Traffic Characteristics based Performance Analysis Model for Efficient Random Access in OFDMA-PHY System

Hyun-Hwa Seo[1], Byung-Han Ryu[1], Choong-Ho Cho[2], and Hyong-Woo Lee[3]

[1] Department of Mobile Transmission Technology Research
Electronics and Telecommunication Research Institute, Daejeon, Korea 305-350
{hhseo, rubh}@etri.re.kr
[2] Department of Computer and Information Science[2], Electronics and Information Engineering[3],Korea University, Choongnam, Korea 339-700
{chcho, hwlee}@korea.ac.kr

Abstract. Currently, IEEE 802.16a wireless MAN supports the contention based OFDMA-CDMA ranging subsystem for ranging operation (Initial Ranging, Periodic Ranging, Bandwidth Request)[1]. This system uses essentially the slotted ALOHA protocol. However, the number of code-slots/frame for random access transmission tends to be much greater than one and a simple Markov chain analysis may not be numerically feasible. At the same time, the frame size in number of code-slots may be dynamically adjusted based on the traffic load. In order to evaluate delay-throughput performance and stability measure of the random access protocol, we first examine the possible traffic load to be carried throughput the ranging subchannel. We the present performance analysis and numerical examples.

1 Introduction

In the IEEE 802.16a CDMA based OFDMA-PHY, Subscriber Stations(SS) access to Base Station(BS) that uses ranging subchannel to transmit code based Initial Ranging(IR), Periodic Ranging(PR) and Bandwidth Request(BR) in random access mode[1]. Initial ranging transmission is used in initial maintenance interval when beginning connection. Bandwidth-request transmissions are for requesting uplink allocations from the BS, and periodic ranging transmissions are sent for periodic adjustment for the system reflecting the channel conditions (timing and power adjustments). This procedure can reduce guard time through timing alignment on uplink and achieve power adjustment. For this, the SS shall choose randomly a ranging slot (i.e. OFDM symbol number, subchannel, etc.) as the time to perform the ranging, then it chooses randomly a ranging code and sends it to the BS(as a CDMA code). This system is basically similar to Slotted ALOHA[2], but the frame size (code-slot/frame) allocated for random access or ranging request is relatively big and can be altered by the dynamic adjustment of downlink and uplink subframe size on the MAC frame of OFDMA-PHY.

Although the performance of existing random access protocols such as slotted ALOHA [2], WLAN [3] and HIPERLAN[4][5] has been widely studied, to the best of our knowledge an analysis of the random access protocol employed by the OFDMA-CDMA ranging subchannel is not available in the open literature.

Considering that one of the characteristics of the system is that the frame size in number of code-slots can be dynamically adjusted according toe traffic load, we will briefly examine traffic load of IR, PR and BR. The estimated traffic load, then, can be used to appropriately allocate code-slots to different types of traffic according to the performance requirements.

This paper is organized as following. In section II, IEEE 802.16a OFDMA-CDMA environment is described. In section III, each traffic model is explained by generation rate of initial ranging, periodic ranging and bandwidth-request to do random access. In section IV, performance analysis is given. Numerical examples are examined in section V, followed by a conclusion in section VII.

2 Ranging Procedure [1]

System environment for analytic model in our study have followed std. IEEE 802.16a CDMA based OFDMA-PHY. It was designed based on TDD, and each frame consists of uplink and downlink subframes. The structure of ranging sub-channel in an OFDMA MAC frame.

Initial ranging transmission is used in initial maintenance interval when beginning connection. Bandwidth-request transmissions are for requesting uplink allocations from the BS, and periodic ranging transmissions are sent for periodic adjustment for the system reflecting the channel conditions (timing and power adjustments).

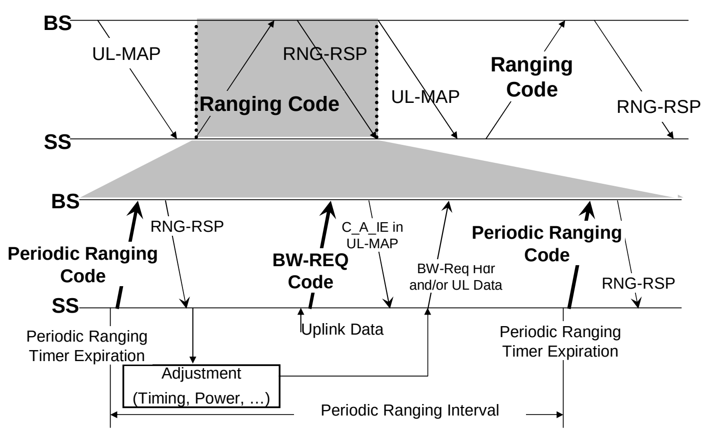

Fig. 1. Ranging procedure: initial ranging, periodic ranging, bandwidth-request

3 Traffic Modeling for Ranging Requests

3.1 Initial Ranging and Periodic Ranging

Initial ranging transmissions are used by SSs who want to synchronize system channel. The arrival of initial ranging packets can be modeled by a Poisson process whose rate depends on the number of SSs which are in idle state.

Periodic ranging is performed by active SSs periodically to adjust system parameters such as timing and power level. If the previous request was successful, the SS sends a request to BS at every 30 seconds. Hence, periodic ranging packets can be generated at every 30 seconds by an active SS.

3.2 Bandwidth-Request

Bandwidth-request transmissions are used to demand required bandwidth when a SS has data to send. Since bandwidth request transmission of each traffic class is different.

Real-time Service Real-time traffic is voice and video. A lot of studies of voice and video traffic model are reported until now. The ON/OFF model is the most simple and representative of voice traffic model. Here, bandwidth-request is generated at the beginning point of ON period (i.e. Bandwidth-request transmission is occurred at the beginning instant of every an ON period, and regenerated at every 2.35 seconds in average). Also, bandwidth-request transmission in video traffic[10],[11] can property piggyback subsequent transmission of data if first transfer succeeds.

HTTP A model for web browsing in shown in [6]. In the model a session consists of alternating cycles of HTTP ON and HTTP OFF periods. A HTTP ON period, in turn, consists of a main object and a number of embedded objects. One main object includes multiple embedded objects. But, this is just downlink traffic model; uplink traffic model is not established yet. Therefore, we propose traffic model for uplink transmission.

First, the main/embedded objects in TCP are transmitted as IP packets, which cannot be larger than the maximum transfer unit (i.e., the MTU in Ethernet system is 1500 bytes). In response to a main/embedded object, ACKs are generated by a destination SS. To measure the number of ACKs according to the size of the main/embedded objects, we use TCP dump based ANYPA-LAN tool, and execute excommunications with different file sizes (1/3/5 Mbytes), different web servers, and different up and download times. In these experiments, we identified that an ACK is generated for every two data packets on the average which coincides with the *delayed ACK policy*[12] in TCP and *ACK-every-other-segment policy*[13][14]. Therefore, possible number of BRs including the number of ACKs is given by $\lceil \text{Main(or Embedded) Object size}/(2MTU) \rceil$ number of objects. And each BR size can be modeled as shown in Fig. 2.

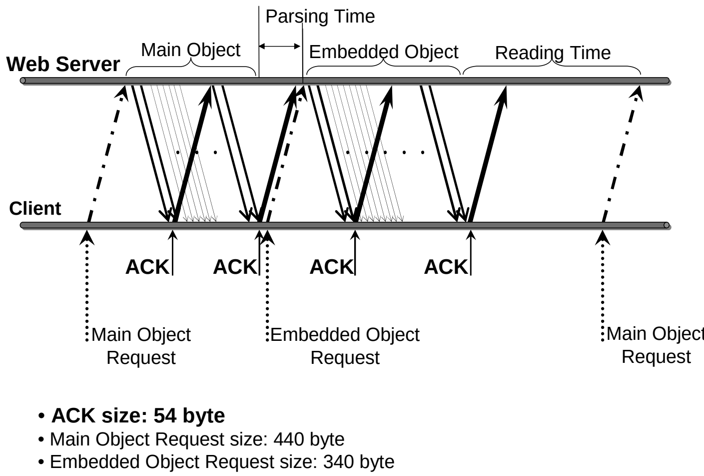

Fig. 2. Web traffic model at SSs.

Best Effort Service Messages arrives to mailboxes can be modeled by Poisson process. In case of receiving messages, the client host sends approximately a half times of the maximum number of packets[3] as shown in Fig. 2. On the other hand, in case of uplink emailing, the number of $\lceil$e-mail size$/MTU\rceil$ is transmitted because the whole message becomes a IP packet by capsulization. Therefore, from the client host, In cases of reception and transmission, approximately $\lceil$e-mail size$/MTU/2\rceil$ number of ACKs and $\lceil$e-mail size$/MTU\rceil$ number of IP packets are generated respectively.

4 Performance Analysis Model

To gain some understanding of the ranging subsystem, we formulate the analytical model by two steps as follow. First, we use a Discrete Time Markov Chain(DTMC)model for a more precise model. This model is divided into "Exact analysis" that applies Feller's classical occupancy theorem[6] and "Approximate analysis" (or "Asymptotic analysis") that uses an approximation by a Poisson distribution for the number of successful transmission in a frame. However, the DTMC model has numerical precision problem when the frame size is large. In order to circumvent the numerical precision problem, we ignore the discrete nature of the frame structure and use a Continuous Time Markov Chain(CTMC) which is a generalized M/M/1 queue. Therefore, the DTMC model is given for the derivation of CTMC and understanding the characteristics of each performance model under the small frame size. Finally, numerical results are shown illustrating the trading relations among the throughput-mean delay time and

[3] The maximum number of packets is $\lceil$file size$/MTU\rceil$.

throughput-first exit time for system stability under the big frame size by using generalized M/M/1.

4.1 Assumptions

We assume the following.

Table 1. A Summary of System Parameters

Parameters	value
System	OFDMA/TDD
New ranging request arrival rate(λ)	$[0.02\ 0.3] * n$
The number of code-slot/frame(n)	$5\ 50, 12, 576$
The number of symbol and subchannel per frame	fixed
Multiple Access Interference(MAI) between codes	no
Retransmission randomization	no
Immediate feedback	yes
Frame period	5 ms

4.2 Mean Delay Time(MDT) Analysis; Discrete Time Markov Chain

For a more precise analysis, we derive a DTMC model, which is analyzed exactly by applying Feller's classical occupancy theorem [6] and approximately by an asymptotic application of the Poisson distribution.

Let $X^{(t)}$ be a random variable representing the total number of collided code slots in the t^{th} frame. Let us also define $\pi_k^{(t)}$ to be the probability of finding the system in channel state k in the t^{th} frame, that is,

$$\pi_k^{(t)} = P[X^{(t)} = k] \tag{1}$$

where, k is the number of collided code-slot.
We then define the probability vector $\underline{\pi}^{(t)}$ in the t^{th} frame as

$$\underline{\pi}^{(t)} = (\pi_0^{(t)}, \pi_1^{(t)}, ..., \pi_k^{(t)}, ...) \quad and\ so \quad \underline{\pi}^{(t+1)} = \underline{\pi}^{(t)} P \tag{2}$$

where, $P = p_{ij}$ is the transition probability matrix[4].

Under the assumption that the system is stable, the steady state probability can be written as $\pi_k = \lim_{t\to\infty} \pi_k^{(t)}$, which satisfies the following equations:

$$\underline{\pi} = (\pi_0, \pi_1, ..., \pi_k, ...), \underline{\pi} = \underline{\pi} P, \tag{3}$$

[4] transition probability is $p_{ij} = Pr[X^{(t+1)} = j | X^{(t)} = i]$.

$$\sum_{k=0}^{k_{max}} \pi^{(t+1)} = 1 \quad (4)$$

From these, we calculate the sequence of values $\underline{\pi}$ and then the MDT($\overline{D}$) can be written as:

$$\overline{D} = \frac{\sum k\pi_k}{\lambda} \quad (5)$$

In order to obtain p_{ij}, we use

$$p_{ij} = \sum_{ij} P[R_{new}]\cdot P[(i+a-j)R_{suc}|(i+a)R_{trn}] \quad (6)$$

where, $R_{new}, R_{suc}, R_{tran}$ are each request of a new arrivals, request of a successes, and request of a transmissions. In (6), the conditional probability of the successful code-slot in a frame is obtained by applying the classical occupancy theorem derived by Feller[6].

$$\frac{(-1)_{i+a-j}n!(i+a)!}{(i+a-j)!n^{i+a}} \times \sum_{l=i+a-j}^{min(n,i+a)} (-1)^l \frac{(n-l)^{i+a-l}}{(l-i-a+j)!(n-l)!(i+a-l)!} \quad (7)$$

where i is the number of collided code-slot in the previous frame, a is the number of new BRs or ranging requests, and n is the maximum number of code-slot per frame.

Also, as proven by Feller[6], if the number of successful code-slot in a frame is approximated by a Poisson distribution with parameter α $P[(i+a-j)R_{suc}|(i+a)R_{trn}]$ can be approximated as (9).

$$\alpha = (i+a)e^{-(i+a)/n} \quad (8)$$

$$P[(i+a-j)R_{suc}|(i+a)R_{trn}] = \frac{e^{-\alpha}\alpha^{i+a-j}}{(i+a-j)!} \quad (9)$$

4.3 MDT and FET; Continuous Time Markov Chain

If we regard the behavior of the continual ranging subframes as continuous time and do not distinguish between new bandwidth requests and retransmission requests, then we can assume that the bandwidth request arrival process is Poisson, the service times are exponentially distributed, and there is a single server. Consequently, the system can be modeled as a birth-death process with arrival rate $\lambda_k = \lambda$ and a state dependent service rate $\mu_k = \mu (= ke^{-k/n})$, where k includes the total number of backlogged and new bandwidth requests and n is the number of code-slot per frame. Under M/M/1 formulation, we first can get probability of the number of packets in system, k, p_k as

$$p_k = \frac{\lambda_{k-1}}{\mu_k} p_{k-1} = p_0 \left(\prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} \right) = \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{\lambda_i}{\mu_i} \right]^{-1} \left(\prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} \right) \quad (10)$$

The MDT, $\overline{D}$ is given by

$$\overline{D} = \left(\frac{1}{\lambda} \sum_{k=1}^{\infty} k p_k \right) - \frac{1}{2} \tag{11}$$

where 1/2 accounts for the average waiting frame time for new bandwidth requests due to frame synchronization which was not included in the discrete time models.

Secondly, we consider the First Exit Time(FET). FET is the average time for the system to make a first exit into the unstable region starting from an initially empty system, which means that all code-slot in the frame are empty. In other words, the stability definition is as follows. A ranging subchannel is said to be stable if its service rate μ_k is greater than λ. Otherwise, the channel is said to be unstable.

Define k_{cr} to be the critical state of channel, which indicates the minimum k for the system to enter the unstable state for the first time.

$$k_{cr} = \min_{\mu_k < \lambda} k \tag{12}$$

Let the transition probabilities from state i to $i+1$ and from state i to $i-1$ be $\theta_i = \frac{\lambda_i}{\lambda_i + \mu_i}$ and $\overline{\theta}_i = 1 - \theta$, respectively. Then the mean transition time $t_{i,i+1}$ from state i to $i+1$ can be written as

$$t_{i,i+1} = \frac{1}{\theta_i} \left[\frac{\theta_i}{\lambda} + \overline{\theta}_i \left(\frac{1}{\mu_i} + t_{i-1,i} \right) \right] \tag{13}$$

where

$$t_{01} = \frac{1}{\lambda}. \tag{14}$$

Using (12),(13), we can express the FET $t_{0,k_{cr}}$ as

$$t_{0,k_{cr}} = \sum_{i=0}^{k_{cr}-1} t_{i,i+1} \tag{15}$$

5 Analysis Result

In order to verify the validity of the analysis and to provide a number of examples of how these analyses can be used to evaluate and compare the performance of the contention based OFDMA-CDMA ranging subchannel, this section presents some numerical and simulation results on the MDT and the FET by numerical computations and simulations using MATLAB. In the simulations of the ranging subchannel, we assume that the arrival process of access requests at each SS is Poisson with rate λ, and let the number of backlogged requests vary according to the specified protocol. The duration of each simulation run is 1000 seconds. Other basic assumptions are the same as given in Section 3. Using the equations in the previous subsections, we numerically compare with the results of the "exact analysis" and "approximate (or asymptotic) analysis" using DTMC. We also compare the analytical results from DTMC and CTMC with simulation results for a system with a small number of code-slots/frame. Applying the CTMC model to a system with a large frame size, we investigate the tradeoff between throughput and FET. Finally, we observe system stability to be determined by CTMC M/M/1 for estimating the arrival rate, the critical state, and the number of backlogged users.

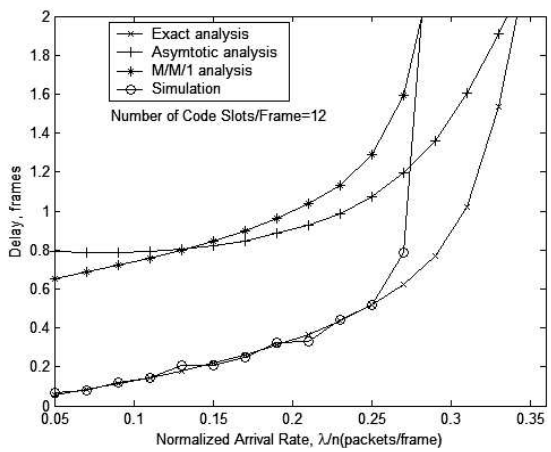

Fig. 3. Web traffic model at SSs.

In Fig. 3, we compare the MDT from the three analytical models and simulations, under a small frame size (number of code-slots/frame n=12). As expected MDT increases as the normalized arrival rate λ/n(throughput) of bandwidth requests per code slot is increased. We observe that the results from the exact analysis and simulations are very close over a broad range of arrival rates below saturation, which are in turn lower than those from the approximate and M/M/1 models. We also observe that MDTs from the approximate analysis and M/M/1 are greater than the corresponding values of the exact analysis and simulations. We can conclude from these results that analysis using the approximate M/M/1 model yields more pessimistic results than the performance of the real system.

To further understand the stability characteristics of the system, we compute FET from (15) over the range of normalized arrival rates of interested identified above, i.e., between 0.3 and 0.35. Results are shown in Fig. 4. It is apparent that FET increases with the number of code-slots, n, and decreases as normalized arrival rate is increased. This is because the number of collisions increases when n is reduced or the arrival rate is increased, resulting in a decrease in FET. Currently, the expected number of code-slot in the OFDMA-CDMA ranging

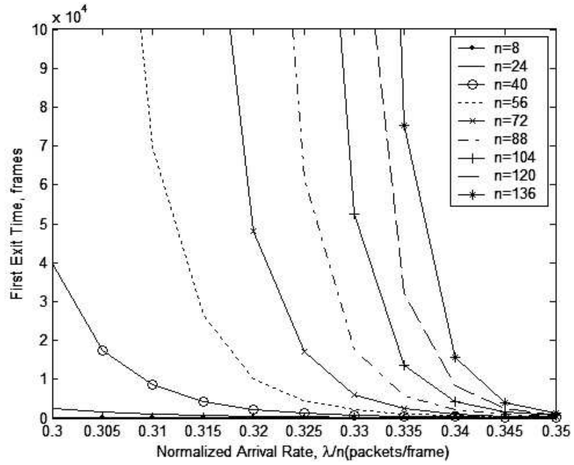

Fig. 4. Web traffic model at SSs.

subchannel of IEEE 802.16a is about 5 to 12 symbols per frame and 5 to 48 codes/symbol, for a total of 25 to 576 available code slots per frame.

6 Conclusion

We have presented a theoretical analysis of the contention based OFDMA-CDMA ranging subchannel in IEEE 802.16a wireless MANs and validate the theoretical models by simulations. To evaluate delay-throughput tradeoffs and system stability, the MDT and FET performance of the ranging subchannel have been analyzed by means of an exact DTMC model that becomes untractable for systems with a large number of code slots/frame, and an approximate CTMC M/M/1 model based upon the DTMC model that can be applied to systems with any number of code slots/frame. Numerical results show that the approximate CTMC M/M/1 model yields somewhat pessimistic results that are nevertheless close to those obtained using the exact model. The approach presented in this paper is also applicable to other multi-channel multi-slot random access systems such as the random access channel in HIPERLAN2 [5].

In addition, radio environment has various restriction(i.e., limited resources, channel state, mobile, etc.). For example, even if channel status is good, it is difficult to assign radio resource efficiently. Most packets in Internet traffic are bigger in size than MTU. The packets in TCP are transmitted as IP packets. Then as a reply to successful transfer of packets that exceed MTU size, ACKs are generated in TCP. The ACK packets consume radio resource for random access. So, in case of Internet traffic modelling, must consider ACK certainly.

Performance analysis result of this paper can be performance reference mark for random access protocol in OFDMA-CDMA ranging system and traffic modeling can be used to do to prepare network resources appropriately in radio network.

References

1. IEEE Standard for Local and Metropolitan Area Networks: Part16: Air Interface for Fixed Broad Wireless Access Systems-Medium Access Control Modification and Additional Physical Layer Specification for 2-11 GHz", IEEE Std. 802.16aTM-2003 (2003)
2. Leonard, Kleinrock, Simon, S. Lam: Packet Switching in a Multiaccess Broadcast Channel:Performance Evaluation IEEE Transaction on Communications, Vol. Com-23, No.4, April (1975), pp.410-423
3. S. Simoens, P. Pellati, J. Gosteau, and K. Gosse: The Evolution of 5 GHz WLAN toward Higher Throughputs IEEE Wireless Communications, Vol. 10, pp.6-13, December 2003
4. G.-H. Hwang and D.-H. Cho: Adaptive random channel allocation scheme in HIPERLAN type 2, IEEE Communications Letters, vol. 6, pp. 40-42, January (2002)
5. G. Anastasi and L. Lenzini: HIPERLAN/1 MAC Protocol: Stability and Performance Analysis, IEEE Journal on Selected Areas in Communications, vol. 18, pp. 1787-1798, September (2000)
6. William Feller: An Introduction to Probability Theory and its Applications, Vol. 1 Comell University, Jan (1950)
7. Leonard Kleninrock: Queueing Systems, Vol.1:Theory, 417 pp. John Wiley and Sons-Interscience, (1975)
8. Alex Brand, Hamid Aghvami: Multiple Access Protocols for Mobile Communications: GPRS, UMTS and Beyond, WILEY, (2001)
9. Traffic Models for IEEE 802.20 MBWA System Simulations, IEEE C802.20-03/66
10. 1xEV-DV Evaluation Methodogy, 3GPPe/TSG-C. R1002. Document available in 802.20 drop-box
11. Hun-Jeong Kang, Myung-Sup Kim, and James W. Hong: Streaming Media and Multimedia Conferencing Traffic Analysis Using Payload Examination, ETRI Journal, vol. 26, no.3, June (2004), pp.203-217.
12. Gary R. Wright W. Richard Stevens: TCP/IP Illustrated, Vol. 2, Addison-wesley professional computing series.
13. [RFC793] Transmission Control Protocol
14. [RFC879] The TCP Maximum Segment Size and Related Topics