(n_d = i)} \quad b \in \Omega_d$$

where $\Omega_v$ and $\Omega_d$ are the sets of possible number of voice and data users in each service class, $\lfloor x \rfloor$ denotes the largest integer not exceeding $x$.

$$\Omega_v := \{0, 1, \ldots, \left\lfloor \frac{K}{\tau_v} \right\rfloor \}$$

$$\Omega_d := \{0, 1, \ldots, \left\lfloor \frac{K - \Delta K}{\tau_d} \right\rfloor \}$$

$q(n_v)$ and $q(n_d)$ are the relative state probabilities, calculated as follows:

$$q(n_v = a) = \begin{cases} p_v(a \tau_v) \cdot \sum_{j=0}^{K-\Delta K} p_d(j) & \text{if } a \leq \left\lfloor \frac{\Delta K}{\tau_v} \right\rfloor \\ p_v(a \tau_v) \cdot \sum_{j=0}^{K-\Delta K} p_d(j) & \text{if } a > \left\lfloor \frac{\Delta K}{\tau_v} \right\rfloor \end{cases}$$

$$q(n_d = b) = p_d(b \tau_d) \cdot \sum_{j=0}^{K-\Delta K} p_v(j)$$

(5)
where $p_c(\cdot)$ and $p_d(\cdot)$ are the state probabilities of each traffic stream as if it is alone in the system:

$$
egin{align*}
  p_c(i \cdot \tau_v) &= \left\{
    \begin{array}{ll}
      \frac{\alpha^i}{\Gamma(1+\alpha)}\frac{(1-\alpha)^{\alpha^i}}{\sum_{j=0}^{\alpha^i}(1-\alpha)^{\alpha^i-j}} & \text{if } i \in \Omega_v \\
      0 & \text{else}
    \end{array}
  \right. \\
  p_d(i \cdot \tau_d) &= \frac{\alpha^i}{\sum_{j=0}^{\alpha} j^\alpha} \quad \text{if } i \in \Omega_d \\
  p_d(\cdot) &= 0 \quad \text{else}
\end{align*}
$$

where $\alpha = 1 - Z_v$.

Based on the state probabilities, the closed-form of the call blocking probability of each traffic stream can be obtained by using the convolution algorithm, shown in Eqn. (8) & (9).

**B. Analysis of Outage Probability**

In AMC scheme, the modulation mode and coding rate is chosen depending on the time-varying channel conditions. As a consequence, the number of time slots allocated to each user is varying on a frame by frame basis. Outage is defined to occur when the total number of time slots required by the admitted users exceeds the total available time slots.

1) AMC: For flat fading channels, we adopt the general Nakagami-$m$ model to describe the received signal-to-noise ratio (SNR) $\gamma$ statistically, which is a random variable with Gamma probability density function [1]:

$$
p_c(\gamma) = \frac{m^m \gamma^{m-1}}{\overline{\gamma}^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\overline{\gamma}}\right)
$$

where $\overline{\gamma} := E[\gamma]$ is the average received SNR, $\Gamma(m) := \int_0^\infty t^{m-1} \exp(-t) dt$ is the Gamma function and $m$ is the Nakagami fading parameter ($m \geq 1/2$).

The design objective of AMC is to maximize the data rate by adjusting the transmission parameters according to channel conditions, while maintaining a prescribed packet error rate (PER) $P_b$. Let $N$ denote the total number of transmission modes available ($N = 5$ in this paper). Assuming constant power transmission, we partition the entire SNR range into $N + 1$ non-overlapping consecutive intervals with boundaries denoted as $\{\gamma_n\}_{n=1}^{N+1}$. Specifically, mode $n$ is chosen when $\gamma \in [\gamma_n, \gamma_{n+1})$. Therefore, mode $n$ will be chosen with probability:

$$
P_c(n) = \int_{\gamma_n}^{\gamma_{n+1}} p_c(\gamma) d\gamma = \frac{\Gamma(m, \frac{m\gamma_n}{\overline{\gamma}}) - \Gamma(m, \frac{m\gamma_{n+1}}{\overline{\gamma}})}{\Gamma(m)}
$$

where $\Gamma(m, x) := \int_x^\infty t^{m-1} \exp(-t) dt$ is the complementary incomplete Gamma function.

The closed-form of the average PER corresponding to mode $n$ is obtained as [1]:

$$
\text{PER}_n = \frac{1}{P_c(n)} \int_{\gamma_n}^{\gamma_{n+1}} a_n \exp(-g_n \gamma) p_c(\gamma) d\gamma
$$

where $a_n$, $g_n$ are the mode dependent parameters shown in Table I.

The algorithm for determining the threshold $\{\gamma_n\}_{n=1}^{N+1}$ with the prescribed PER $P_b$ is described in details in [1].

<table>
<thead>
<tr>
<th>Mode</th>
<th>Mode 1</th>
<th>Mode 2</th>
<th>Mode 3</th>
<th>Mode 4</th>
<th>Mode 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modulation</td>
<td>BPSK</td>
<td>QPSK</td>
<td>QPSK</td>
<td>16QAM</td>
<td>64QAM</td>
</tr>
<tr>
<td>Coding rate</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>$R_n$ (bits/sym)</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>$a_n$</td>
<td>274.7229</td>
<td>90.2512</td>
<td>67.6181</td>
<td>53.3987</td>
<td>35.3508</td>
</tr>
<tr>
<td>$g_n$</td>
<td>7.9932</td>
<td>3.4998</td>
<td>1.6883</td>
<td>0.3756</td>
<td>0.0900</td>
</tr>
<tr>
<td>$\gamma_{\text{ext}}$ (dB)</td>
<td>-1.5331</td>
<td>0.1042</td>
<td>3.9722</td>
<td>10.2488</td>
<td>15.9784</td>
</tr>
</tbody>
</table>

**TABLE I**

**TRANSMISSION MODES WITH CONVOLUTIONALLY CODED MODULATION** [1]

2) Outage Probability: Assume that there are $G_v$ voice users and $G_d$ data users admitted in the system, respectively. Each user has an ON/OFF activity model, represented by a random variable $\xi$, with probability $P(\xi_v = 1) = \alpha_v$ for voice users, and $P(\xi_d = 1) = \alpha_d$ for data users. We further assume that voice users require a constant bit rate of $R_v$ bits per frame, and data users require a minimum throughput of $R_d$ bits per frame. Hence, the total amount of required bandwidth in terms of time slots per frame is given by:

$$
Z = \sum_{i=1}^{G_v} \xi_{v,i} N_{v,i} + \sum_{j=1}^{G_d} \xi_{d,j} N_{d,j}
$$

where $\xi_{v,i}$ and $\xi_{d,j}$ are the activity factors for the $i^{th}$ voice user and the $j^{th}$ data user, respectively. $G_v$ and $G_d$ are random variables representing the number of voice and data users admitted in the system. $N_{v,i}$ and $N_{d,j}$ denote the number of time slots required by the $i^{th}$ voice user and the $j^{th}$ data user.
respectively, depending on the current channel conditions.

\[ \begin{align*}
N_{v,i} &\in \mathcal{N}_v, \quad \mathcal{N}_v := \left\{ \left\lfloor \frac{R_{c_i}}{sR_n} \right\rfloor, \ n = 0, \cdots, N \right\} \\
N_{d,j} &\in \mathcal{N}_d, \quad \mathcal{N}_d := \left\{ \left\lfloor \frac{R_d}{sR_m} \right\rfloor, \ n = 0, \cdots, N \right\}
\end{align*} \]  

(14)

where \( s \) is the fixed number of symbols per time slot, \( R_n \) denotes the number of bits carried per symbol in transmission mode \( n \) shown in Table I, and \( \lfloor x \rfloor \) denotes the smallest integer larger than \( x \). \( N_{v,i} \) and \( N_{d,j} \) are random variables with probability mass function:

\[ \begin{align*}
P(N_{v,i} = \left\lfloor \frac{R_{c_i}}{sR_n} \right\rfloor) &= P_r(n) \quad n = 0, \cdots, N \\
P(N_{d,j} = \left\lfloor \frac{R_d}{sR_m} \right\rfloor) &= P_r(m) \quad m = 0, \cdots, N
\end{align*} \]  

(15)

Since \( Z \) is a sum of random variables, we can approximate \( Z \) to be a Gaussian random variable by applying the central limit theorem. Therefore, the outage probability of the system can be approximated by:

\[ P_{\text{outage}} = P\{Z > K\} \approx Q\left( \frac{K - E[Z]}{\sqrt{\text{Var}(Z)}} \right) \]  

(16)

It is assumed that \( N_{v,i}, N_{d,j}, \xi_{v,i} \) and \( \xi_{d,j} \) are independent and identically distributed random variables. Furthermore, \( \xi_{v,i}, N_{v,i} \) and \( \xi_{d,j}, N_{d,j} \) are independent of each other. \( G_v \) and \( G_d \) follow probability mass function shown in Eqn. (2) & (3), respectively. The mean and variance of \( Z \) are given by:

\[ \begin{align*}
E[Z] &= E\left\{ E[Z \mid (G_v, G_d)] \right\} \\
\text{Var}(Z) &= E\left\{ V[Z \mid (G_v, G_d)] \right\} + V\left\{ E[Z \mid (G_v, G_d)] \right\} \\
\end{align*} \]  

(18)

C. Calculation of Erlang Capacity

The Erlang capacity in this paper is defined as the maximum value of offered traffic in each traffic stream \( \{A_v, A_d\} \) that the system can support when the blocking probability and the outage probability do not exceed certain thresholds.

From the analytical models developed in the previous two subsections, we note that the blocking probability is determined by the traffic loads \( \{A_v, A_d\} \) and the average number of time slots occupied per connection \( \{\bar{\tau}_v, \bar{\tau}_d\} \), which depend on the outage probability. On the other hand, the outage probability is determined by the distributions of the number of voice and data users \( \{G_v, G_d\} \), which depend on the blocking probability. Therefore, the analytical models of the blocking and the outage probabilities are tightly coupled. In this subsection, we propose a joint algorithm to calculate the Erlang capacity of multi-class TDMA systems with AMC.

The joint calculation algorithm consists of iterations of two steps. In the first step, given the traffic loads of voice users and data users \( \{A_v, A_d\} \), the blocking probabilities are obtained such that the outage probability requirement \( P_{\text{outage}} \) is satisfied. Specifically, the proposed algorithm first sets the initial values of \( \bar{\tau}_v = E[N_v] \) and \( \bar{\tau}_d = E[N_d] \), which are the average number of time slots needed for establishing one connection of voice and data services, respectively, and calculates the user distributions, the blocking probabilities, and the outage probability by using Exp. (2), (3), (8), (9) & (16). Then it checks whether \( P_{\text{outage}} \) meets its requirement or not at the current values of \( \bar{\tau}_v \) and \( \bar{\tau}_d \). The values of \( \bar{\tau}_v \) and \( \bar{\tau}_d \) are increased if \( P_{\text{outage}} \) is larger than the threshold, and are decreased otherwise. Unless \( P_{\text{outage}} \) meets its requirement, \( \bar{\tau}_v \) and \( \bar{\tau}_d \) are further increased or decreased. Once \( P_{\text{outage}} \) is satisfied, the blocking probabilities are recorded under given traffic loads. Then the values of \( \{A_v, A_d\} \) are modified and the above procedure is repeated. After we get the blocking probabilities under various traffic loads, the algorithm proceeds to the second step, where the Erlang capacity region is obtained by extracting the maximum values of \( \{A_v, A_d\} \) that meet both requirements of blocking and outage probabilities. The flow chart of the joint algorithm is shown in Fig. 2.

IV. NUMERICAL RESULTS

In this section, we present some numerical results based on the analytical models developed in Section III. The system parameters are listed in Table II.

We set the target outage probability to be 2%, which is usually considered to be an acceptable QoS requirement. Fig. 3 shows the outage probability versus different traffic loads of voice and data services in Erlangs. We can see
from the figure that our proposed algorithm can well keep the outage probability around the predefined threshold under various traffic loads. This is achieved by properly adjusting the values of \{\tau_v, \tau_d\}, which in turn changes the distribution of the number of voice and data users in the system.

A two dimensional Erlang capacity region at 2% blocking probability is shown in Fig. 4 over different outage probabilities and over different bit rates of data users. We can see in all scenarios that the increase of data traffic will severely reduce the capacity for voice users, and this phenomenon becomes more obvious when the bit rate of data users increases from 128 bits/frame to 256 bits/frame. This is generally because data traffic usually has higher bit rate and activity factor, thus results in the loss of voice Erlang capacity more sharply. Furthermore, we also observe that by changing the threshold of the outage probability from 2% to 6%, the Erlang capacity region of the system can only be marginally increased.

V. CONCLUSION

Adaptive modulation and coding (AMC) has been widely used to match transmission parameters to time-varying channel conditions. In this paper, we have investigated the performance of a TDMA system that carries both voice and data traffics with AMC in terms of the Erlang capacity. We have developed analytical models to calculate the blocking and the outage probabilities. As the outage probability depends on the distribution of the number of users in the system, which is tightly coupled with the blocking probability, we proposed a joint algorithm to determine the Erlang capacity that the system can support when both the blocking and the outage probabilities are within certain thresholds. Numerical results have shown that the voice traffic capacity is severely affected by the data traffic. Moreover, we have also observed that the Erlang capacity region of the system can only be marginally increased by loosing the target outage probability.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Frame duration</td>
<td>1 ms</td>
</tr>
<tr>
<td>Number of time slots in the downlink, K</td>
<td>800</td>
</tr>
<tr>
<td>Number of symbols per time slot, s</td>
<td>4</td>
</tr>
<tr>
<td>Reserved time slots for voice users, ΔK</td>
<td>100</td>
</tr>
<tr>
<td>Nakagami fading parameter, ( m )</td>
<td>1</td>
</tr>
<tr>
<td>Average SNR, ( \gamma )</td>
<td>15 dB</td>
</tr>
<tr>
<td>Target PER, ( P_{T} )</td>
<td>( 10^{-4} )</td>
</tr>
<tr>
<td>Constant bit rate for voice users, ( R_{c} )</td>
<td>64 bits/frame</td>
</tr>
<tr>
<td>Minimum bit rate for data users, ( R_{d} )</td>
<td>128 bits/frame</td>
</tr>
<tr>
<td>Voice activity factor, ( \xi_{v} )</td>
<td>0.4</td>
</tr>
<tr>
<td>Data activity factor, ( \xi_{d} )</td>
<td>0.6</td>
</tr>
<tr>
<td>Peakedness of voice traffic, ( Z_{v} )</td>
<td>0.5</td>
</tr>
<tr>
<td>Peakedness of data traffic, ( Z_{d} )</td>
<td>1</td>
</tr>
</tbody>
</table>

**TABLE II**

SYSTEM PARAMETERS USED FOR THE NUMERICAL EVALUATION

**REFERENCES**


