Scientific
Research
Publishing

A Priority Queuing Model for HCF Controlled Channel Access (HCCA) in Wireless LANs

Reza GHAZIZADEH[1], Pingzhi FAN[1], Yi PAN[2]

[1]*Keylab of Information Coding and Transmission, Institute of Mobile Communications,*
Southwest Jiaotong University, Chengdu, China
[2]*Department of Computer Science, Georgia State University, Atlanta, GA, USA*
Emails: r_ghazizadeh@yahoo.com, p.fan@ieee.org, pan@cs.gsu.edu

Abstract

Recently, there has been a rapid growing interest in new applications requiring quality of service (QoS) guarantees through wireless local area networks (WLAN). These demands have led to the introduction of new 802.11 standard series to enhance access medium supporting QoS for multimedia applications. However, some applications such as variable bit rate (VBR) traffic address some challenges in the hybrid coordination function (HCF) nominated to provide QoS. This paper presents a novel priority queuing model to analyze a medium access in the HCF controlled channel access (HCCA) mode. This model makes use of a MAP (Markovian Arrival Process)/PH (Phase Type)/1 queue with two types of jobs which are suitable to support VBR traffic. Using a MAP for traffic arrival process and PH distribution for service process, the inclusion of vacation period makes our analysis very general and comprehensive to support various types of practical traffic streams. The proposed priority queuing model is very useful to evaluate and enhance the performance of the scheduler and the admission controller in the HCCA mechanism.

1. Introduction

Increasing demands to access to network in hotspots areas at airports, hotels, coffee shops have led wireless local area network (WLAN) to be a key technology for high-speed local access in public and private areas. Furthermore, in the near future, WLAN will play a key role within the hybrid wireless systems and also it is the best candidate to connect home devices to wireless networks. Therefore, it should be able to allow users to ubiquitously access a large variety of services. On the other hand, demands for new applications such as real time traffic, multimedia video and voice over IP are increasing rapidly. These applications have created need for QoS support.

However, IEEE802.11 is unsuitable for multimedia applications to support QoS in the MAC layer. Therefore, IEEE802.11 working group has been developing a new protocol, IEEE802.11e, which will be able to provide QoS features. IEEE802.11e introduces the hybrid co-ordination function comprised of two medium access mechanisms: contention-based channel access referred to as *enhanced distributed channel access* (EDCA) and controlled channel access referred to as *HCF controlled channel access* (HCCA) [1].

Although contention-based channel access is very simple and robust for best effort traffic, it can not provide QoS guarantees easily. These can be achieved with the polling-based medium access through the HCCA. The HCCA provides a *hybrid coordinator* (HC) with ability to assign a contention free time interval during contention period and contention free period to packet transmission. Therefore, *transmission opportunity* (TXOP) and service interval (SI) are very important parameters to provide QoS guarantees. A reference scheduler calculates these parameters with the reservation information achieved through the negotiation with the end users. Using average values, such as average packet length and average data rate, to compute transmission parameters cause some challenges to QoS support in VBR traffic. Therefore, modification of the scheduler to provide such traffic is very crucial. In this way, analytical system analysis is very useful to improve and develop the system.

This paper presents a priority queuing model to analyze a medium access in the HCCA. According to the HCCA characteristics, the model is based on a MAP/PH/1 queue with vacation and time-limited service in the presence of two priority levels. Some important performance measures are presented after solving the queuing model through the matrix-analytic method [2,3]. To our best knowledge, although there are some papers which have investigated the HCCA and the EDCA through simulation [4−9], and the EDCA analytically [10−12], there is only one published analytical work for the HCCA [13], where a queuing model without priority levels has been developed. However, priority is a key point to separate various traffic streams with different QoS requirements, especially for real-time and non-real-time traffic. In addition, to manage resources, use bandwidth efficiently and provide QoS for various traffic streams, it is necessary that different traffic streams appear in the different statistic or dynamic priority levels in the system. Therefore, IEEE802.11e introduces four priority levels for eight groups of traffic streams. Consequently, to modify and investigate the performance of the scheduler and the admission controller, priority analytical model in the presence of prioritized traffic is very helpful. In [14], a priority model in a medium without a vacation period and time-limited service was introduced. That model can not be applied to the HCCA medium access which is based on a vacation period and time-limited service. This paper will present a priority queuing model to analyze medium access in the HCCA by making use of an MAP/PH/1 queue with two types of jobs which are suitable to support vast practical traffic streams.

The rest of the paper is organized as follows: in Section 2, we briefly describe the HCCA mechanism, phase type (PH) distribution and discrete Markovian arrival process (DMAP); our proposed model is presented in Section 3; the related performance measures are analyzed in Section 4; numerical and simulation results are given in Section 5; and finally conclusions are drawn in Section 6.

2. HCCA and System Parameters

2.1. HCCA

IEEE802.11e/HCCA is a polling-based medium and centralized scheduling which is controlled by the HC. Each station that requires a strict QoS support is allowed to send QoS requirement packets to the HC and the HC assign a corresponding transmission opportunity to the station. The HC can start a polling period at any time during a contention period after the medium remains idle for at least *point coordination function* (PCF) inter-frame space interval. Each station can transmit a sequences of data packets separated by a short inter-frame space during own TXOP allocated by the HC in a contention free period. Therefore, as it has shown in Figure 1, a sequence of transmission opportunities will be assigned

to the stations during each SI. Consequently, each station is polled once per SI and allowed to transmit its packets until its TXOP duration elapses. Uplink and downlink TXOPs are initiated by the scheduler in the HC and end when there is no packet in the queues for transmission or TXOP duration expires. To provide QoS, each station manages QoS control field added to the legacy frames. Consequently, the scheduler receives separated reservation information of different traffic streams to calculate an aggregated service schedule. Some of this information is the mean data rate, delay bound, maximum burst size, minimum physical rate, user priority and peak data rate. The scheduler, first of all calculates the maximum service interval according to the delay bound for each traffic stream. Then, it selects the smallest service interval among all the maximum service intervals corresponding with the traffic streams as a service interval for all stations. The scheduler after that determines TXOP for each traffic stream according to the negotiated reservation information. Allocated TXOP for each station is sum of all TXOPs of station's traffic streams. TXOP of jth station that has made *n* traffic streams is computed as follows.

$$TXOP_j = \sum_{i=1}^{n} \max\left(\frac{\left\lceil \frac{SI \times \rho_i}{L_i} \right\rceil \times L_i}{R_i} + o, \frac{M_i}{R_i} + o\right) \qquad \textbf{(1)}$$

where R denotes the minimum physical data rate, L and M represents the nominal and the maximum size of packet respectively, ρ denotes the mean data rate and O represents the overhead due to the physical and MAC headers, acknowledgment and polling frames. According to the service interval duration, the number of active stations and the TXOP duration in each station, an admission controller manages the number of active stations to provide QoS.

2.2. Discrete Markovian Arrival Process (DMAP)

Consider a discrete time Markov chain with a transition matrix D and two sub-stochastic matrix D_0, D_1, where $D = D_0 + D_1$. Suppose at the time t, the Markov chain is in the state i, $1 \le i \le n$. Then, at the time epoch t+1 with probability $(D_k)_{ij}$, $k = 0,1$ arrival process enters state j, $1 \le j \le n$ and starts a batch of k arrivals. Therefore, D_0 corresponds to a transition matrix with no arrival and D_1 corresponds to a transition matrix with only one arrival per time slot.
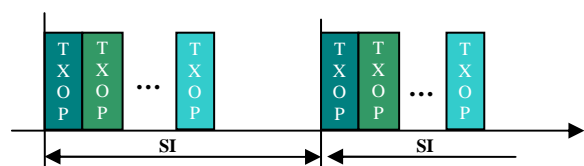

Figure 1. TXOP allocation in contention free period.

2.3. Phase Type (PH) Distribution

Phase type distributions can approximate many of general distributions encountered in queuing systems. Therefore, it is appropriate in service time modeling [2]. Consider an m+1 states Markov chain with one absorbing state, and initial probability vector $[\beta, \beta_{m+1}]$ with m+1 elements where β is a row vector with m element while $\beta e + \beta_{m+1} = 1$ (e is column vector of one). Let the transition probability matrix of the Markov chain be P: $P = \begin{bmatrix} S & S^0 \\ 0 & 1 \end{bmatrix}$.

where S is a sub-stochastic transition matrix, I-S is nonsingular (I is identity matrix) and $S^0 = e - Se$. The absorbing Markov chain can initialize from any states according to the initial vector and gets absorbed to the absorbing state. Therefore, the time to absorption in such a Markov chain is said to have phase type distribution which is represented with the pair (β, S).

3. Proposed Priority Queuing Model

To clarify the model, suppose that there is one station communicating with the HC in the system. The HC can be considered as a single server which serves queues (high and low priority) of the station no more than T slots (maximum TXOP duration is divided to T slots) during each SI. In the view of the station, as soon as T time slots is used up or queues become empty, the server goes on a vacation (i.e. the server serves other stations or becomes idle until the next visit). Hence, as it is illustrated in Figure 2, the minimum vacation duration is subtraction of the SI and the maximum TXOP duration. It is assumed that the HC is a server which serves priority queues in non-preemptive priority case during a TXOP period. In a non-preemptive case, no service interruption is applied upon arrival of a high priority packet when a low priority packet is being served. To analyze the discrete time Markov chain (DTMC) describing the queuing model, arrival process, service process and vacation model are defined. The arrival process is modeled by a discrete Markovian arrival process (DMAP) to allow correlation among the inter-arrival times within packets (within each priority and between two priorities packets) and support various types of traffic streams, especially VBR traffic which generate packets in random arrival intervals. On the other
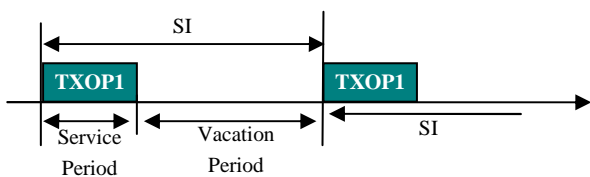
Figure 2. Service and vacation periods for one station in HCCA.

hand, it is obvious that the packet transmission time is corresponding to the service time which depends on the packet length while the channel data rate is fixed. Therefore, to support various packet length distributions and make the model of service process more general and comprehensive, a phase type (PH) distribution is proposed for a service process model. Consequently, the introduced priority queuing model is based on a MAP/PH/1 queue with vacation and time-limited service. The proposed model is based on the work of [14] which makes use of matrix-geometric solution for analysis priority queues without vacation and time limitation in service.

Some of the notations and symbols which will be used throughout the paper are introduced as follows: e is a column vector of one (with appropriate order equals to the number of columns of the matrix or to the vector length that it is multiplied with), $e_T(v)$ represents a column vector of zeroes with T length except at the vth position that is one, e' is the transpose of e vector, I_k denotes an identity matrix of dimension $k \times k$ and H_k represents $H_k = \begin{bmatrix} 0 & I_{k-1} \\ 0 & 0 \end{bmatrix}$.

3.1. Arrival Process

The arrival process is modeled by a discrete Markovian arrival process (DMAP). The DMAP, an extension of the Markov modulated Bernoulli process, can support many types of traffic flows such as VBR traffic generating variable packets length in variable inter-arrival periods. Suppose there are two independent types of traffic corresponding to two priorities where each traffic source is able to generate only one packet per time slot. Hence, each traffic flow will have two sub-stochastic matrices (Section 2.2) and consequently there are four sub-stochastic matrixes (D_{00}, D_{11}, D_{12}, D_2) corresponding to join packet arrival (high priority and low priority packet arrival) where D_{00} denotes a transition matrix with no packet arrival, D_{11} is a transition matrix when one high priority packet arrives, D_{12} represents a transition matrix when one low priority packet arrives, D_2 is a transition matrix when two high and low priority packet arrive(one of each priority packet) and also $D = D_{00} + D_{11} + D_{12} + D_2$ where D represents stochastic matrix. The arrival rate is $\lambda_i = \psi(D_{1i} + D_2)$ $i = 1, 2$ where $\psi = \psi e$ and $\psi e = 1$ (e is a column vector of one).

As mentioned above, four sub-stochastic matrices can be expressed by the sub-stochastic matrices of both traffic streams. Suppose $D_0(1)$ and $D_1(1)$ are the transition matrices in the high priority traffic when no packet and one packet arrives at a time slot, respectively. Furthermore assume, $D_0(2)$ and $D_1(2)$ are the transition matrices in the low priority traffic when no packet and one packet arrives at the time slot respectively.

In this case, $D_{00} = D_0(1) \otimes D_0(2)$, $D_{11} = D_1(1) \otimes D_0(2)$, $D_{12} = D_0(1) \otimes D_1(2)$ and $D_2 = D_1(1) \otimes D_1(2)$, where $\otimes$ is the Kronecker product.

3.2. Service Process

The service process is corresponding to the transmission time. The total transmission time of a frame is sum of transmission time of data packet, its necessary headers added by the MAC and physical layer, ACK, and short inter-frame space (SIFS). We assume that the channel data rate, ACK, SIFS durations and header size are fixed. Hence, the service time of a packet can be considered as a random variable which varies only with the length of the packet. Consequently, to generalize the model and support different packet length distributions, we consider phase type distribution for both high priority and low priority service processes. Let (β_1, S_1) and (β_2, S_2) denote PH distribution for high priority and low priority services, respectively where S_1, S_2 are transition matrices of dimensions $m_1 \times m_1$, $m_2 \times m_2$, respectively and β_1, β_2 represent initial vectors. $S_1^0 = e - S_1 e$, $S_2^0 = e - S_2 e$ are transition to absorption vectors for the high priority and low priority services, respectively.

3.3. Vacation Model

In the service period, whenever there is no packet in the queues or the TXOP duration expires, the server enters a vacation period. Therefore, the vacation duration depends on the service duration. A vacation with the maximum duration begins whenever the server visits the station at the first slot of TXOP and the station has no packet to transmit. Consequently, a vacation model can be represented by a (δ_i, V) phase type distribution while the Markov chain can initialize from any states depending on the vacation duration. Therefore, the initial vector, the transition matrix and the transition to absorption vector will be $\delta_i = e_{SI}(i)$ $1 \le i \le T + 1$, $V = H_{SI}$ and $V^0 = e - V e$, respectively.

3.4. State Space and Transition Matrix of the DTMC

In this subsection, we introduce state space and the transition matrix of the discrete time Markov chain (DTMC). The state space can be divided into two main groups that are vacation and service state spaces. Each of these states are described by the number of the packets in the high priority and low priority queues (i_1, i_2), the phase of the arrival process (j), the phase of the high priority or low priority service processes (k_1, k_2), the phase of the vacation (l) and the phase of the TXOP (t). Therefore, the states can be expressed as follows.

$$d^v = \{(i_1, i_2, j, l), \ 0 \le i_1 \le Q \ ; 0 \le i_2 \le Q \ ; j$$
$$= 1, 2, \ldots, n; \ l = 1, \ldots, r\}$$

$$d_0^s = \{(0, i_2, j, k_2, t), \ 1 \le i_2 \le Q \ ; \ j = 1, 2, \ldots, n; \ t$$
$$= 1, \ldots, T; \ k_2 = 1, 2, \ldots, m_2\}$$

$$d_1^s = \{(i_1, i_2, j, k_2, t), 1 \le i_1 \le Q \ ; 1 \le i_2 \le Q \ ; j = 1, 2, \ldots, n; t$$
$$= 1, \ldots, T; \ k_2 = 1, 2, \ldots, m_2\} \qquad \textbf{(2)}$$

$$d_2^s = \{(i_1, i_2, j, k_1, t), 1 \le i_1 \le Q \ ; 0 \le i_2 \le Q \ ; j = 1, 2, \ldots, n; t$$
$$= 1, \ldots, T; \ k_1 = 1, 2, \ldots, m_1\}$$

where Q is the buffer size in the number of packets (high priority and low priority), d^v denotes the vacation states while the number of high priority and low priority packets in the queues are i_1 and i_2 respectively and the packet arrival is in j th phase as well as vacation is in l th phase, d_0^s represents service state space when there are only low priority packets in the system. Therefore, a low priority packet is being served while the service is in the phase K_2 at t th time slot and the packet arrival is in j th phase as well. In the same way, d_1^s is service state space when there is at least one high priority packet in the system and a low priority packet is being served, and d_2^s represents service state space when there is at least one high priority packet in the system and a high priority packet is being served.

The transition matrix of the discrete time Markov chain can be expressed as follows.

$$\begin{bmatrix} \text{probability of remaining in vacation} & \text{probability of switching from vacation to low priority service} & \text{probability of switching from vacation to high priority service} \\ \text{probability of switching from low priority service to vacation} & \text{probability of remaining in low priority service} & \text{probability of switching from low priority service to high priority service} \\ \text{probability of switching from high priority service to vacation} & \text{probability of switching from high priority service to low priority service} & \text{probability of remaining in high priority service} \end{bmatrix}$$

Figure 3. General form of the transition probability sub matrix.

$$P = \begin{bmatrix} B_{00} & B_{01} \\ B_{10} & A_0 & A_1 \\ & A_{-1} & A_0 & A_1 \\ & & \ddots & \ddots & \ddots \\ & & & A_{-1} & A_0 & A_1 \\ & & & & A_{-1} & A_0+A_1 \end{bmatrix} \quad \textbf{(3)}$$

where the rows of P matrix correspond to the number of packets in the high priority queue. Therefore, matrix will have $Q+1$ rows. As it is assumed that each type of traffic can generate only one packet and only one packet can be served per time slot (less than or equal to one), the structure of P matrix is quasi-birth-death. Consequently, the elements of matrix P represent block transition matrices in which the number of packets in the high priority queue increases (B_{01}, A_1) or decreases (B_{10}, A_{-1}) by one, or remains invariant (B_{00}, A_0) after transition at the current time slot. On the other hand, each element of P matrix also represents one matrix describing low priority queue size. Therefore, each row of B_{00}, B_{01}, B_{10}, A_{-1}, A_0, A_1 block matrices, described as follows, represents the number of low priority packets in the queue.

To describe each sub matrix (with general form B_a^b or A_a^b), a general sub matrix form is defined in Figure 3. The general sub matrix can be understood as the transitions probability matrix governing switch among a vacation, a high priority and a low priority service. Note that the maximum service duration is T slots and transition can happen at any time slots. Therefore, service period in the general matrix is divided to T slot (in high priority and low priority). It is obvious that some of the state transitions in the general matrix may not happen. Therefore, those states will be zero. To reduction of the matrix dimensions, those rows and columns of the general matrix which are zero will be removed if the general matrix can match with the other matrices in the P matrix. Now, in the rest of this sub section, we express the sub matrices by considering the possible state transitions in the general matrix form.

$$B_{00} = \begin{bmatrix} B_{00}^{00} & B_{00}^{01} \\ B_{00}^{10} & B_{00}^0 & B_{00}^1 \\ & B_{00}^{-1} & B_{00}^0 & B_{00}^1 \\ & & \ddots & \ddots & \ddots \\ & & & B_{00}^{-1} & B_{00}^0 & B_{00}^1 \\ & & & & B_{00}^{-1} & B_{00}^0+B_{00}^1 \end{bmatrix} \quad \textbf{(4)}$$

Block matrix B_{00}^{00} represents state transitions when high priority and low priority queues are empty and remain empty after transition. Transitions occur whenever no packet arrives (D_{00}) and the server is on vacation (V), or completes vacation (V^0) and starts it again. As there is no packet to be served, the maximum vacation duration will be initialized (δ_1).

$$B_{00}^{00} = D_{00} \otimes (V + V^0\delta_1) \quad \textbf{(5)}$$

B_{00}^{01} matrix denotes transitions in which the number of packets in the low priority queue increases by one while the both queues are empty. Transitions occur whenever only a low priority packet arrives (D_{12}) and the server stays on vacation (V) or completes vacation (V^0) and starts the low priority service (β_2) at the first slot of the TXOP.

$$B_{00}^{01} = \begin{bmatrix} D_{12} \otimes V & e_1'(T) \otimes D_{12} \otimes V^0\beta_2 \end{bmatrix} \quad \textbf{(6)}$$

It is assumed that whenever the both queues become empty the server goes on vacation. The vacation can begin from different states of its Markov chain which is dependent on the instant that the queues become empty in the service period. B_{00}^{10} supports state transitions that no packet arrives (D_{00}), the process of the last low priority packet is completed. Consequently, queue become empty and the vacation period begins (transitions can happen at any time slots in the TXOP period) ($S_2^0\delta_i \; i=1,..,T$).

$$B_{00}^{10} = \begin{bmatrix} 0 \\ D_{00} \otimes S_2^0\delta_1 \\ \vdots \\ D_{00} \otimes S_2^0\delta_T \end{bmatrix} \quad \textbf{(7)}$$

$B_{00}^0, B_{00}^1, B_{00}^{-1}$ block matrices represent transitions in which the number of packets in the low priority queue remains invariant, increases, decreases by one respectively while the high priority queue remains empty after transition, and there is at least one packet in the low priority queue before transition. These conditions can happen on the vacation, or in the service (low priority queue is served).

Now, we explain possible state transitions in the block matrix B_{00}^0. We divide transitions into two cases. 1) no packet arrives (D_{00}), and a) the server remains on vacation (V), b) the server ends vacation and goes on the low priority service ($V^0\beta_2$) at the first slot of TXOP (i.e. $e_1'(T) \otimes D_{00} \otimes V^0\beta_2$). c) the server remains on the processing of low priority packet (S_2), d) the server leaves the service processing due to the TXOP expiration ($e_T(T)$) and enters vacation ($e_{m_2}\delta$). 2) a low priority packet arrives (D_{12}), and a) the processing of a low priority packet is completed and a new low priority processing begins ($S_2^0\beta_2$), b) the processing of a low priority packet is completed while the TXOP duration expires as well and the server goes on vacation ($S_2^0\delta$)

$$B_{00}^0 = \begin{bmatrix} D_{00} \otimes V & e_1'(T) \otimes D_{00} \otimes V^0\beta_2 \\ B_{00}^{0Lv} & H_T \otimes [D_{00} \otimes S_2 + D_{12} \otimes S_2^0\beta_2] \end{bmatrix} \quad \textbf{(8)}$$

$$B_{00}^{0Lv} = e_T(T) \otimes [D_{00} \otimes e_{m_2}\delta + D_{12} \otimes S_2^0\delta]$$

I. J. Communications, Network and System Sciences, 2009, 1, 1-89

B_{00}^1 contains the state transitions increasing the number of packets in the low priority queue by one and remaining the high priority queue empty while there is at least one packet in the low priority queue before transition. One can easily compute the possible transitions such above discussions

$$B_{00}^1 = \begin{bmatrix} D_{12} \otimes V & e_1'(T) \otimes D_{12} \otimes V^0 \beta_2 \\ e_T(T) \otimes D_{12} \otimes e_{m_2} \delta & H_T \otimes D_{12} \otimes S_2 \end{bmatrix} \quad (9)$$

B_{00}^{-1} represents state transitions in which the number of packets in the low priority queue decreases by one and at least one packet is in the low priority queue before transition as well as high priority queue remains empty. These transitions happen when no packet arrives and the processing of a low priority packet is completed, and a) the processing of another one begins, b) the TXOP duration expires and a vacation period begins.

$$B_{00}^{-1} = \begin{bmatrix} 0 & 0 \\ e_T(T) \otimes [D_{00} \otimes S_2^0 \delta] & H_T \otimes D_{00} \otimes S_2^0 \beta_2 \end{bmatrix} \quad (10)$$

$B_{01}, B_{10}, A_{-1}, A_0, A_1$ and their elements can be computed in the same manner. The block matrices are given in Appendix A.

4. Performance Measures

According to the structure of P matrix, its steady state distribution vector can be obtained by applying the matrix-geometric method. Let probability steady state distribution vector be

$x = [x_0 \ x_1 \ x_2 \ ... \ x_Q]$ where $x_i = [x_{i0} \ x_{i1} \ x_{i2} \ ... \ x_{iQ}]$,

$x_{ij} = [x_{ij}^v \ x_{ij}(1) \ x_{ij}(2)]$ and $x_{ij}(k) = [x_{ij}^1(k) \ x_{ij}^2(k) \ ... \ x_{ij}^T(k)]$

$0 \leq i \leq Q, 0 \leq j \leq Q, k = 1,2$ where $x_{ij}^n(k)$ is the probability

that the number of packets in the high priority and in the low priority queues are i and j respectively while type k packet ($k = 1$: high priority, $k = 2$: low priority) is being served at the n th time slot of the TXOP period.

Using balanced equations ($x = xP$, $xe = 1$) and the matrix-geometric method, the steady state vector x can be calculated. For more details of how to find out steady state vector, readers can refer to [2].

4.1. Queue Length Distribution

Let $f_h(l)$ ($f_l(l)$) be the probability that there are l high priority packets (low priority packets) in the queue. The length of the high priority queue will be l if there are l high priority packets in the system and the server is not on the processing of the high priority packet (i.e. is on vacation or in the processing of the low priority packet) or, $l+1$ high priority packet are in the system while one high priority packet is being served.

$$f_h(L=l) = \begin{cases} x_{00} + \sum_{j=1}^{Q} x_{0j}^v + \sum_{j=1}^{Q} x_{0j}(2) + \sum_{j=0}^{Q} x_{1j}(1) & l=0 \\ \sum_{j=0}^{Q} x_{lj}^v + \sum_{j=0}^{Q} x_{lj}(2) + \sum_{j=0}^{Q} x_{l+1j}(1) & 1 \leq l \leq Q \end{cases} \quad (11)$$

$$f_l(L=l) = \begin{cases} x_{00} + \sum_{i=1}^{Q} x_{i0} + x_{01}(2) & l=0 \\ \sum_{j=1}^{Q} x_{il} + x_{0l}^v + x_{0l+1}(2) & 1 \leq l \leq Q \end{cases} \quad (12)$$

Probability of the queue length at the end of the TXOP duration can be calculated in the similar manner.

4.2. Packet Loss Rate

Packet loss occurs whenever a new packet arrives and the target buffer is full. These conditions can happen during service processing (at any time slots of the TXOP duration) or vacation.

The high priority packet will be lost when the number of packets in the high priority queue is Q (regardless of the number of packets in the low priority queue) and a new high priority packet arrives (by itself or together with a low priority packet) ($D_{11} + D_2$). Therefore, the packet loss probability will be sum of all possible probabilities among vacation and service period satisfying above conditions. As an example, $\lambda_1^{-1} \sum_{j=0}^{Q} \sum_{k=1}^{T-1} x_{Qj}^k(1)[(D_{11} + D_2) \otimes (S_1)]$ shows sum of the

probabilities in which the server stays on the high priority processing while a new high priority packet arrives and the other mentioned conditions has been satisfied. Consequently the high priority Packet loss rate which is normalized with the high priority arrival rate (λ_1) is expressed as follow.

$$\begin{aligned} P_{Lh} = & \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{Q_1 j}^v [(D_{11} + D_2) \otimes (V + V^0 \beta_1)] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} \sum_{k=1}^{T-1} x_{Q_1 j}^k(1)[(D_{11} + D_2) \otimes (S_1)] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{Q_1 j}^T(1)[(D_{11} + D_2) \otimes e_{m_1} \delta] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} \sum_{k=1}^{T-1} x_{Q_1 j}^k(2)[(D_{11} + D_2) \otimes (S_2)] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{Q_1 j}^T(2)[(D_{11} + D_2) \otimes e_{m_2} \delta] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} \sum_{k=1}^{T-1} x_{Q_1 j}^k(2)[(D_{11} + D_2) \otimes (S_2^0 \beta_1)] \\ & + \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{Q_1 j}^T(2)[(D_{11} + D_2) \otimes S_2^0 \delta] \end{aligned} \quad (13)$$

At the same way, Packet loss in the low priority queue occurs when the low priority queue is full and a new low priority packet arrives ($D_{12} + D_2$). Therefore all possible states satisfying those conditions could be expressed as follows.

$$
\begin{aligned}
P_{b2} = & \lambda_2^{-1} x_{0Q_2}^v [(D_{12} + D_2) \otimes (V + V^0 \beta_2)] \\
& + \lambda_2^{-1} \sum_{i=1}^{Q_1} x_{iQ_2}^v [(D_{12} + D_2) \otimes (V + V^0 \beta_1)] \\
& + \lambda_2^{-1} \sum_{i=1}^{Q_1} \sum_{k=1}^{T-1} x_{iQ_2}^k (1)[(D_{12} + D_2) \otimes (S_1)] \\
& + \lambda_2^{-1} \sum_{i=1}^{Q_1} x_{iQ_2}^T (1)[(D_{12} + D_2) \otimes e_{m_1} \delta] \\
& + \lambda_2^{-1} \sum_{i=0}^{Q_1} \sum_{k=1}^{T-1} x_{iQ_2}^k (2)[(D_{12} + D_2) \otimes (S_2)] \\
& + \lambda_2^{-1} \sum_{i=0}^{Q_1} x_{iQ_2}^T (2)[(D_{12} + D_2) \otimes e_{m_2} \delta] \\
& + \lambda_2^{-1} \sum_{i=1}^{Q_1} \sum_{k=1}^{T-1} x_{iQ_2}^k (1)[(D_{12} + D_2) \otimes (S_1^0 \beta_1)] \\
& + \lambda_2^{-1} \sum_{i=1}^{Q_1} x_{iQ_2}^T (1)[(D_{12} + D_2) \otimes S_1^0 \delta]
\end{aligned} \tag{14}
$$

4.3. Access Delay Distribution

In this subsection, we introduce access delay distribution for the high priority and the low priority packets. Access delay is the required time in which an arriving packet at the target queue reaches the head of queue. Access delay can be studied as an absorbing Markov chain. The chain initializes when the packet arrives the queue, and gets absorbed when the packet reaches the head of the queue. Therefore, the access delay is the required time to absorption in the Markov chain.

In the high priority queue, experienced delay is the period of the time in which an arriving packet has to wait until all high priority packets ahead of it are served, and the process of a low priority packet, which is being processed at the arrival time, is completed Therefore, the access delay in the high priority queue depends on the number of the high priority packets ahead of an arriving packet. Suppose z defines the initial probability vector in the high priority access delay.

$$
z = [z_0\, z_1 \ldots z_{Q_1}] \quad z_i = [z_i^v\, z_{iL}^1 \ldots z_{iL}^T\, z_{iH}^1 \ldots z_{iH}^T] \tag{15}
$$

where z_i^v, z_{iL}^k and z_{iH}^k ($1 \le k \le T$) denote the probability of the arriving high priority packet finding i high priority packet ahead of it with the server: in vacation, in the low priority processing at the k th slot of the TXOP and, in the high priority processing at the k th slot of the TXOP, respectively.

Probability vector z_0^v represents the probability of arriving high priority packet ($D_{12} + D_2$) (regardless of low priority packet arrival), finding no high priority packet ahead of it with server: on vacation. The possible

scenarios are a) there is no packet in the high priority queue and the server stays on vacation ($V, V^0 \delta$), b) the server completes the processing of the last high priority packet in the Tth slot (the last slot of TXOP) and goes on vacation, c) the server leaves the low priority processing due to the TXOP expiration and goes on vacation, d) the server completes the low priority processing at the Tth slot and goes on vacation. Consequently, z_0^v which is normalized by λ_1 can be expressed as follows.

$$
\begin{aligned}
z_0^v = & \lambda_1^{-1} x_{00} (D_{11} + D_2) \otimes (V + V^0 \delta) \\
& + \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{0j}^v (D_{11} + D_2) \otimes V \\
& + \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{1j}^T (1)(D_{11} + D_2) \otimes S_1^0 \delta \\
& + \lambda_1^{-1} \sum_{j=1}^{Q_2} x_{0j}^T (2)(D_{11} + D_2) \otimes (S_2^0 \delta + e_{m_2} \delta)
\end{aligned} \tag{16}
$$

z_{0L}^k represents the probability vector in which arriving high priority packet finding no high priority packet ahead of it while server is serving a low priority packet at the kth slot of the TXOP. It can occur when the server stays on the processing of a low priority packet at the kth slot of the TXOP while high priority packet arrives.

$$
z_{0L}^k = \lambda_1^{-1} \sum_{j=1}^{Q} x_{0j}^{k-1} (2)(D_{11} + D_2) \otimes S_2 \quad 2 \le k \le T \tag{17}
$$

The other elements which are given in Appendix B can be calculated using similar above discussions Now, to find the time till absorption in a Markov chain, the transition matrix for high priority packet access delay (P_h) is required. This matrix is defined as follows.

$$
P_h =
\begin{bmatrix}
\tilde{B}_{00} & & & & & \\
\tilde{B}_{10} & \tilde{A}_0 & & & & \\
& \tilde{A}_{-1} & \tilde{A}_0 & & & \\
& & \ddots & \ddots & & \\
& & & \tilde{A}_{-1} & \tilde{A}_0 & \\
& & & & \tilde{A}_{-1} & \tilde{A}_0
\end{bmatrix} \tag{18}
$$

It is obvious that the access delay for an arriving high priority packet only depends on the number of high priority packets ahead of the arriving packet. The number of packets which arrives after desired packet has no effect on the access delay. Therefore, the arrival transition matrix will be I. Now, each element of P_h matrix can be computed with the similar discussions in the Section 3.4. For example, $\tilde{B}_{10}$ represent state transitions when the number of packets ahead of arriving packet changes from 1 to 0 at the end of transition. It is obvious that transitions occur when the high priority processing is completed.

$$
\tilde{B}_{10} =
\begin{bmatrix}
0 & 0 & 0 \\
e_T(T) \otimes I \otimes S_1^0 \delta & 0 & H_T \otimes I \otimes S_1^0
\end{bmatrix} \tag{19}
$$

The other elements are given in Appendix B. Finally, the probability vector after elapsing n+1 time slot will be

$$z^{n+1} = z^n P_h \qquad (20)$$

where $z^0 = z$. To reduction of computation, the set of following equation can be used.

$$\begin{cases} z_0^{n+1} = z_0^n \tilde{B}_{00} + z_1^n \tilde{B}_{10} \\ z_i^{n+1} = z_i^n \tilde{A}_0 + z_{i+1}^n \tilde{A}_{-1} \end{cases} \qquad (21)$$

Finally, let W_{HT} be the probability that the waiting time of high priority packet is less than or equal to T, then

$$W_{HT} = z_0^T P_h \qquad (22)$$

where z_0^T is the probability vector that the arrived high priority packet finding no packet a head of it after T slot.

The low priority access delay is calculated in Appendix C.

5. Numerical and Simulation Results

In this section, first we provide a simple example of wireless multimedia communications to demonstrate how can apply the computational algorithm.

It is assumed that the wireless network can transmit a fixed size data block during one time slot, and each packet is segmented into a number of data blocks. Suppose a station can transmit voice and video traffic. Furthermore, the priority of voice traffic is higher than the video traffic. Voice traffic is modeled by an ON/OFF source as depicted in Figure 4a. Therefore, $D_0(1)$ and $D_1(1)$ can be calculated as follows [15].

$$D_1(1) = R * D \qquad , \qquad D_0(1) = (I - R) * D$$

$$D = \begin{bmatrix} 1-\gamma_1 & \gamma_1 \\ \mu_1 & 1-\mu_1 \end{bmatrix} \qquad R = \begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix}$$

where A is the probability of the packet arrival per time slot. Now assume that the voice packet length is fixed and is three times more than data block size. Therefore,
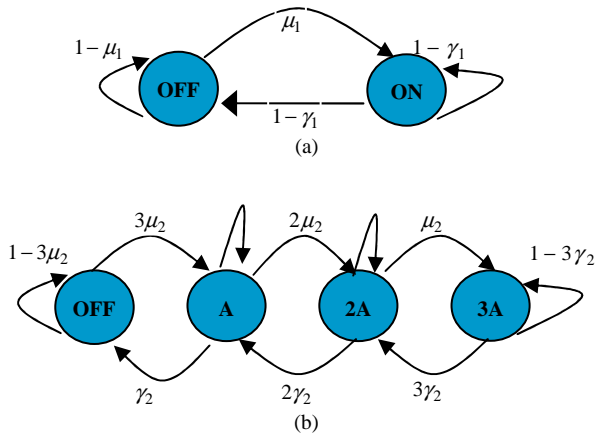
(a)

(b)

Figure 4. ON/OFF traffic model for Voice and VBR traffic.

$$S_1 = H_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad S_1^0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \beta_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

The VBR traffic is modeled by three independent ON/OFF sources as showed in Figure 4b. D, $D_0(2)$ and $D_1(2)$ can be easily calculated like voice traffic matrices. Readers to find more details can refer to [15]. If we assume that the maximum video packets size is 8 times more than data block size, and video packets size follow log-normal distribution with a probability mass function (p_v) in terms of the number of data blocks such a following example [3]

$p_v =$

$[0.002 \quad 0.153 \quad 0.427 \quad 0.286 \quad 0.099 \quad 0.025 \quad 0.006 \quad 0.002]$

The S_2, S_2^0 and β_2 will be

$$S_2 = H_8 \quad S_2^0 = e - S_2 e$$

$\beta_2 =$

$[0.002 \quad 0.006 \quad 0.025 \quad 0.099 \quad 0.286 \quad 0.427 \quad 0.153 \quad 0.002]$

Let SI and TXOP durations be 100 and 10 slots, respectively. Then,

$$V = H_{100} \, , V^0 = e - Ve \quad \delta_i = e_i(100) \ \ 1 \le i \le 11 \ \text{T=10}$$

Using the above information, one can easily find out system performance through the introduced model.

Now, according to the 802.11e and characteristics of the applied traffic streams which are described as follow, the numerical results obtained from the analytical model are compared with simulation results. We analyze the queue length and the access delay distribution as well as the packet loss rate for the high priority and the low priority packets. Similar to above example, one can easily match the proposed model with the introduced traffic and system parameters.

It is assumed that the voice traffic is handled with a higher priority than video traffic. The voice traffic is modeled by an ON/OFF source which generates 160 octet message periodically with a bit rate 64 kb/s during active period. The CBR video traffic has only a ON state and always stays in that. The VBR traffic is modeled by three independent ON/OFF sources with the mean bit rate 200 Kb/s. however, the PhFit program [16] can be used to find out the phase type distribution of service times of real traffic. Table 1 summarizes the different traffic used for the analytical analysis and simulations. It is assumed that the queue buffer size is seven and the channel data rate is 12Mbps.

Simulations are performed using program which is written in C++ medium. There are two queues in each station, and the server processes packets in each of the queues in the FIFO fashion. There are ten stations which are communicating with the access point. All stations enjoy ON/OFF voice traffic as high priority traffic. Five

stations generate the CBR video traffic and the others send the VBR video traffic as low priority traffic. Arrivals to the queues are independent of whether the server is in service or on vacation. The TXOPs duration are calculated through Eqn (1), according to the traffic information in each station. Each station can transmit its data during its TXOP period.

Figure 5 shows cumulative distribution function (CDF) of the high priority and the low priority queue length for VBR, CBR video and voice traffic streams. Although the CBR packet arrival rate is much larger than that of the VBR traffic stream, the queue length in the CBR traffic is less than that one in the VBR traffic considerably. As it is obvious from the mentioned figure, the probability that the length of the queue get less than or equal to six is about 98 percent for the CBR traffic while that is about 74 percent for the VBR traffic. It means that most of the time the VBR packets remain in the queue and unable to be transmitted. Therefore, the packet loss goes up and lots of packets drop. Simulation and numerical results show that the packet loss rate is about 28 percent for the VBR traffic while it is about 0.6 percent for the CBR video traffic. Consequently, although there is enough bandwidth to support QoS guarantee in the VBR video traffic but the scheduler is unable to use it.

The CDF of the queue length at the end of the TXOP for all traffic, plotted in Figure 6, confirms created challenges through the VBR traffic. Therefore, the modification of the scheduling algorithm and introduction of a dynamic scheduler to adapt with the bursty arrivals are unavoidable. Dynamic conditions in the scheduler can be obtained through adjusting the TXOP and the SI durations based on the packet queue length statistics. Scheduler can get the information from the stations and find the optimal TXOP and SI through the employing the model to maintain an empty queue at the end of TXOP duration.

Figure 7 shows CDF of the access delay and packet blocking in the high priority and the low priority traffic through analysis and simulations. It is observed that all packets in the CBR video traffic experience access delay less than about 35 ms, while only 23 percent of the VBR video packets experience such a cumulative access delay. Although there is enough bandwidth to serving the VBR traffic, the scheduler does not have essential flexibility to support of bursty arrival rate. Consequently, the queue will be full and 28 percent of arrived packets are blocked and dropped.

Finally, from Figures 5–7, it can be readily seen that the validation of analytical model is confirmed by the numerical results obtained from analytical model and the simulation results under the same conditions.

Table 1. Description of different traffic streams.

Application	Arrival rate(Kb/s)	Packet size(byte)
Voice	64	160
VBR video	200	660
CBR video	3200	910
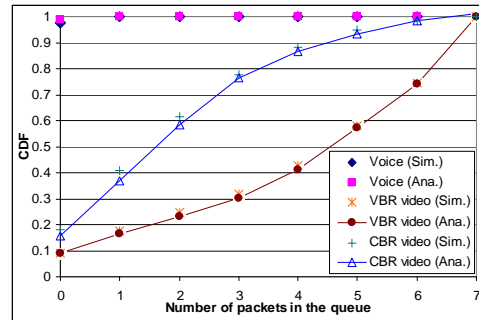
Figure 5. CDF of packet queue length for different traffic streams (Simulation (Sim.) & Analytical (Ana.)).
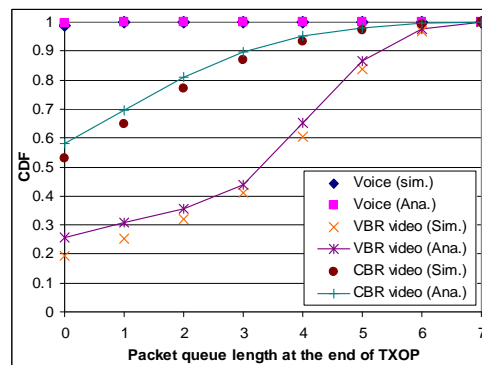
Figure 6. CDF of packet queue length for different traffic streams at the end of TXOP (Simulation (Sim.) & Analytical (Ana.)).
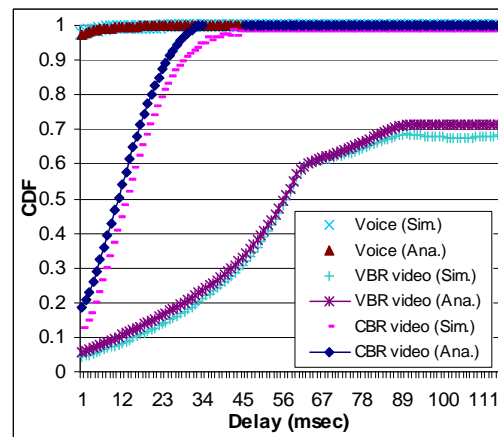
Figure 7. CDF of access delay for different traffic streams (Simulation (Sim.) & Analytical (Ana.)).

6. Conclusions

The scheduling algorithm which is introduced by HCCA /IEEE802.11e to support QoS in multimedia applications enjoys separated queues with specified priority levels and transmission opportunity according to the traffic streams characteristics. The transmission opportunity is found out based on the mean values. Therefore, some multimedia traffic streams such as VBR traffic address

some challenges in this medium. Consequently, adapting algorithms to new conditions in order to provide desired QoS is on the focus of researchers. To investigate and improve the scheduler, analytical model is very useful. This paper introduced a priority queuing model for the HCCA. Using of the MAP/PH/1 queue makes the model more comprehensive and provides it to support different practical traffic streams. The important performance measures in the high priority and the low priority queues are calculated which enable us to investigate the effect of the SI and the TXOP durations on QoS guarantees and find out the optimal TXOP values according to the queue length and the access delay statistics to provide QoS. It is shown by the numerical and the simulation results that the analytical model is quite accurate, and thus useful in the practical system design and performance evaluation.

7. Acknowledgments

8. References

[1] IEEE 802.11e/D13.0, Part 11, Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Medium access control (MAC) enhancements for quality of service (QoS), draft supplement to IEEE 802.11 Std, January 2005.

[2] M. F. Neuts, "Matrix geometric solutions in stochastic models−an algorithmic approach," John Hopkins University Press Baltimore, MD, 1981.

[3] J. A. Zhao, B. Li, X. Cao, and I. Ahmad, "Matrix-analytic solution for DMAP/PH/1 priority queue," Queuing Systems Journal, Springer Netherlands, Vol. 53, No. 3, pp. 127−145. July 2006.

[4] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," IEEE Wireless Communications, pp. 40−50, December 2003.

[5] P. Ansel, Q. Ni, and T. Turletti, "FHCF: A Simple and an efficient scheduling scheme for IEEE 802.11e wireless LAN," Mobile Networks and Applications, No. 11, pp. 391−403, April 2006.

[6] A. Grilo, M. Macedo and M. Nunes, "A Scheduling Algorithm for QoS Support in IEEE802.11e Networks," IEEE Wireless Communications, pp. 36−43, June 2003.

[7] A. Iera, A. Molinaro, G. Ruggeri, and D. Tripodi, "Improving QoS and throughput in single and multihop WLANs through dynamic traffic prioritization," IEEE Network, pp. 35−44, June 2005.

[8] N. Vaidya, A. Dugar, S. Gupta, and P. Bahl, "Distributed fair scheduling in a wireless LAN," IEEE Transactions on Mobile Computing, Vol. 4, No. 6, pp. 616−629, November 2005.

[9] H. Zhai, X. Chen and Y. Fang, "How well can the IEEE 802.11 wireless LAN support quality of service?" IEEE Transactions on Wireless Communications, Vol. 4, No. 6, pp. 3084−3093, November 2005.

[10] Z. Kong, D. Tsang and B. Bensaou, "Performance analysis of IEEE802.11e contention-based channel access," IEEE Journal on Selected Areas in Communications, Vol. 22, No. 10, pp. 2095−2106, December 2004.

[11] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," IEEE Transactions on Wireless Communications, Vol. 4, No. 4, pp. 1506−1515, July 2005.

[12] H. Zhu and I. Chlamtac, "Performance analysis for IEEE 802. 11e EDCF service differentiation," IEEE Transactions on Wireless Communications, Vol. 4, No.4, pp. 1779−1788, July 2005.

[13] M. M. Rashid and E. Hossain, "Queuing analysis of 802.11e HCCA with variable bit rate traffic," IEEE International Conference on Communications, Vol. 10, pp. 4792 −4798, 2006.

[14] A. S. Alfa, "Matrix geometric solution of discrete time MAP/PH/1 priority queue," Naval Research Logistics, Vol. 45, pp. 23−50, July 1997.

[15] C. Blondia and O. Casals, "Performance analysis of statistical multiplication of VBR sources," Conference of the IEEE Computer and Communications Societies, pp. 828−838, 1992.

[16] Horvath and M. Telek, "Phfit: A general phase type fitting tool," in proceedings of 12th performance TOOLS, Lecture Notes in Computer Science, Vol. 2324, London, UK, April 2002.

Appendix A

Block matrices of transition matrix P (Section 3.4)

$$B_{01} = \begin{bmatrix} B_{01}^{00} & B_{01}^{01} & & & & \\ B_{01}^{-1} & B_{01}^{0} & B_{01}^{1} & & & \\ & B_{01}^{-1} & B_{01}^{0} & B_{01}^{1} & & \\ & & \ddots & \ddots & \ddots & \\ & & & B_{01}^{-1} & B_{01}^{0} & B_{01}^{1} \\ & & & & B_{01}^{-1} & B_{01}^{0}+B_{01}^{1} \end{bmatrix}$$

$$B_{01}^{00} = \begin{bmatrix} D_{11}\otimes V & 0 & e'_1(T)\otimes D_{11}\otimes V^0\beta_1 \end{bmatrix}$$

$$B_{01}^{01} = \begin{bmatrix} D_2\otimes V & 0 & e'_1(T)\otimes D_2\otimes V^0\beta_1 \end{bmatrix}$$

$$B_{01}^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ B_{01}^{-1Lv} & H_T\otimes D_{11}\otimes S_2 & H_T\otimes D_{11}\otimes S_2^0\beta_1 \end{bmatrix}$$

$$B_{01}^{-1Lv} = e_T(T)\otimes[D_{11}\otimes S_2^0\delta + D_{11}\otimes e_{m_2}\delta]$$

$$B_{01}^{1} = \begin{bmatrix} D_2\otimes V & 0 & e'_1(T)\otimes D_2\otimes V^0\beta_1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_{10} = \begin{bmatrix} B_{10}^{00} & B_{10}^{1} & & & & \\ & B_{10}^{0} & B_{10}^{1} & & & \\ & & B_{10}^{0} & B_{10}^{1} & & \\ & & & \ddots & \ddots & \\ & & & & B_{10}^{0} & B_{10}^{1} \\ & & & & & B_{10}^{0}+B_{10}^{1} \end{bmatrix}$$

$$B_{10}^{00} = \begin{bmatrix} 0 \\ 0 \\ D_{00}\otimes S_1^0\delta_1 \\ \vdots \\ D_{00}\otimes S_1^0\delta_T \end{bmatrix}$$

$$B_{10}^{1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ e_T(T)\otimes D_{12}\otimes S_1^0\delta & H_T\otimes D_{12}\otimes S_1^0\beta_2 \end{bmatrix}$$

$$B_{10}^{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ e_T(T)\otimes D_{00}\otimes S_1^0\delta & H_T\otimes D_{00}\otimes S_1^0\beta_2 \end{bmatrix}$$

$$A_0 = \begin{bmatrix} A_0^0 & A_0^1 & & & & \\ & A_0^0 & A_0^1 & & & \\ & & A_0^0 & A_0^1 & & \\ & & & \ddots & \ddots & \\ & & & & A_0^0 & A_0^1 \\ & & & & & A_0^0+A_0^1 \end{bmatrix}$$

$$A_0^0 = \begin{bmatrix} D_{00}\otimes V & 0 & e'_1(T)\otimes D_{00}\otimes V^0\beta_1 \\ A_0^{0Lv} & H_T\otimes D_{00}\otimes S_2 & H_T\otimes D_{00}\otimes S_2^0\beta_1 \\ A_0^{0Hv} & 0 & H_T\otimes[D_{00}\otimes S_1+D_{11}\otimes S_1^0\beta_1] \end{bmatrix}$$

$$A_0^{0Lv} = e_T(T)\otimes[D_{00}\otimes e_{m_2}\delta + D_{00}\otimes S_2^0\delta]$$

$$A_0^{0Hv} = e_T(T)\otimes[D_{00}\otimes e_{m_1}\delta + D_{11}\otimes S_1^0\delta]$$

$$A_0^1 = \begin{bmatrix} D_{12}\otimes V & 0 & e'_1(T)\otimes D_{12}\otimes V^0\beta_1 \\ A_0^{1Lv} & H_T\otimes D_{12}\otimes S_2 & H_T\otimes D_{12}\otimes S_2^0\beta_1 \\ A_0^{1Hv} & 0 & H_T\otimes[D_{12}\otimes S_1+D_2\otimes S_1^0\beta_1] \end{bmatrix}$$

$$A_0^{1Lv} = e_T(T)\otimes[D_{12}\otimes e_{m_2}\delta + D_{12}\otimes S_2^0\delta]$$

$$A_0^{1Hv} = e_T(T)\otimes[D_{12}\otimes e_{m_1}\delta + D_2\otimes S_1^0\delta]$$

$$A_1 = \begin{bmatrix} A_1^0 & A_1^1 & & & & \\ & A_1^0 & A_1^1 & & & \\ & & A_1^0 & A_1^1 & & \\ & & & \ddots & \ddots & \\ & & & & A_1^0 & A_1^1 \\ & & & & & A_1^0+A_1^1 \end{bmatrix}$$

$$A_1^0 = \begin{bmatrix} D_{11}\otimes V & 0 & e'_1(T)\otimes D_{11}\otimes V^0\beta_1 \\ A_1^{0Lv} & H_T\otimes D_{11}\otimes S_2 & H_T\otimes D_{11}\otimes S_2^0\beta_1 \\ A_1^{0Hv} & 0 & H_T\otimes D_{11}\otimes S_1 \end{bmatrix}$$

$$A_1^{0Lv} = e_T(T)\otimes[D_{11}\otimes e_{m_2}\delta + D_{11}\otimes S_2^0\delta]$$

$$A_1^{0Hv} = e_T(T)\otimes D_{11}\otimes e_{m_1}\delta$$

$$A_1^1 = \begin{bmatrix} D_2\otimes V & 0 & e'_1(T)\otimes D_2\otimes V^0\beta_1 \\ A_1^{1Lv} & H_T\otimes D_2\otimes S_2 & H_T\otimes D_2\otimes S_2^0\beta_1 \\ A_1^{1Hv} & 0 & H_T\otimes D_2\otimes S_1 \end{bmatrix}$$

$$A_1^{1Lv} = e_T(T)\otimes[D_2\otimes e_{m_2}\delta + D_2\otimes S_2^0\delta]$$

$$A_1^{1Hv} = e_T(T)\otimes D_2\otimes e_{m_1}\delta$$

$$A_{-1} = \begin{bmatrix} A_{-1}^0 & A_{-1}^1 & & & & \\ & A_{-1}^0 & A_{-1}^1 & & & \\ & & A_{-1}^0 & A_{-1}^1 & & \\ & & & \ddots & \ddots & \\ & & & & A_{-1}^0 & A_{-1}^1 \\ & & & & & A_{-1}^0+A_{-1}^1 \end{bmatrix}$$

$$A_{-1}^0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_T(T)\otimes D_{00}\otimes S_1^0\delta & 0 & H_T\otimes D_{00}\otimes S_1^0\beta_1 \end{bmatrix}$$

$$A_{-1}^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_T(T)\otimes D_{12}\otimes S_1^0\delta & 0 & H_T\otimes D_{12}\otimes S_1^0\beta_1 \end{bmatrix}$$

Appendix B

State probability vectors in the high priority access delay (Section 4.3)

$$z_{0H}^1 = \lambda^{-1} \sum_{j=1}^{Q_2} x_{0j}^v (D_{11} + D_2) \otimes V^0$$

$$z_{0H}^k = \lambda_1^{-1} \sum_{j=1}^{Q_2} x_{0j}^{k-1}(2)(D_{11} + D_2) \otimes S_2^0$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{1j}^{k-1}(1)(D_{11} + D_2) \otimes S_1^0 \quad 2 \leq k \leq T$$

$$z_i^v = \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^v (D_{11} + D_2) \otimes V$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^T (1)(D_{11} + D_2) \otimes e_{m_1} \delta$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{i+1j}^T (1)(D_{11} + D_2) \otimes S_1^0 \delta$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^T (2)(D_{11} + D_2) \otimes (S_2^0 \delta + e_{m_2} \delta)$$

$$z_{iL}^k = \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^{k-1}(2)(D_{11} + D_2) \otimes S_2 \quad 2 \leq k \leq T$$

$$z_{0H}^1 = \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^v (D_{11} + D_2) \otimes V^0 \beta_1$$

$$z_{iH}^k = \lambda_1^{-1} \sum_{j=1}^{Q_2} x_{ij}^{k-1}(1)(D_{11} + D_2) \otimes S_1$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{i+1j}^{k-1}(1)(D_{11} + D_2) \otimes S_1^0 \beta_1$$
$$+ \lambda_1^{-1} \sum_{j=0}^{Q_2} x_{ij}^{k-1}(2)(D_{11} + D_2) \otimes S_2^0 \beta_1 \quad 2 \leq k \leq T \quad 1 \leq i < Q_1$$

Block matrices of transition matrix P_h

$$\tilde{A}_0 = \begin{bmatrix} I \otimes V & 0 & e_1'(T) \otimes I \otimes V^0 \beta_1 \\ \tilde{A}_0^{Lv} & H_T \otimes I \otimes S_2 & H_T \otimes I \otimes S_2^0 \beta_1 \\ \tilde{A}_0^{Hv} & 0 & H_T \otimes I \otimes S_1 \end{bmatrix}$$

$$\tilde{A}_0^{Lv} = e_T(T) \otimes [I \otimes e_{m_2} \delta + I \otimes S_2^0 \delta]$$
$$\tilde{A}_0^{Hv} = e_T(T) \otimes I \otimes e_{m_1} \delta$$

$$\tilde{B}_{00} = \begin{bmatrix} I \otimes V & 0 & e_1'(T) \otimes I \otimes V^0 \\ \tilde{B}_{00}^{Lv} & H_T \otimes I \otimes S_2 & H_T \otimes I \otimes S_2^0 \\ \tilde{B}_{00}^{Hv} & 0 & H_T \otimes I \end{bmatrix}$$

$$\tilde{B}_{00}^{Lv} = e_T(T) \otimes [I \otimes e_{m_2} \delta + I \otimes S_2^0 \delta]$$
$$\tilde{B}_{00}^{Lv} = e_T(T) \otimes I \otimes e_{m_1} \delta$$

$$\tilde{A}_{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_T(T) \otimes I \otimes S_1^0 \delta & 0 & H_T \otimes I \otimes S_1^0 \beta_1 \end{bmatrix}$$

Appendix C

Access delay Distribution for the low priority packets (Section 4.3)

The arriving low priority packet will reach the head of its queue and will be ready to transmit if all the low priority packets ahead of it are served. Since the low priority packet is unable to be transmitted while there are high priority packets in the system, the arriving low priority packet has to wait for completion of the transmission of all high priority and low priority packets which are in the system and those high priority packets which will enter during the period of the time that the arriving packet moves towards head of queue. Therefore, the number of packet ahead of arriving low priority packet at its arrival time is all high priority and low priority packets in the system (including any high priority packet which might have arrived jointly with it). Suppose y defines initial probability vector in the low priority access delay.

$$y = [y_0 \ y_1 ... y_{Q_1}] \quad \text{where} \quad y_i = [y_{i0} \ y_{i1} ... y_{iQ_2}]$$
$$y_{ij} = [y_{ij}^v \ y_{ijL}^1 ... y_{ijL}^T \ y_{ijH}^1 ... y_{ijH}^T]$$

where y_{ij}^v, represents probability of an arriving low priority packet finding i high priority and j low priority packet in the system with the server on vacation and y_{ijL}^k (y_{ijH}^k) $1 \leq k \leq T$ is the probability of an arriving low priority packet finding i high priority and j low priority packet in the system with the server in the low priority (high priority) processing at the kth slot of TXOP. All the probability vectors can be calculated in a similar manner which has discussed in the Section 4.3.

$$y_{00}^v = \lambda_2^{-1} \{ x_{00}[D_{12} \otimes V + V^0 \delta_1] + x_{10}^T(1)[D_{12} \otimes S_1^0 \delta]$$
$$+ x_{01}^T(2)[D_{12} \otimes S_2^0 \delta] + \sum_{k=2}^T x_{10}^{k-1}(1)[D_{12} \otimes S_1^0 \delta_k]$$
$$+ \sum_{k=2}^T x_{01}^{k-1}(2)[D_{12} \otimes S_2^0 \delta_k] \}$$

$$y_{0j}^v = \lambda_2^{-1} \{ x_{0j+1}^T(2)[D_{12} \otimes S_2^0 \delta] + x_{0j}^T(2)[D_{12} \otimes e_{m_2} \delta]$$
$$+ x_{1j}^T(1)[D_{12} \otimes S_1^0 \delta] + x_{0j}^v[D_{12} \otimes V] \} \quad 1 \leq j < Q_2$$

$$y_{0jL}^1 = \lambda_2^{-1} \{ x_{0j}^v[D_{12} \otimes V^0 \beta_2] \} \quad 1 \leq j < Q_2$$

$$y_{0jL}^k = \lambda_2^{-1} \{ x_{0j}^{k-1}[D_{12} \otimes S_2] \} \quad 1 \leq j < Q_2 \quad 2 \leq k \leq T$$

$$y_{1j}^v = \lambda_2^{-1} \{ x_{1j}^v[D_{12} \otimes V] + x_{1j}^T(1)[D_{12} \otimes e_{m_1} \delta]$$
$$+ x_{1j}^T(1)[D_2 \otimes S_1^0 \delta] + x_{1j}^T(2)[D_{12} \otimes S_2^0 \delta]$$
$$+ x_{2j}^T(1)[D_{12} \otimes S_1^0 \delta] + x_{0j+1}^T(1)[D_2 \otimes S_2^0 \delta]$$
$$+ x_{0j}^v[D_2 \otimes V] + x_{1j}^T(2)[D_{12} \otimes e_{m_2} \delta]$$
$$+ x_{0j+1}^T(2)[D_2 \otimes e_{m_2} \delta] \} \quad 0 \leq j < Q_2$$

$$y_{1jL}^k = \lambda_2^{-1}\{x_{1j}^{k-1}(2)[D_{12}\otimes S_2] + x_{0j}^{k-1}(2)[D_2\otimes S_2]\}$$
$$0 \le j < Q_2 \quad 2 \le k \le T$$

$$y_{1jH}^1 = \lambda_2^{-1}\{x_{1j}^v[D_{12}\otimes V^0\beta_1] + x_{0j}^v[D_2\otimes V^0\beta_1]\} \quad 0 \le j < Q_2$$

$$y_{1jH}^k = \lambda_2^{-1}\{x_{1j}^{k-1}(1)[D_{12}\otimes S_1] + x_{1j}^{k-1}(2)[D_{12}\otimes S_2^0\beta_1]$$
$$+ x_{2j}^{k-1}(1)[D_{12}\otimes S_1^0\beta_1] + x_{1j}^{k-1}(1)[D_2\otimes S_1^0\beta_1]$$
$$+ x_{0j+1}^{k-1}(2)[D_2\otimes S_2^0\beta_1]\} \quad 0 \le j < Q_2 \quad 2 \le k \le T$$

$$y_{ij}^v = \lambda_2^{-1}\{x_{ij}^v[D_{12}\otimes V] + x_{ij}^T(1)[D_{12}\otimes e_{m_1}\delta]$$
$$+ x_{i-1j}^T(1)[D_2\otimes e_{m_1}\delta] + x_{i+1j}^T(1)[D_{12}\otimes S_1^0\delta]$$
$$+ x_{ij}^T(2)[D_{12}\otimes S_2^0\delta] + x_{ij}^T(1)[D_2\otimes S_1^0\delta]$$
$$+ x_{i-1j}^T(2)[D_2\otimes S_2^0\delta] + x_{i-1j}^v[D_2\otimes V]$$
$$+ x_{ij}^T(2)[D_{12}\otimes e_{m_2}\delta] + x_{i-1j}^T(2)[D_2\otimes e_{m_2}\delta]\}$$
$$i \ge 2 \quad 0 \le j < Q_2$$

$$y_{ijL}^k = \lambda_2^{-1}\{x_{ij}^{k-1}(2)[D_{12}\otimes S_2] + x_{i-1j}^{k-1}(2)[D_2\otimes S_2]\}$$
$$2 \le i \le Q_1 \quad 2 \le k \le T \quad 0 \le j < Q_2$$

$$y_{1jH}^1 = \lambda_2^{-1}\{x_{ij}^v[D_{12}\otimes V^0\beta_1] + x_{i-1j}^v[D_2\otimes V^0\beta_1]\}$$
$$2 \le i \le Q_1 \quad 0 \le j < Q_2$$

$$y_{ijH}^k = x_{ij}^{k-1}(1)[D_{12}\otimes S_1] + x_{i-1j}^{k-1}(2)[D_2\otimes S_1]$$
$$+ x_{i+1j}^{k-1}(1)[D_{12}\otimes S_1^0\beta_1] + x_{ij}^{k-1}(1)[D_2\otimes S_1^0\beta_1]$$
$$+ x_{ij}^{k-1}(2)[D_{12}\otimes S_2^0\beta_1] + x_{i-1j}^{k-1}(2)[D_2\otimes S_2^0\beta_1]$$
$$2 \le i \le Q_1 \quad 0 \le j < Q_2$$

Transition matrix for the low priority packet access delay:

$$P_l = \begin{bmatrix} \tilde{B}_{00} & \tilde{B}_{01} & & & & \\ \tilde{B}_{10} & \tilde{A}_0 & \tilde{A}_1 & & & \\ & \tilde{A}_{-1} & \tilde{A}_0 & \tilde{A}_1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \tilde{A}_{-1} & \tilde{A}_0 & \tilde{A}_1 \\ & & & & \tilde{A}_{-1} & \tilde{A}_0 + \tilde{A}_1 \end{bmatrix}$$

$$\tilde{B}_{00} = \begin{bmatrix} \tilde{B}_{00}^{00} & & & & & \\ \tilde{B}_{00}^{10} & \tilde{B}_{00}^0 & & & & \\ & \tilde{B}_{00}^{-1} & \tilde{B}_{00}^0 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \tilde{B}_{00}^{-1} & \tilde{B}_{00}^0 & \\ & & & & \tilde{B}_{00}^{-1} & \tilde{B}_{00}^0 \end{bmatrix}$$

$$\tilde{B}_{00}^{00} = (D_{00}+D_{12})\otimes(V+V^0\delta_1)$$

$$\tilde{B}_{00}^{10} = \begin{bmatrix} 0 \\ (D_{00}+D_{12})\otimes S_2^0\delta_1 \\ \vdots \\ (D_{00}+D_{12})\otimes S_2^0\delta_T \end{bmatrix}$$

$$\tilde{B}_{00}^0 = \begin{bmatrix} (D_{00}+D_{12})\otimes V & e_1'(T)\otimes(D_{00}+D_{12})\otimes V^0\beta_2 \\ e_T(T)\otimes(D_{00}+D_{12})\otimes e_{m_2}\delta & H_T\otimes(D_{00}+D_{12})\otimes S_2 \end{bmatrix}$$

$$B_{00}^{-1} = \begin{bmatrix} 0 & 0 \\ e_T(T)\otimes(D_{00}+D_{12})\otimes S_2^0\delta & H_T\otimes(D_{00}+D_{12})\otimes S_2^0\beta_2 \end{bmatrix}$$

$$\tilde{B}_{01} = \begin{bmatrix} \tilde{B}_{01}^{00} & & & & \\ \tilde{B}_{01}^{-1} & \tilde{B}_{01}^0 & & & \\ & \tilde{B}_{01}^{-1} & \tilde{B}_{01}^0 & & \\ & & \ddots & \ddots & \ddots \\ & & & \tilde{B}_{01}^{-1} & \tilde{B}_{01}^0 \\ & & & & \tilde{B}_{01}^{-1} & \tilde{B}_{01}^0 \end{bmatrix}$$

$$\tilde{B}_{01}^{00} = \begin{bmatrix} (D_{11}+D_2)\otimes V & 0 & e_1'(T)\otimes(D_{11}+D_2)\otimes V^0\beta_1 \end{bmatrix}$$

$$\tilde{B}_{01}^0 = \begin{bmatrix} (D_{11}+D_2)\otimes V & 0 & e_1'(T)\otimes(D_{11}+D_2)\otimes V^0\beta_1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\tilde{B}_{01}^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ \tilde{B}_{01}^{-1Lv} & H_T\otimes(D_{11}+D_2)\otimes S_2 & H_T\otimes(D_{11}+D_2)\otimes S_2^0\beta_1 \end{bmatrix}$$

$$\tilde{B}_{01}^{-1Lv} = e_T(T)\otimes[(D_{11}+D_2)\otimes S_2^0\delta + (D_{11}+D_2)\otimes e_{m_2}\delta]$$

$$\tilde{B}_{10} = \begin{bmatrix} \tilde{B}_{10}^{00} & & & & \\ & \tilde{B}_{10}^0 & & & \\ & & \tilde{B}_{10}^0 & & \\ & & & \ddots & \\ & & & & \tilde{B}_{10}^0 & \\ & & & & & \tilde{B}_{10}^0 \end{bmatrix}$$

$$\tilde{B}_{10}^0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ e_T(T)\otimes(D_{00}+D_{12})\otimes S_1^0\delta & H_T\otimes(D_{00}+D_{12})\otimes S_1^0\beta_2 \end{bmatrix}$$

$$\tilde{B}_{10}^{00} = \begin{bmatrix} 0 \\ 0 \\ (D_{00}+D_{12})\otimes S_1^0\delta_1 \\ \vdots \\ (D_{00}+D_{12})\otimes S_1^0\delta_T \end{bmatrix}$$

$$\tilde{A}_0 = I\otimes\tilde{A}_0^0 \qquad \tilde{A}_1 = I\otimes\tilde{A}_1^0 \qquad \tilde{A}_{-1} = I\otimes\tilde{A}_{-1}^0$$

$$\tilde{A}_0^0 = \begin{bmatrix} (D_{00}+D_{12})\otimes V & 0 & e_1'(T)\otimes(D_{00}+D_{12})\otimes V^0\beta_1 \\ \tilde{A}_0^{0Lv} & H_T\otimes(D_{00}+D_{12})\otimes S_2 & H_T\otimes(D_{00}+D_{12})\otimes S_2^0\beta_1 \\ \tilde{A}_0^{0Hv} & 0 & \tilde{A}_0^{0HH} \end{bmatrix}$$

$$\tilde{A}_0^{0Lv} = e_T(T)\otimes[(D_{00}+D_{12})\otimes e_{m_{21}}\delta + (D_{00}+D_{12})\otimes S_2^0\delta]$$
$$\tilde{A}_0^{0Hv} = e_T(T)\otimes[(D_{00}+D_{12})\otimes e_{m_1}\delta + (D_{00}+D_{12})\otimes S_1^0\delta]$$
$$\tilde{A}_0^{0HH} = H_T\otimes[(D_{00}+D_{12})\otimes S_1 + (D_{11}+D_2)\otimes S_1^0\beta_1]$$

$$\tilde{A}_1^0 = \begin{bmatrix} (D_{11}+D_2)\otimes V & 0 & e_1'(T)\otimes(D_{11}+D_2)\otimes V^0\beta_1 \\ \tilde{A}_1^{0Lv} & H_T\otimes(D_{11}+D_2)\otimes S_2 & H_T\otimes(D_{11}+D_2)\otimes S_2^0\beta_1 \\ \tilde{A}_1^{0Hv} & 0 & H_T\otimes(D_{11}+D_2)\otimes S_1 \end{bmatrix}$$

$$\tilde{A}_1^{0Lv} = e_T(T)\otimes[(D_{11}+D_2)\otimes e_{m_2}\delta + (D_{11}+D_2)\otimes S_2^0\delta]$$
$$\tilde{A}_1^{0Hv} = e_T(T)\otimes(D_{11}+D_2)\otimes e_{m_1}\delta$$

$$\widetilde{A}_{-1}^0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_T(T) \otimes (D_{00} + D_{12}) \otimes S_1^0 \delta & 0 & H_T \otimes (D_{00} + D_{12}) \otimes S_1^0 \beta_1 \end{bmatrix}$$

Suppose $y^0 = y$. Then, $y^{n+1} = y^n P_l$

To reduction of computation, the set of following equation can be used.

$$\begin{cases} y_0^{n+1} = y_0^n \widetilde{B}_{00} + y_1^n \widetilde{B}_{10} \\ y_1^{n+1} = y_0^n \widetilde{B}_{01} + y_1^n \widetilde{A}_0 + y_2^n \widetilde{A}_{-1} \\ y_i^{n+1} = y_{i-1}^n \widetilde{A}_1 + y_i^n \widetilde{A}_0 + y_{i+1}^n \widetilde{A}_{-1} \quad i \geq 2 \end{cases}$$

Finally, let W_{LT} be the probability that the waiting time of low priority packet is less than or equal to T. Then, $W_{LT} = y_{00}^T P_l$.