

Teletraffic Analysis for VoIP Services in WLAN Systems with Handoff Capabilities

Sergio Sánchez-Reyes¹, Mario E. Rivero-Angeles², Noé Torres-Cruz¹

¹ Instituto Politécnico Nacional, Telematics Department, UPIITA-IPN, Mexico

² Instituto Politécnico Nacional, CIC-IPN, Mexico

{ssanchezr0602, mriveroa, ntorresc}@ipn.mx

Abstract. In this work, a TDMA-based system is studied for VoIP services with and without mobility. The system is considered to be dedicated for voice-only services such as a WLAN environment under the PCF mode or a cellular system. In such environments, mobiles usually have mobility. This introduces the need to study the system when a handoff procedure is enabled. The teletraffic analysis considers call arrivals, call completions, and ON/OFF activity processes for individual VoIP sessions. The VoIP system is simulated in order to verify the accuracy of the analytical results. Since many papers published in the literature consider a fixed number of VoIP users in the system, this analysis can be useful to estimate the system's performance when the system is in statistical equilibrium. As an additional feature of this work, a simple fluid model to calculate the number of active and inactive users in a VoIP system with and without handoff capabilities is proposed and developed.

Keywords. Teletraffic analysis, packet delay, packet dropping, fluid model, VoIP, on/off traffic, active users.

1 Introduction

Voice over Internet protocol (VoIP) is the technology that enables people to use the Internet as the transmission medium for voice communications. Recently, this technology has drawn much interest in wireless networks because of the convergence of all IP architecture in wireless and wireline networks. Compared to traditional resource-dedicated Public Switched Telephone

Network (PSTN), IP network is resource shared. For this reason, IP-based VoIP applications are cost-efficient. Hence, they are being considered as the principal means of voice transmission in cellular systems (LTE networks) and other systems such as WLAN.

Wireless Local Area Networks (WLANs) are expected to carry an important amount of VoIP traffic either through stationary desktops or by mobile devices connected to the system due to the universal access of such networks. As such, it is important to evaluate the system's performance for both scenarios. Traditionally, data exchange in WLANs is done using the DCF (Distributed Coordination Function) mode under the CSMA/CA protocol. However, the standard defines the PCF (Point Coordination Function) mode based on a centralized polling scheme in order to transmit delay sensitive data through the access point (AP).

In the DCF operation mode data transmissions are not guaranteed due to the random nature of transmissions while in the PCF operation, there are no collisions and all stations waiting for transmission occupy the channel in an ordered fashion under the control of the AP. However, the PCF mode has not been extensively used or studied since it is considered to be inefficient [8]. The reason for this is that in order to initiate the PCF mode, the AP has to take control of the channel competing with the rest of the nodes in the DCF mode. As such, the start of

the PCF mode cannot be guaranteed to start in a periodical fashion. Nonetheless, in [10] and [11], the performance of the PCF mode rendered numerical results that demonstrate the adequate performance of the system for time-sensitive data under certain conditions.

Furthermore, if the WLAN is considered to be dedicated to voice services, this main drawback of the PCF mode no longer holds since there are no transmissions in the DCF mode. Thus there is no *stretching* of the DCF period. This could provide a low cost solution to implement voice services with an acceptable QoS.

For cellular systems, the use of VoIP services has been considered before. Compared to traditional resource-dedicated Public Switched Telephone Network (PSTN), IP network is resource shared. For this reason, IP-based VoIP applications are cost-efficient. However, traditional IP networks and protocols only provide best effort service and do not address the Quality of Service (QoS) requirements of a real-time and interactive application like voice telephony.

As such, transport of voice over an IP-based network demands a different set of QoS metrics than those required for the transport of data. As such, different mechanisms have been proposed in [12] - [14]. However, the mobility of nodes has not been addressed in these works.

The Markov modulated Poisson process (MMPP) has been employed for modeling the aggregated traffic generated by the VoIP users in different wireless systems [1] – [5]. It has been shown that a superposition of a number of ON/OFF voice sources can be approximated by a two-state MMPP [6] – [7]. As such, the authors of [1] – [5] developed their performance analysis for a fixed total number of VoIP users (i.e., VoIP sessions). Hence, these analyses fail to capture the dynamic of session arrivals and departures.

In this paper, we aim to study the average buffer length and average packet delay in a WLAN under the PCF operation mode considering the arrival and departure of users in the system as well as the active and inactive periods of users currently connected to an AP.

Also, the case where users are mobile and perform a handoff procedure when moving away

from their current AP or Base Station (BS) towards a different AP or BS with better channel conditions. Specifically, for the WLAN environment, either at home or at the office it is not uncommon to have more than one AP in order to increase the system's coverage. As such, in VoIP services it is important for the system to guarantee that an ongoing call is not dropped when a user moves inside the building. For the cellular system, the handoff procedure is of major importance since it is considered to be a basic QoS parameter. In this work, we focus on the case where only two APs or BSs are installed. However, the system can be extended to consider multiple APs. Note that this extension is straightforward when homogenous conditions are assumed.

As an additional feature of this work, a simple fluid model to calculate the number of active/inactive VoIP users when the system is in statistical equilibrium is proposed. The expressions obtained represent an alternative to the *conventional* average values of active/inactive users. The main advantage of this model is that it achieves closer results to the simulated system while providing simple expressions that can be used as reference for a more elaborated analysis.

The teletraffic analysis considers call arrivals, call completions, and ON/OFF activity processes for individual VoIP sessions. This analysis can be particularly useful for different radio systems such as the ones studied in [15], [16] where users make use of an ON/OFF type of service, such as, but not limited to, VoIP. Indeed, in modern radio systems, it is important to know the number of active/inactive users in order to analyze the system's capacity and possible interference among users. It is important to notice that this fluid model can be easily extended to other types of services with an ON/OFF characteristic such as the ones studied in [17] and [18].

The analytical model is compared to simulated results in order to verify its accuracy. Throughout the paper, the terms active state and ON state are used interchangeable to represent the state where VoIP users generate packets while the terms inactive state and OFF state are used interchangeable to represent the state where users remain silent. Hence, they are not generating

voice packets. Also the term AP and BS are used interchangeably since we consider that either a WLAN under the PCF mode or a cellular system have the same TDMA basic operation.

The rest of the paper is organized as follows: Section II summarizes the related work previously published in the literature. Section III presents the model and main assumptions used throughout the paper. Section IV focuses on the development of the proposed fluid model while Section V develops the teletraffic analysis. Finally, section VI presents some relevant numerical results. The paper concludes with the main results of this work.

2 Related Work

In [23] and [24], the aggregated talk spurt is modeled by a Poisson process, where its average arrival rate depends on the average talk spurt duration, the voice activity factor, and the offered traffic load. The Poisson arrival model, however, cannot capture the interframe dependency between consecutive frames. To overcome this drawback, the Markov modulated Poisson process (MMPP) has been employed for modeling the aggregated traffic generated by VoIP users in different wireless systems [3]- [4]. It has been shown that a superposition of a number of ON/OFF voice sources can be approximated by a two-state MMPP. Nevertheless, in these works, performance analysis is developed for a fixed total number of VoIP users (i.e., VoIP sessions) and, consequently, it is not possible to capture the dynamic of arrival and departure of VoIP sessions. All of the previously described works only consider packet delay as performance metric.

Concerning the study of VoIP services in WLAN (VoWLAN), little attention has been given to the performance analysis of the PCF mode. In [10] an in-depth operation description of the PCF mode and a voice capacity analysis is performed considering the active/inactive nature of packetized voice services when silence suppression is used. However, this analysis is performed for a maximum fixed number of nodes with a combined operation of the DCF and PCF modes.

As such, the arrival and departure of nodes is not considered. Also, the handoff capabilities is

not considered. In [11] the PCF mode is evaluated for VoIP traffic with active/inactive periods through simulations. But again, a fixed number of nodes is considered and no handoff capacity is assumed.

Many studies have considered the DCF mode to transmit voice services, such as [8] and [9]. However, these studies consider a fixed number of stations which constantly transmit information, i.e., no silence suppression is considered under saturation conditions. In contrast to these studies, we consider both the arrival and departure of nodes in the system as well as the use of silence suppression schemes where nodes with an ongoing transmission can either be active (sending packets to the AP) or inactive (in a silence period). Also, handoff capabilities are considered. To the best of our knowledge, the handoff procedure has not been considered before in such networks.

The transmission of VoIP services in cellular networks have been considered in such works as [12]- [13]. In these works different Call Admission Control strategies are reviewed and studied. However, these works do not consider explicitly the mobility of nodes. Specifically, no handoff procedures are considered.

3 System Model

In this work we are interested on studying the performance of TDMA-based systems like the ones used in cellular networks, such as GSM, GPRS or EDGE. Also, wireless local area networks such as the IEEE 802.11 standard under the PCF mode, which is specifically used for time sensitive data transmissions.

In the basic model, we assume a single access point or base station that emulates the data transmission over N shared channels for a general case. In the rest of the paper, we use interchangeably the terms of Access Points and Base Stations to indicate the wireless transmitter or receptor where many users are connected.

In the PCF based system in a WLAN scenario, where a polling scheme is used, the system can be approximated as a TDMA-like scheme where each transmission opportunity can be considered as a time slot. As such, each channel can have up to N polling opportunities which are equivalent

to N time slots. Since only voice packets are transmitted, the system does not use the DCF mode. Hence, this time structure can be assumed to be periodic with duration T_F sec. In the cellular system scenario, N represents the number of traffic channels in the network.

It is assumed that users requesting a VoIP session are able to connect (and stay connected up to the end of their sessions) to this access point, that is, no handoff mechanism is considered (typical in office and home scenarios where users access the network through a static computer).

In the handoff scenario, a system with only two access points is considered. Users can go from one access point to the other without having to reinitialize the call. This corresponds to a mobility scenario in the office or at home where users access the network using a mobile device. We consider that two access points are enough to study the impact of mobility in a WLAN or cellular system under VoIP services.

Statistical characterization of telephone conversations resulted in the main finding that a telephone conversation can be represented by the ONOFF pattern [22]. We assume that these periods are geometrically distributed with parameters $P_{N/F}$ and $P_{F/N}$, respectively. The parameters $P_{N/F}$ and $P_{F/N}$ represent the probability that a call that is currently in the active or silent period turns to a silent or active period respectively. During the active periods, we assume that the voice samples are generated with a constant bit rate. In the VoIP call model used in this work, each voice packet is composed of protocol headers and the payload.

It is considered that all users use the same voice codec, and hence both the size of VoIP packet and packet interval are identical. In our model, we consider that during the active periods the voice packets are deterministically generated every T_F seconds, which is the frame duration. A typical value of T_F is 20 ms [24]. It is considered that the frame duration is equal to T_F seconds. In each frame, the system can transmit as many voice packets (from different active users) as the transmission capacity of the system allows. This capacity depends on the transmission capability of each user (i.e., modulation and coding scheme).

The user arrival model considers that the time between consecutive arrival calls is exponential distributed with mean $1/\lambda$. This continuous time process can also be described by a discrete time process as in [21]. On the other hand, call duration is also considered to be exponentially distributed with mean $1/\mu$. Building on this, the arrival process is modeled as a Bernoulli process with parameter $P_A = \lambda T_F$. Also, the probability that one user finishes its session in a frame duration can be approximated by a Geometric process with parameter $P_m = \mu T_F$.

In a similar way, an active user becomes inactive with probability $P_{N/F} = \beta T_F$ and an inactive user becomes active with probability $P_{F/N} = \alpha T_F$. Without loss of generality, we consider that all arriving users are in the active (ON) state.

4 Fluid Model

The teletraffic model is based on a simple fluid model such as the one proposed on [19]. New users arrive to the system with rate λ according to a Poisson process and they are considered to be in an active mode. The model for the system with no handoff, also called *basic model* and with handoff is presented.

4.1 Basic Model

The call duration is assumed to be exponentially distributed with mean $1/\mu$. An ON/OFF traffic is considered where users remain in the ON (OFF) state for a random time exponentially distributed with mean $1/\alpha$ ($1/\beta$). Let x be the instantaneous number of active users, i.e., in the ON state, and y be the instantaneous number of inactive users, i.e., in the OFF state. From the previous description, the evolution with time x and y satisfies:

$$x'(t) = \lambda + \beta y(t) - (\alpha + \mu)x(t), \quad (1)$$

$$y'(t) = \alpha x(t) - (\beta + \mu)y(t). \quad (2)$$

This dynamical system is analyzed and its equilibrium point, denoted by (\bar{x}, \bar{y}) , is studied. The equilibrium is obtained by solving the system

$\{x'(t) = 0, y'(t) = 0\}$ and by replacing $x(t)$ by \bar{x} and $y(t)$ by \bar{y} . The result is:

$$\bar{x} = \frac{\lambda(\beta + \mu)}{\mu(\mu + \beta + \alpha)}, \quad \bar{y} = \frac{\lambda\alpha}{\mu(\mu + \beta + \alpha)}. \quad (3)$$

Note that, as expected, the total number of users in steady state is given by $\bar{U}_S = \bar{x} + \bar{y} = \lambda/\mu$. On the other hand, the average number of users in the ON and OFF states can be calculated as $(\bar{U}_S)(p_{ON})$ and $(\bar{U}_S)(p_{OFF})$ respectively, where p_{ON} is the probability that a user is in the ON state and p_{OFF} is the probability that a user is in the OFF state. These probabilities can be calculated as in [20] and are given by $p_{ON} = (\beta)/(\alpha + \beta)$ and $p_{OFF} = (\alpha)/(\alpha + \beta)$ respectively. Hence, the average number of active and inactive users in the system can be found as:

$$E[x] = \frac{\lambda\beta}{\mu(\alpha + \beta)}, \quad E[y] = \frac{\lambda\alpha}{\mu(\alpha + \beta)}. \quad (4)$$

4.2 Handoff Model

In this case, VoIP users can enter or leave a particular coverage area with rate γ . As such, the evolution with time of the system can be described as:

$$x'(t) = \lambda + \gamma + \beta y(t) - (\alpha + \mu + \gamma)x(t), \quad (5)$$

$$y'(t) = \gamma + \alpha x(t) - (\beta + \mu + \gamma)y(t). \quad (6)$$

As in the previous case, the system is analyzed and its equilibrium point, denoted by (\bar{x}, \bar{y}) , is studied. The result is:

$$\bar{x} = \frac{\gamma^2 + \gamma\mu + \gamma\lambda + \mu\lambda + 2\gamma\beta + \lambda\beta}{(\gamma + \mu)(\gamma + \alpha + \mu + \beta)}, \quad (7)$$

$$\bar{y} = \frac{\gamma^2 + 2\gamma\alpha + \gamma\mu + \alpha\lambda}{(\gamma + \mu)(\gamma + \alpha + \mu + \beta)}. \quad (8)$$

As such, the total number of users in the system is given by $\bar{x} + \bar{y} = (\lambda + 2\gamma)/(\gamma + \mu)$.

5 Teletraffic Analysis

In this section, a delay and packet dropping probability analysis is performed for a VoIP service in a TDMA-based system. The mathematical analysis presented assumes that the system evolves in a discrete time basis where the unit of time is considered to be the time slot (time required for the transmission of a voice packet including all the control packet associated with this procedure).

Additionally, since the frame duration is considered to be sufficiently small compared to the time between arrivals, service time, cell dwelling times, etc., it is assumed that there can only be one user arrival to the system, or one user ending its session, or one user in OFF mode becoming active, or one in ON mode going to stand by.

Using the system model explained in previous sections, the probability that one new user arrives to the system, one user leaves the system, goes from the ON state to the OFF state or vice versa can be calculated approximately, respectively, as:

$$P_A = e^{-\lambda T_F} \lambda T_F, \quad (9)$$

$$P_M = U P_m (1 - P_m)^{U-1}, \quad (10)$$

$$P_{N/F} = x P_{n/f} (1 - P_{n/f}^{x-1}), \quad (11)$$

$$P_{F/N} = (U - x) P_{f/n} (1 - P_{f/n}^{U-x-1}), \quad (12)$$

where P_A is the probability that only one user arrives to the system, P_M is the probability that one user leaves the system at the end of the frame, $P_{N/F}$ is the probability that one of the active users goes to the inactive state, and $P_{F/N}$ is the probability that one inactive user becomes active.

Notice also that there can be up to N simultaneous transmissions in a time slot duration, and there are M time slots per frame. Also, the buffer can allocate up to Q packets waiting for resources to be transmitted. In order to avoid congestion, a limit on the number of users with an ongoing call is set at $(MN) + Q$.

Hence, if a user arrives and all the resources are occupied and the buffer is filled, the arriving user is blocked. From this description, we analyze the case of the basic model and the system with mobility, where handoffs are enabled.

5.1 Basic Model

The aforementioned system can be modeled as a discrete-time Markov chain $W(t) = (U(t), x(t), y(t), s(t))$ where $U \geq 0$, $0 \leq x \leq U$, $0 \leq y \leq MN + Q$, and $1 \leq s \leq M$. $W(t)$ is a discrete time homogeneous Markov chain, with initial state $W(0) = (0, 0, 0, 1)$ that takes into consideration the total number of users in the system U , the number of active users x , the number of packets to be transmitted y and the number of time slot inside the frame, s . In the following, we denote $U(t)$ by U , $x(t)$ by x , $y(t)$ by y and $s(t)$ by s . Starting from state (U, x, y, s) the possible transitions are as follows:

- To state $(U, x, y-N, s+1)$ for $y > N$ and $s > M$ with probability 1. This transition corresponds to the case when the number of packets waiting for transmission is higher than the capacity of the system inside the frame period. As such, not all packets can be transmitted in slot s . Note that the number of users, and active users is not changed inside the frame period.
- To state $(U, x, 0, s+1)$ for $y \leq N$ and $s > M$ with probability 1. This transition corresponds to the case when the number of packets waiting for transmission is lower than the capacity of the system inside the frame period. As such, all awaiting packets can be transmitted in the slot s . Also note that the number of users, and active users is not changed inside the frame period.
- To state $(U+1, x+1, y, 1)$ for $s = M$, with probability $P_A(1 - P_M)(1 - P_{N/F})(1 - P_{F/N}) + P_A(1 - P_M)P_{N/F}P_{F/N}$. This transition corresponds to the case when the system is at the end of the frame and one new user arrives to the system but no active user becomes inactive and no inactive user becomes active and non of the users leave the system. Note that the system goes to the beginning of the frame ($s = M$). It also corresponds to the case where one user arrives to the system, there are no departures of the system but one active user becomes inactive and one inactive user becomes active.
- To state $(U-1, x, y, 1)$ for $s = M$, with probability $(1 - P_A)\{(1 - P_m)^x\}\{(U - x)P_m(1 - P_m)^{U-x-1}\}(1 - P_{N/F})(1 - P_{F/N}) + (1 - P_A)\{(xP_m(1 - P_m)^{x-1})\}(1 - P_m)^{U-x}P_{F/N}(1 - P_{N/F})$. This transition corresponds to the case when the system is at the end of the frame and no new users arrive to the system, no active user becomes inactive and no inactive user becomes active and one inactive user leaves the system. This can also occur when there are no arrivals to the system, but one active user leaves the system and one inactive user becomes active and no active user goes inactive.
- To state $(U-1, x-1, y, 1)$ for $s = M$, with probability $(1 - P_A)\{(1 - P_m)^{U-x}\}\{xP_m(1 - P_m)^{x-1}\}(1 - P_{N/F})(1 - P_{F/N}) + (1 - P_A)\{(U - x)P_m(1 - P_m)^{U-x-1}\}(1 - P_m)^xP_{N/F}(1 - P_{F/N})$. This transition corresponds to the case when the system is at the end of the frame and no new users arrive to the system, no active user becomes inactive and no inactive user becomes active and one active user leaves the system. This also occurs when there are no arrivals to the system, one inactive user leaves the system, one active user becomes inactive and no inactive users become active.
- To state $(U, x+1, y, 1)$ for $s = M$, with probability $(1 - P_A)(1 - P_M)(1 - P_{N/F})(P_{F/N}) + P_A\{(1 - P_m)^{U-x}\}\{xP_m(1 - P_m)^{x-1}\}(1 - P_{N/F})(1 - P_{F/N})$. This transition corresponds to the case when the system is at the end of the frame and no new users arrive to the system, no active user becomes inactive and one inactive user becomes active and no user leaves the system. But also to the case when there is one arrival and one inactive user leaves the system while no active users become inactive and no inactive users become active.
- To state $(U, x-1, y, 1)$ for $s = M$, with probability $(1 - P_A)(1 - P_M)(P_{N/F})(P_{F/N}) + P_A\{(xP_m(1 - P_m)^{x-1})\}(1 - P_m)^{U-x}P_{N/F}(1 - P_{F/N})$. This transition corresponds to the case when the system is at the end of the frame and no new users arrive to the system, one

active user becomes inactive and no inactive user becomes active and no user leaves the system. This can also occur when there is one arrival to the system, one active user leaves, one active user becomes inactive and no inactive user becomes active.

- To state $(U, x, y, 1)$ for $s = M$, with probability $(1 - P_A)(1 - P_M)(P_{N/F})(P_{F/N}) + P_A\{(1 - P_m)^x\}\{(U - x)P_m(1 - P_m)^{U-x-1}\}(1 - P_{N/F})(1 - P_{F/N})$. This transition corresponds to the case when the system is at the end of the frame and no new users arrive to the system, one active user becomes inactive and one inactive user becomes active and no user leaves the system. But also to the case when there is one arrival and one active user leaves the system while no inactive user become active and no inactive users become active.
- To state $(U+1, x+2, y, 1)$ with probability $P_A(1 - P_M)P_{F/N}(1 - P_{N/F})$. This transition occurs when there is one arrival to the system, there are no departures, one inactive user becomes active and no active users become inactive.

This chain is numerically solved in order to find the steady state distributions Π , i.e., the vector of the probabilities that the system is in state $\pi_{U,x,y,s}$. When the system is in such conditions that the number of packets waiting to be transmitted is $MN + Q$, i.e., the buffer is full and no more packets can be transmitted, any arriving user can be blocked. This happens with probability:

$$P_B = \pi_{U,x,MN+Q,M}\{P_A(1 - P_M)(1 - P_{N/F}), \\ + (1 - P_A)(P_{F/N})(1 - P_M)(1 - P_{N/F})\}.$$

Building on this, the average length of the queue can be obtained as:

$$L = \sum_{u=0}^{MN+Q} \sum_{k=0}^u \sum_{i=MN+1}^{MN+Q} \sum_{j=1}^M \pi_{U=u,x=k,y=i,s=j}(i - MN). \quad (13)$$

From this and using Little's theorem, the average packet delay for the active users can be calculated as:

$$W = \frac{L}{\lambda P_{ON}}, \quad (14)$$

where $P_{ON} = \alpha/(\alpha + \beta)$.

5.2 Handoff Model

In this scenario, users can go from one access point to another without losing the established connection with rate γ . As such, the total arrival rate to an access point is given by $\lambda' = \lambda + (\gamma U)$ while the total departure rate from the access point is given by $\mu' = \mu + (\gamma U)$. Since a user in the handoff procedure is always active, the transition probabilities considered in the basic model still holds replacing λ by λ' and μ by μ' .

6 Numerical Results

In this section, the WLAN dedicated for VoIP under the PCF mode is evaluated considering the average packet delay and packet loss probability with and without handoff capabilities between only two access points. Also, the accuracy of the fluid model proposed is studied. For these numerical results, the system's parameters are as follows unless stated otherwise: $N = 1$, $\lambda=0.2$, $1/\mu=180$ sec., $\alpha = 0.02$, and $\beta=0.01$.

6.1 Fluid Model

Fig. 1 presents both the average number of active users (obtained through simulations and analytically) and the number of active users in steady state. Fig. 2 presents the same results for the inactive users. It is clear from these results that when α is low, the number of active users in the system is high and the number of inactive users is low. This is because VoIP users spend more time in the active state. On the other hand, when β is low, the number of active users is low and the number of inactive users is high.

This is clearly due to the fact that users spend more time in the inactive state. In a more detailed analysis, it can be seen that the results obtained through the fluid model in (3) closely match the average values obtained through the simulations for any value of α and β .

Conversely, the average number of active users calculated using (4) are not as close to the simulation results. Specifically, for low values of α and β the difference is more noticeable. This implies that there is a relatively good match

between (4) and the simulation results when users spend more time in either the active or the inactive state, and there are not as much transitions. Also note that, for high values of β , i.e., when the average time in the OFF state is low, the average values calculated with (4) are the closest to the simulation results. Since the fluid model has been validated, in the following results, we only show the fluid model instead of the average values.

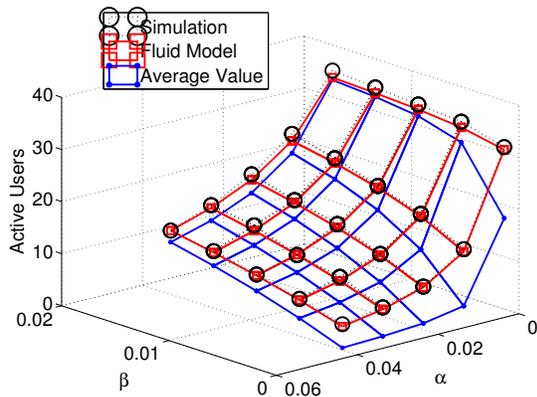


Fig. 1. Number of Active Users for $\lambda=0.2$, $1/\mu=180$

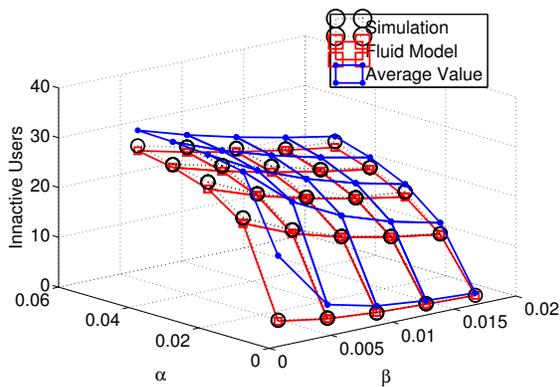


Fig. 2. Number of Inactive Users for $\lambda=0.2$, $1/\mu=180$

Now, we investigate the effect of the arrival rate in the number of users in the system. In Fig, 3 the number of active, inactive and total users is presented. Note that as λ increases, the increase on the number of active nodes is much higher than for the inactive nodes. This is because it is assumed that users enter the system as active

nodes. This greatly increases the traffic of the system. As such, it can be considered as the worst case compared to users entering the system as inactive users.

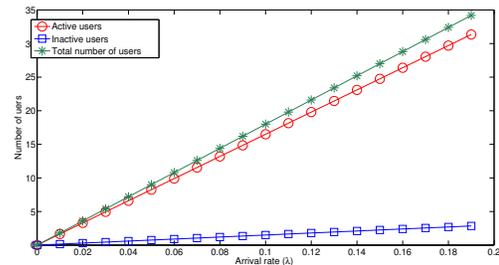


Fig. 3. Number of Active, Inactive and Total Users in the System for $1/\mu=180$

When users can perform a handoff procedure between the two access points, the system is described by (7) and (8). Since the fluid model has been validated in the previous results, only the fluid model results are presented. In Fig. ?? we present the total number of users for different values of α and γ .

In general, the total number of users in the system is insensitive to both parameters except for rather high values of γ where resources are constantly being released and requested, causing a higher resource occupancy and a non zero failed transfer probability. This effect has been thoroughly studied in such systems as cellular networks under the ping-pong effect. Indeed, as a mobile user leaves the current access point, packets queued in the buffer occupy this released resources. Hence, queued packets have priority over newly arrived users. This in turns increment the blocking and failed call transfer probabilities.

6.2 Basic Model

Now, the performance of the system with no handoff capabilities is studied in terms of packet delay and packet loss probability. From the system's administrator perspective, the main control variables are the number of available time slots or transmission opportunities (M) and the buffer size (Q). It can be seen that the system is highly sensitive to these main parameters as

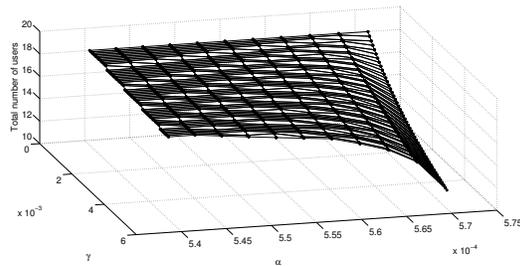


Fig. 4. Total Number of Users in the System for $\lambda = 0.2$, $1/\mu=180$, $\beta = 0.01$, with Handoff Capabilities Enabled

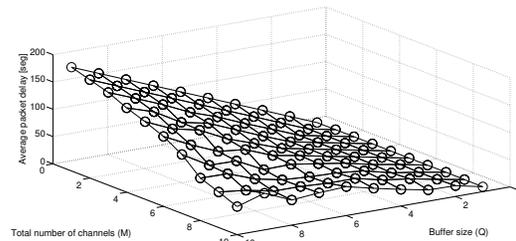


Fig. 5. Average Delay for $\lambda = 0.2$, $1/\mu=180$, $\alpha = 0.02$, $\beta = 0.01$

shown in Fig. 5. Recall that the number of channels in this system correspond to the number of nodes that transmit their data in the PCF mode, i.e, the number of nodes registered to the AP to transmit in the collision-free phase.

As expected, as the number of channels increases, the average packet delay decreases. This is because there is more bandwidth assigned to the VoIP users, reducing the time to transmit each packet. On the other hand, as the buffer size increases, the blocking probability and the packet loss probability decrease, allocating more users in the system.

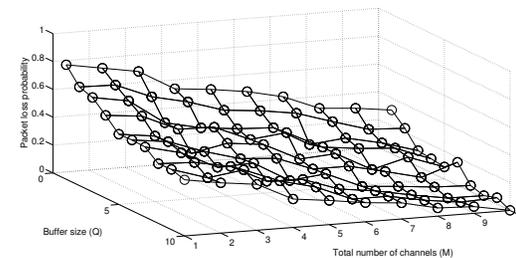


Fig. 6. Packet Loss Probability for $\lambda = 0.2$, $1/\mu=180$, $\alpha = 0.02$, $\beta = 0.01$

However when the VoIP packets do not find available resources they are queued in the buffer, increasing the packet delay. As such, in order to achieve an acceptable packet delay (less than 400 msec for the case of VoIP services) the buffer size has to be relatively small. For instance, for $M = 10$, the buffer size should not be higher than 2, while for lower values of M the buffer size should not be higher than 1. This is a relevant result since this system does not require to have a high capacity buffer, such as the ones used for non-real time services.

In Fig. 6 the packet loss probability is presented. In this system, a VoIP packet is lost when all the M resources are used and the buffer is full. As such, when the buffer size increases there are less packets lost at the cost of increasing the packet delay. The same effect occurs when increasing M . Nonetheless, the main performance parameter is packet delay and a high value of Q is not recommended.

6.3 Handoff Model

In the following results, the system with handoff capabilities enabled is studied. It can be seen in Fig. 7 that the general behavior of the average packet delay is the same as the system with no handoffs. However, for high values of M and Q the average packet delay is lower. This is because as users leave one AP in order to connect to the other AP, some users cannot perform the handoff procedure successfully as some of them experience a fail call transfer. As such, resources are more available and users that remain in the system until the end of the call have a lower packet delay.

Consider for instance the case where a mobile user with an ongoing call leaves the current access point in order to request a handoff to the other access point. In this case, when the user leaves his access point, a queued packet at the head of the buffer immediately occupies the released channel, reducing indeed the packet delay. The same effect occurs to the packet loss probability

as seen in Fig. 8. A lower packet loss probability is experienced at the expense of a higher failed call transfer. Again, whenever a user leaves the system, if there are packets in the buffer, one of those packets is served and at the same time a location in the buffer becomes available.

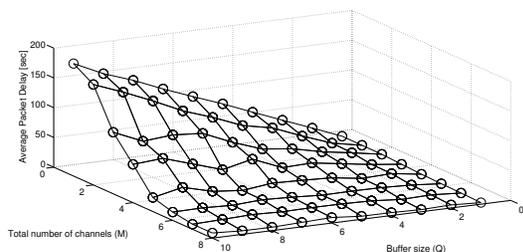


Fig. 7. Average Delay for $\lambda = 0.2$, $1/\mu=180$, $\alpha = 0.02$, $\beta = 0.01$, $\gamma = 0.1$ with Handoff Capabilities Enabled

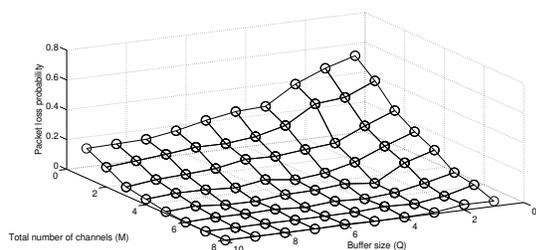


Fig. 8. Packet Loss Probability for $\lambda = 0.2$, $1/\mu=180$, $\alpha = 0.02$, $\beta = 0.01$, $\gamma = 0.1$ with Handoff Capabilities Enabled

7 Conclusion

A voice service under a TDMA-based system is studied with and without handoff capabilities. It is considered that due to the extensive use of the VoIP protocol in conjunction with the ubiquity of the WLAN and cellular infrastructure, VoIP services will be greatly used as an alternative to the conventional telephone services for both the PTSN and cellular systems.

A teletraffic analysis based on the arrival and departure of users as well as the active/inactive characteristic of packetized voice is performed.

Also, a simple fluid model is performed in order to calculate the number of active and inactive users in a VoIP system. The analysis provides close and simple expressions that closely matches simulation results.

This model is presented due to the fact that any VoIP system (wired or wireless) requires a first estimation on the number of active and inactive users for a given set of parameters (arrival rate, service time, activity factor of the ON/OFF traffic) that captures the dynamic nature of the system. The system performance is evaluated in terms of packet delay and packet loss probability considering a buffered network.

References

1. So, J. W. (2007). A downlink performance analysis of VoIP services over an IEEE 802.16e OFDMA system, *IEEE Commun. Letters*, Vol. 11, No. 2, pp. 155–157.
2. So, J. W. (2008). Performance analysis of VoIP services in the IEEE 802.16e OFDMA system with in band signaling. *IEEE Trans. Veh. Technol.*, Vol. 57, No. 3, pp. 1876–1886.
3. Lee, H. & Cho, D. H. (2009). Multi-user queuing analysis considering AMC for wireless VoIP services. *Proc. IEEE ICC Workshops*. DOI: 10.1109/ICCW.2009.5208097.
4. Lee, H. & Cho, D. H. (2009). VoIP capacity analysis in cognitive radio systems. *IEEE Commun. Letters*, Vol. 13, No. 6.
5. Lee, H. & Cho, D. H. (2010). Capacity improvement and analysis of VoIP service in a cognitive radio system. *IEEE Trans. Veh. Technol.*, Vol. 50, No. 4.
6. Heffes, H. & Lucantoni, D. M. (1986). A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J. Select. Areas Commun.*, Vol. 4, No. 6, pp. 856–868.
7. Sriram, K. & Whitt, W. (1986). Characterizing superposition arrival processes in packet multiplexers for voice and data. *IEEE J. Select. Areas Commun.*, Vol. 4, No. 6, pp. 833–846.
8. Brouzioutis, C., Vitsas, V., & Chatzimisios, P. (2010). Studying the impact of data traffic on voice capacity in IEEE 802.11 WLANs. *Communications (ICC), IEEE International Conference on*.

9. **Shawish, A., Jiang, X., & Horiguchi, S. (2008).** S-Voice capacity analysis and enhancement in wireless LAN. *Wireless Communications and Networking Conference, (WCNC)*.
10. **Malathi, V., Cocker, N., & Moors, T. (2001).** Support of voice services in IEEE 802.11 wireless LANs. *INFOCOM, Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 1*.
11. **Chen, D. & et al. (2002).** *Proceeding of OPNETWork*.
12. **VoIP Call Admission Control CISCO (2003).** Available online: http://www.cisco.com/en/US/docs/ios/solutions_docs/voip_solutions/CAC.html, pp. 1–50.
13. **Hoque, M. A. & Afroz, F. (2008).** Call admission control: QoS issue for VoIP. *Proceeding COMSWARE*, pp. 757–761.
14. **McGovern, P., Perry, P., Murphy, S., & Murphy, L. (2011).** Endpoint Based Call Admission Control and Resource Management for VoWLAN. *IEEE Trans. on Mobile Comp.*, Vol. 10, No. 5, pp. 684–699.
15. **Jianmin, Z., Zhaoyang, Z., Haiyan, L., Wei, W., & Guanding, Y. (2012).** Initial spectrum access control with QoS protection for active users in cognitive wireless networks. *Wiley International Journal of Communication Systems*, Vol. 25, No. 5, pp. 636–651.
16. **Qianxi, L., Tao, P., Wei, W., Wenbo, W., & Chao, H. (2012).** Utility-based resource allocation in uplink of OFDMA-based cognitive radio networks. *Wiley International Journal of Communication Systems*, Vol. 23, No. 2, pp. 252–274.
17. **Whai-En, C. & Pin-Jen, L. (2010).** A performance study for IPv4-IPv6 translation in IP multimedia core network subsystem. *Wiley International Journal of Communication Systems*, Vol. 23, No. 8, pp. 929–944.
18. **Nidal, N. & Sghaier, G. (2010).** Performance analysis of a cell-based call admission control scheme for QoS support in multimedia wireless networks. *Wiley International Journal of Communication Systems*, Vol. 23, No. 6.
19. **Qiu, D. & Srikant, R. (2004).** Modeling and performance analysis of BitTorrent-like peer-to-peer networks. *Proc. ACM SIGCOMM*.
20. **Lee, H. & Cho, D. H. (2009).** VoIP capacity analysis in cognitive radio system. *IEEE Communications Letters*, Vol. 13, No. 6, pp. 393–395.
21. **Niyato, D. & Hossain, E. (2007).** A novel analytical framework for integrated cross-layer study of call-level and packet-level QoS in wireless mobile multimedia networks. *IEEE Trans. Mobile Comp.*, Vol. 6, No. 3, pp. 322–335.
22. **Brady, P. T. (1969).** A model for generating ON-OFF speech patterns in two-way conversations. *Bell Syst. Tech. Journal*, Vol. 48, No. 7, pp. 2445–2472.
23. **Bi, Q., Chen, P. C., Yang, Y., & Zhang, Q. (2006).** An analysis of VoIP service using 1xEVDO revision A system. *IEEE J. Select. Areas Commun.*, Vol. 24, No. 1, pp. 36–45.
24. **Bi, Q., Vitebsky, S., Yang, Y., & Zhang, Q. (2008).** Performance and capacity of cellular OFDMA systems with voice-over-IP traffic. *IEEE Trans. Veh. Technol.*, Vol. 57, No. 6, pp. 3641–3652.

Article received on 03/04/2017; accepted on 09/05/2018.
Corresponding author is Mario E. Rivero-Angeles.