Modeling Throughput and Delay in 802.11 Infrastructure Mode Networks with QoS Support from the Point Coordination Function

Mahasweta Sarkar, Christopher P. Paolini, and Santosh Nagaraj

Abstract— **This paper analyses a hybrid medium access control (MAC) scheme that alternates between contention and polling periods in an infrastructure mode wireless local area network (WLAN). The scheme can be thought of as a standard 802.11 channel access scheme operating both in the Distributed Coordination Function (DCF) and Point Coordination Function (PCF) modes for enhanced Quality of Service (QoS) performance. An analytical model has been proposed by the authors for such a hybrid MAC, and dimensionless expressions depicting throughput and delay have been derived by the authors. Graphs displaying throughput and delay characteristics of the hybrid MAC as a function of packet collision probability and high priority station count are presented. The author's dimensionless expressions characterizing throughput and delay as a function of the contention free period ratio is significant because the model can be used in infrastructure mode wireless networking by an access point to dynamically adjust the contention free period ratio to an optimal value that minimizes delay while maximizing throughput, based on current high priority station count.**

Index Terms—**WLAN; QoS; Throughput Analysis; Delay Analysis, Infrastructure Mode; Hybrid MAC Model; Point Coordination Function; Distributed Coordination Function**

I. Introduction

Traffic generated in any data network system is neither uniform nor homogeneous. Performance sensitive traffic such as voice and video applications require stringent delay constraints while data packets of a file transfer application, for example, can operate over a much broader delay and throughput requirement. In order to provide differentiated service to different categories of traffic, the IEEE 802.11e MAC standard [1] has the provision of traffic classification and prioritization. The standard classifies network traffic into four different priority level or access categories (ACs). Nodes maintain separate queues for each AC and packets at the head-of-line (HOL) of each queue contend for channel access using AC-specific parameters.

Manuscript received July 24, 2012; revised August 10, 2012.

M. Sarkar is with San Diego State University, San Diego, CA 92182 USA (phone: 619-594-7797; fax: 619-594-2654; e-mail: msarkar2@mail.sdsu.edu).

C. P. Paolini is with San Diego State University, San Diego, CA 92182 USA (e-mail: paolini@engineering.sdsu.edu).

S. Nagaraj is with San Diego State University, San Diego, CA 92182 USA (e-mail: snagaraj@mail.sdsu.edu).

On the other hand, legacy 802.11 [2] has the provision for nodes to be operating in two different modes – (a) DCF or Distributed Coordination Function and (b) PCF or Point Coordination Function. While DCF is based on the contention based CSMA/CA [3] mode of channel access, PCF is based on the polling mechanism.

Limited QoS support in the legacy 802.11 standard is available through the use of the PCF. Previous work on analyzing throughput and delay in wireless networks has focused on modeling performance aspects of the IEEE 802.11 standard [3,4,5,6,7,8,9] while considering the distributed coordination function (DCF) only. Other publications that analyze 802.11e have provided an overview of the QoS enhancement [10,11], or an analysis of performance for particular applications [12], but have not attempted to model throughput and delay in analytical form. This paper takes a first step toward finding analytical expressions for modeling throughput and delay characteristics of a MAC protocol that mimics the IEEE 802.11e in every essential respect. We first propose a simplified model of the IEEE 802.11e MAC. This model can be thought of as a hybrid MAC model which operates in both the contention and contention free phases alternately akin to a legacy 802.11 MAC protocol with both its DCF and PCF mode enabled. The rest of the paper is organized as follows: Section III describes our 802.11e-like Hybrid MAC model before enumerating the details of the system model in Section IV. Section V presents analytical expressions for dimensionless throughput and delay for our Hybrid MAC. Section VI presents graphs to visualize delay and throughput as a function of collision probability. The paper is concluded in Section VII.

II. Nomenclature

MAC *media access control* is a sublayer of the data link layer that provides channel access control facilities to allow several stations to have multiple access to a wireless network.

LP *low priority* packets belonging to background or best effort protocols such as FTP and SMTP.

HP *high priority* packets belonging to video or voice protocols such as RTP, RTSP, and VoIP.

AC *access category* is a traffic priority classification. In this paper we use two ACs: HP and LP.

DCF *distributed coordination function* is a fundamental access method that uses the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance)

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

scheme to sense a channel and wait until the channel is free before transmitting. If the channel is busy, the DCF will backoff a number of slot times before attempting to transmit again.

PCF *point coordination function* is a channel access method that is centrally controlled by a point coordinator implemented in an access point. The point coordinator maintains a list of stations eligible for polling and polls stations in a round-robin manner to give each station guaranteed channel access.

CP *contention period* is a reoccurring time in which stations compete for channel access using the CSMA/CA scheme. Stations sense a channel and wait until the channel is free before transmitting. If the channel is busy, the station will backoff for a random period before reattempting channel access.

CFP *contention free period* is a reoccurring time in which stations are polled by an access point's point coordinator and provided guaranteed channel access for a specified time. Stations with delay sensitive traffic (voice and video) can determine an upper bound on packet latency because stations are guaranteed a transmission opportunity after receiving a poll from a point coordinator.

HOL *head-of-line*, the front of a FIFO queue where packets are equeued for transmission.

AIFS *arbitration inter frame spacing* is a variable period of time a station must wait before transmitting packets for each access category (LP or HP). To provide HP packets higher priority to the channel, the AIFS period for HP classified packets is shorter than the AIFS period for LP classified packets.

STA *station* is a mobile node other than an access point.

α value between 0 and 1 that identifies the ratio of the time spent in the CFP to the total time spanned by a 802.11 superframe.

σ *slot time* is a unit of time equal to the sum of the RxTx turnaround time (time for a station to switch from receive to transmit mode), the channel sensing time, the channel propagation delay, and the MAC processing time.

III. Simplified IEEE 802.11e Hybrid MAC Model

The IEEE 802.11e MAC standard provides distributed service differentiation or QoS by employing a priority system. Network traffic is classified into four different priority levels or ACs. Nodes maintain separate queues for each AC and packets at the head-of-line (HOL) of each queue contend for channel access using AC-specific parameters, namely customized back-off and channel sensing durations [3]. Such a mechanism facilitates differentiated QoS where high priority, performance sensitive traffic, such as voice and video applications, will experience less delay and greater throughput, compared to low priority traffic (e.g., FTP and SMTP transfer). In this paper, we analyze the 802.11e MAC protocol and propose a Hybrid-MAC model that resembles the 802.11e MAC in most essential respects. Our MAC model provides us with an abstraction of the essential features of 802.11e MAC, while avoiding complex details. We believe that the insights obtained by using our model are applicable to the 802.11e scenario and can inform future standard evolution.

In our system, applications are classified into two priority levels or ACs and each node maintains two queues, namely a high priority traffic (HP) queue and a low priority traffic (LP) queue. Our model can be generalized to incorporate more ACs. Traffic assigned to the HP category is delay sensitive whereas LP traffic is delay tolerant. The network can operate in both contention and contention-free phases and these phases alternate periodically. During the contention phase, a node with packets to transmit will contend for channel access using the standard CSMA/CA algorithm [6]. QoS differentiation is enforced by allowing packets in the HP queue preferential channel access by enabling the interface to sense the channel prior to data transmission for a shorter period of time (AIFS) and also to back-off for a shorter duration, when faced with a collision or busy channel signal, than the packets must wait in the LP queues. This mechanism is similar to the IEEE 802.11e preferential channel access scheme. We assume that nodes are transmitting to an access point (AP) that can invoke the contention-free period by issuing a poll request to one or more nodes. These polled nodes can then transmit without any contention during the contention-free period. Thus, our protocol is very similar to the 802.11e Hybrid Coordination Function (HCF), with the contention period corresponding to 802.11e's random access or enhanced distributed coordination function (EDCF) functionality and the contention-free period corresponding to the 802.11e polled access or HCCA functionality. A diagram showing an example scenario involving communication between an AP and one node is illustrated in Figure 1.

IV. System Model

Our selected system model is a Basic Service Set (BSS) of N low priority and M high priority traffic flows. We assume that each flow is generated by a node which we refer to as a STA (station) as done in the 802.11 standard. During the CP, each STA uses the basic access mechanism only. That is, no STA is assumed to be hidden from another STA and the RTS/CTS mechanism is not employed. During the CFP, the M high priority traffic STAs are placed in a circular queue and are polled sequentially by the PCF as discussed shortly below. In our simulation, each STA is assumed to have a single IEEE 802.11b transceiver with an omni-directional antenna.

The PCF implements two periods of channel access in a duration of time referred to as the "superframe": (i) a contention free period (CFP) and (ii) a contention period (CP). The proportion of time allocated to each period within a superframe is not defined by the standard; however, the length of time allocated to the CP must be at least long enough to accommodate the transmission of one MAC Service Data Unit (MSDU) with a maximum frame length of 2304 bytes. The period of a superframe is delimited by a beacon frame transmission. The beacon is transmitted by the designated access point (AP) within a Basic Service Set (BSS) and carries with it protocol related parameters that are used by STAs to synchronize local timers and learn when the following beacon frame will be transmitted.

Synchronized data exchange within the CFP is accomplished by polling STAs. The polling process is coordinated by the PCF implementation within an AP. When the CFP begins, the AP waits a brief duration of time known as a short interframe space (SIFS) which serves as a delay between beacon, data, acknowledgement, and end frames that are transmitted during the CFP. The value of SIFS varies by the particular 802.11 standard implemented by a transceiver. For 802.11a, b, and g, the values are 16, 10, and 10 μs, respectively. After waiting an initial SIFS time period, the AP commences with polling by transmitting a Data/CF-Poll frame to the first STA in a polling list. Data/CF-Poll frames serve a dual purpose by piggybacking data carried by the AP which, in an infrastructure mode network, is attached to a wired network via a wired Ethernet interface. The Data/CF-Poll frame polls the receiving STA while simultaneously carrying higher layer datagrams originating from another STA within a BSS or a device external to a BSS via a wired LAN. The collision avoidance (CA) mechanism of CSMA/CA cannot guarantee collisions will not occur. A collision can occur, for example, if two STAs compute exactly the same backoff time after detecting a channel idle for a DCF interframe space duration (DIFS) and then transmit a MPDU when the backoff timer matures. To determine if a transmission resulted in a collision, each data frame (MPDU) must be acknowledged through the transmission of an ACK frame sent by the STA receiving a data frame. If a sending STA does not receive a corresponding ACK after waiting a SIFS period, the sending STA concludes a collision occurred and will repeat the transmission. DIFS values for 802.11a, b, and g are 34, 50, and either 28 or 50 μs, depending on slot time, respectively. In IEEE 802.11g, the slot time can be either 9 μs if no legacy 802.11b STAs are present in the BSS, or 20 μs if the BSS has a mix of 802.11b and 802.11g STAs. DIFS is a function of SIFS and is computed according to

$$DIFS = SIFS + 2\sigma \tag{1}$$

where σ is the slot time defined to be twice the maximum propagation time τ. The slot time is therefore an amount of time a STA waits to determine if another STA has accessed

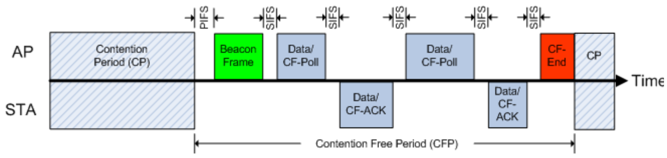

Fig. 1. Example scenario involving communication between the point coordinator and one station during the contention free period.

the channel at the start of the previous slot. Slot time values for 802.11a and b are 9 and 20 μs, respectively, for a PHY that uses a Direct Sequence Spread Spectrum (DSSS) modulation technique and 50 μs for a PHY that uses a Frequency Hopping Spread Spectrum (FHSS) transmission method. Acknowledgement frames may also piggyback data originating from a receiving STA and intended for another STA in the BSS or an external device. If the point coordinator fails to receive a response from a polled STA within a PCF interframe space (PIFS) period of time, the PCF will move on and poll the next STA in its polling list. PIFS is also a function of SIFS and is computed according to

$$PIFS = SIFS + \sigma \tag{2}$$

and thus the PIFS values for 802.11a, b, and g are 25, 30, and either 19 or 30 μs, respectively. The PIFS duration also serves as a gap between the CP and CFP. From (1) and (2) we have the following inequality

$$SIFS < PIFS < DIFS \tag{3}$$

which prevents the PCF from transmitting a poll frame in between a Data/CF-Poll and Data/CF-ACK transaction.

The point coordinator subsystem residing in an AP will continue to poll STAs in its polling list until the CFP duration expires, at which time a special CF-End frame is transmitted by the PCF to mark the end of the CFP.

V. Modeling Throughput

Our analytical model for overall system throughput is a dimensionless multivariable function S of N, M, p, and α,

$$S = S(N, M, p, \alpha) \tag{4}$$

where p is the probability of a successful frame transmission and α is a value between 0 and 1 that identifies the ratio of the time spent in the CFP to the total time spanned by a superframe which forms a repeating interval of contention and contention free time periods,

$$\alpha = \frac{CFP}{CFP + CP} \tag{5}$$

As α tends toward 0, the BSS reverts to a contention only based environment where the point coordinator is not used to poll STAs. With a non-zero α, dimensionless throughput S becomes a weighted sum of time spent in the CP and the CFP,

$$S(N, M, p, \alpha) = (1-\alpha)S_{CP} + \alpha S_{CFP} \tag{6}$$

We define S_{CP} and S_{CFP} as dimensionless throughput for each respective period,

$$S_{CP} = \frac{\bar{U}_{CP}}{\bar{I}_{CP} + \bar{B}_{CP}} \tag{7}$$

$$S_{CFP} = \frac{\bar{U}_{CFP}}{\bar{B}_{CFP}} \tag{8}$$

The definition of S_{CP} from [1] is given by equation (7) where $\bar{U}_{CP}$ is the average duration of time useful data is received by a STA during the CP, $\bar{I}_{CP}$ is the average duration of time the channel remains idle during the CP, and $\bar{B}_{CP}$ is the average duration of time the channel is busy transmitting data and the overhead bits incurred by the data, and time taken handling collisions. Equation (7) is then a dimensionless quantity between 0 and 1 that represents throughput efficiency as the ratio of time the channel is used for sending useful data to total time during the contention period. S_{CFP} is similar to S_{CP}, but does not include the idle term in the denominator since it is assumed the channel is never idle during the CFP. The definitions of $\bar{U}_{CP}$, $\bar{I}_{CP}$, and $\bar{B}_{CP}$ are adopted from [1], with the slight modification that the total STA count N in [1] has been replaced by (N+M), that is

$$\bar{U}_{CP} = \frac{(N+M)Tp}{(1-p)\left[1-(1-p)^{N+M}\right]} \tag{9}$$

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$\bar{I}_{CP} = \frac{\sigma}{1-(1-p)^{N+M}} \quad (10)$$

$$\bar{B}_{CP} = \frac{T_s}{(1-p)^{N+M}} \quad (11)$$

where T_s is the time spent sensing the channel during a successful frame transmission and T is the time spent transmitting useful data in the CP. Substituting (9), (10), and (11) into (7), we obtain

$$S_{CP} = \frac{(N+M)Tp(1-p)^{N+M-1}}{T_s + (\sigma + T_s)(1-p)^{N+M}} \quad (12)$$

The expression for T_s is given by

$$T_s = DIFS + \frac{H+P}{R} + SIFS + \frac{ACK}{R} + 2\tau \quad (13)$$

Our derivation of S_{CFP} proceeds in a similar way. Let q represent the probability a STA has a non-null data frame to transmit during the CFP. $\bar{U}_{CFP}$ is the average time spent during the CFP to transmit *useful* data. By useful data we mean data bits and not bits belonging to beacon, pure ACK, or CF-End frame. If we denote P_{CFP} as the number of data bits transmitted during the CFP, then

$$\bar{U}_{CFP} = \frac{P_{CFP}}{R} \quad (14)$$

where R is the fixed transceiver data rate. In our simulations we use the 802.11b maximum data rate of 11 Mbit/s. To derive an expression for $\bar{B}_{CFP}$, the time the channel is busy in the CFP during a successful polling transaction, we need to account for all the individual frame transmissions shown in Figure 1,

$$\bar{B}_{CFP} = \left(PIFS + \frac{CF_{Beacon}}{R} + \tau \right) + (N+M)\left[q^{(N+M)} \left(\begin{array}{l} 2SIFS + \\ \frac{H+P+CF_{Poll}}{R} + \\ 2\tau + \\ \frac{H+P+CF_{ACK}}{R} \end{array} \right) + (1-q)^{(N+M)} \left(\begin{array}{l} SIFS + \\ \frac{H+P+CF_{Null}}{R} + \\ \tau \end{array} \right) \right] \quad (15)$$

where CF_{Beacon}, CF_{Poll}, CF_{ACK}, and CF_{Null} are the lengths of the beacon, Data/CF-Poll, Data/CF-ACK, and CF-NULL frames, respectively. CF-NULL frames are transmitted by a polled STA if the STA does not have any pending data to send. τ is the propagation delay of the wireless LAN and H is the length of the header and frame check sequence (FCS) of an 802.11 frame. In our simulations we assume each 802.11 frame has a 30 byte header and a 4 byte FCS and thus H=34.

VI. Modeling Delay

Our analytical model for overall system delay is a dimensionless multivariable function D of N, M, p, and α,

$$D = D(N, M, p, \alpha) \quad (16)$$

Observe

$$0 < \frac{D_{ideal}}{D_{actual}} \le 1 \quad (17)$$

where D_{ideal} is the theoretical minimum delay a STA can experience in a superframe while D_{actual} is the true delay experienced. If we define D such that

$$D = \left(1 - \frac{D_{ideal}}{D_{actual}} \right) \quad (18)$$

Then $D \to 0$ as the actual delay approaches the ideal, and $D \to 1$ as actual delay diverges from the ideal.

We first consider delay incurred by the DCF. As described in [3], ideal delay in the CP can be expressed as the sum of ideal head-of-line (HOL) delay and ideal queuing delay,

$$D_{ideal} = D_{ideal}^{HOL} + D_{ideal}^{Queuing} \quad (19)$$

where D_{ideal}^{HOL} represents the minimum time required in the CP to transmit an 802.11 frame successfully, upon the first attempt, and is equal to T_s. Ideal queuing delay is given by the Pollaczek-Khinchine formula [3,13]

$$D_{ideal}^{Queuing} = \frac{\rho}{2\mu(1-\rho)}\left(1 + cv^2\right) \quad (20)$$

that describes the mean time a frame waits in queue to be serviced by the MAC, where the queue is modeled as a M/G/1 queue (a single server with frame arrivals having a Poisson distribution and service time having a general distribution). Total actual delay D_{actual} is modeled in [3] as the sum of (21) and an expression for the expected value of HOL delay which takes into account backoff delay,

$$E\left[D_{actual}^{HOL}\right] = T_s + \beta\left[\frac{CW_{min}}{2\left(1-(1-P_s)^{r_{max}+1}\right)}\right] \left[\frac{P_s\left(1-\left(2(1-P_s)\right)^{r_{max}+1}\right)}{1-2(1-P_s)} - 1 - (1-P_s)^{r_{max}+1}\right] + T_s\left[\frac{1-P_s}{P_s}\right]\left[\frac{(1-P_s)^{r_{max}}(-P_s r_{max} - 1) + 1}{1-(1-P_s)^{r_{max}+1}}\right] \quad (21)$$

where β is the average physical time between two decrements of the backoff counter, CW_{min} is the minimum contention window size, $P_s = (1-p)^{M+N-1}$ is the probability a STA's frame transmission is successful, and r_{max} is the maximum number of retransmissions permitted. In our simulation, CW_{min} is set to 2^4 and CW_{max} is set to 2^{10} which are the values used by a PHY that employs a Frequency-hopping spread spectrum (FHSS) method of transmitting radio signals. In addition, r_{max} is defined as

$$r_{max} = \log_2\left(CW_{max}/CW_{min}\right) \quad (22)$$

D_{ideal}^{HOL} is (21) without any backoff delay,

$$D_{ideal}^{HOL} = T_s \quad (23)$$

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Considering now the PCF, each STA has an opportunity to transmit when polled while the CFP is in progress. If the maximum predetermined duration of the CFP in a given superframe expires before every STA has been polled, STAs that were not given an opportunity are more likely to be polled in the following CFP as the PC uses a circular queue to schedule station polling. Let Ψ represent the expected value of the size of a frame transmitted by a polled STA during the CFP and ψ represent the size of the body of data within this frame,

$$\Psi = 34 + E[\psi] \tag{24}$$

Assuming the length of data in frames transmitted during the CFP is uniformly distributed, the total time for one CFP is given by T_{CFP},

$$T_{CFP} = \frac{CF_{Beacon}}{R} + \frac{(N+M)(\Psi_{PC}+\Psi_{STA})}{R} + \left[2(N+M)+1\right]SIFS + \frac{CF_{End}}{R} + (N+M)\tau \tag{25}$$

In our simulation, CF_{Beacon} and CF_{End} are set to 180 and 20 bytes, respectively. Let D_{CFP} represent the average time a frame must wait at the head-of-line once the CFP begins,

$$D_{CFP} = \frac{CF_{Beacon}}{R} + \left(\frac{N+M}{2}-1\right)\frac{(\Psi_{PC}+\Psi_{STA})}{R} + (N+M-1)SIFS + \left(\frac{N+M}{2}\right)\tau \tag{26}$$

From (19), (20), (23), and (26) we now have

$$D_{ideal} = T_s + \frac{\rho}{2\mu(1-\rho)}\left(1+cv^2\right) + \frac{CF_{Beacon}}{R} + \left(\frac{N+M}{2}-1\right)\frac{(\Psi_{PC}+\Psi_{STA})}{R} + (N+M-1)SIFS + \left(\frac{N+M}{2}\right)\tau \tag{27}$$

Accounting for backoff delay, (27) is modified to give D_{actual},

$$D_{actual} = T_s + \beta\left[\frac{CW_{\min}}{2\left(1-(1-P_s)^{r_{\max}+1}\right)}\right] \left[\frac{P_s\left(1-\left(2(1-P_s)\right)^{r_{\max}+1}\right)}{1-2(1-P_s)} - 1 - (1-P_s)^{r_{\max}+1}\right] + T_s\left[\frac{1-P_s}{P_s}\right]\left[\frac{(1-P_s)^{r_{\max}}(-P_s r_{\max}-1)+1}{1-(1-P_s)^{r_{\max}+1}}\right] + \frac{\rho}{2\mu(1-\rho)}\left(1+cv^2\right) + \frac{CF_{Beacon}}{R} + \left(\frac{N+M}{2}-1\right)\frac{(\Psi_{PC}+\Psi_{STA})}{R} + (N+M-1)SIFS + \left(\frac{N+M}{2}\right)\tau \tag{28}$$

A plot of our analytical expressions for dimensionless throughput and normalized delay based on our derivations is shown in Figures 2 and 3. Parameter values used in both plots are given in Network Simulation Parameters .

TABLE I. NETWORK SIMULATION PARAMETERS

Data Rate (R)	2 Mbps
Frame Data Size	Uniformly distributed in range [0,2312] B
PHY Header Size	24 Bytes
MAC Header	34 Bytes
Slot Time σ	50 µs
SIFS	28 µs
Prop. Delay τ	10 µs
CW_{min}, CW_{max}	2^4, 2^{10}
Beacon Size	90 Bytes
Packet inter-arrival rate (λ)	$2*10^{-8}$

Figures 2 and 3 show a surface plot that quantifies the relationship between collision probability, number of HP users, and the effect these parameters have on system delay and throughput, respectively. In Figure 2 we see that, as the number of HP stations increases, a saturation condition at normalized delay D = 1 is attained with lower values of collision probability p. Collision probability p is defined as the probability a given frame transmission attempt is unsuccessful due to a collision occurring in the CP. Looking at Figure 2, one can see that for a small number of HP stations, the directional derivative *dD/dp* is less than it is for a large number of HP stations. Because the rate of change in delay increases faster with respect to station count as collision probability increases, a saturation condition will arise sooner in a BSS with many high priority traffic stations if stations begin to experience a greater number of collisions in the contention period. Similarly, in Figure 3 we see how small changes in collision probability can greatly affect throughput as the HP station count increases. We also see the appearance of an optimal throughput contour along the maxima of the surface *S*.

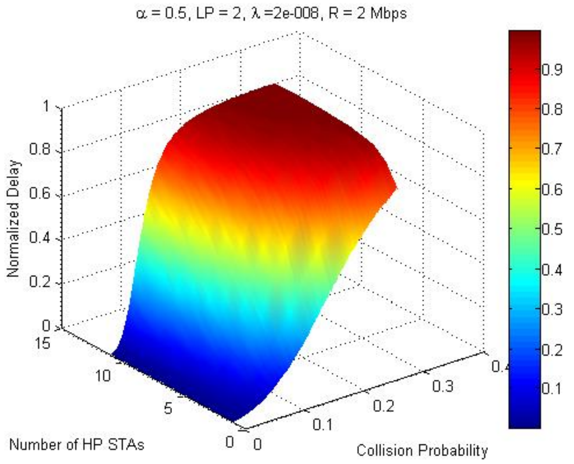

Fig. 2. Normalized delay surface plot $D = D(HP, p)$.

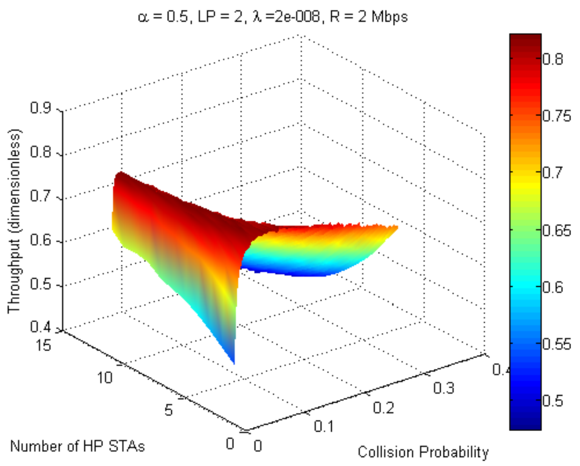

Fig. 3. Dimensionless throughput surface plot $S = S(HP, p)$.

VII. Conclusion

In this paper, the authors have presented analytical expressions to model the throughput and delay of a hybrid MAC scheme akin to IEEE 802.11e. The authors have extended the work of Khalaf and Rubin in [3] and have made the following new contributions:

- the development a dimensionless expression for throughput as a function of the CFP ratio α for two access categories (HP and LP),
- the development a dimensionless expression for delay as a function of α for two access categories,
- expressing dimensionless delay as a function that ranges from 0 to 1, based on ideal and actual delay times,
- expressions for ideal and actual delay times, and
- a code to visualize throughput and delay as a function of HP STA count and collision probability for selected values of α.

The authors consider these contributions are significant because the concepts and expressions can be used in an optimization framework by an access point to dynamically adjust α, in real-time, as HP STAs associate and dissociate from the access point. For example, an access point could repeatedly estimate the collision probability based on past datagram retransmission statistics and select a value for α that minimizes delay and maximizes throughput based on current HP STA count. The authors have shown the value of α has a significant effect on system performance with respect to throughput and delay. An increasing number of HP users create higher contention in the CP phase leading to longer backoff time and thereby a drop in throughput and an increase in delay. A future direction will be to introduce a numerical Min-Max optimization strategy to find the optimal value of α that minimizes delay while maximizing throughput, and to simulate the use of an optimal α in a discrete event network simulator such as ns-3, OPNET Modeler, or QualNet.

References

[1] "Medium Access Control (MAC) Quality of Service (QoS) Enhancements," IEEE 802.11e Working Group, New York, P802.11e/D13.0, January 2005.

[2] "IEEE Standard for Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks-Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE, IEEE Std 802.11-1997, June 12, 2007.

[3] R. Khalaf and I. Rubin, "Throughput and Delay Analysis in Single Hop and Multihop IEEE 802.11 Networks," in *3rd International Conference on Broadband Communications, Networks and Systems, 2006 (BROADNETS 2006)*, San Jose, CA, 2006, pp. 1-9.

[4] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Delay Analysis of IEEE 802.11 in Single-Hop Networks," in *Proceedings of the 11th IEEE International Conference on Network Protocols (ICNP)*, Washington, DC, November 04 - 07, 2003, p. 146.

[5] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535-547, March 2000.

[6] O. Tickoo and B. Sikdar, "Modeling Queueing and Channel Access Delay in Unsaturated IEEE 802.11 Random Access MAC Based Wireless Networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 878-891, August 2008.

[7] T. Suzuki and S. Tasaka, "Performance Evaluation of Integrated Video and Data Transmission with the IEEE 802.11 Standard MAC Protocol," in *Global Telecommunications Conference (GLOBECOM '99)*, vol. 1B, Rio de Janeireo, Brazil, 1999, pp. 580-586.

[8] O. Tickoo and B. Sikdar, "Queueing Analysis and Delay Mitigation in IEEE 802.11 Random Access MAC Based Wireless Networks," in *Proceedings of IEEE INFOCOM*, Hong Kong, P. R. China, 2004, pp. 1404-1413.

[9] L. Yun, L. Ke-Ping, Z. Wei-Liang, and W. Chong-Gang, "Analyzing the Channel Access Delay of IEEE 802.11 DCF," in *IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 5, St. Louis, MO, 2005, pp. 3001-3005.

[10] S. Mangold, S. Choi, G. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS Support in Wireless LANs," *IEEE Wireless Communications*, vol. 10, no. 6, pp. 40-50, December 2003.

[11] A. Lindgren, A. Almquist, and O. Schelén, "Quality of Service Schemes for IEEE 802.11 Wireless LANs: An Evaluation," *Mob. Netw. Appl.*, vol. 8, no. 3, pp. 223-235, June 2003.

[12] A. Köpsel, J.-P. Ebert, and A. Wolisz, "A Performance Comparison of Point and Distributed Coordination Function of an IEEE 802.11 WLAN in the Presence of Real-Time Requirements," Waseda, Tokyo, Japan, October 23-26, 2000.

[13] I. Mitrani, "Modelling of Computer and Communication Systems," Cambridge Computer Science Texts ed.: Cambridge University Press, 1987.