Title
Time-Based Persistence in Channel-Access Protocols with Carrier Sensing

Authors
Garcia-Luna-Aceves, J.J.
Thompson, Spencer
Stern, Joshua

Publication Date
2017-12-10

Peer reviewed

Time-Based Persistence in Channel-Access Protocols with Carrier Sensing

J.J. Garcia-Luna-Aceves*,†, Spencer Thompson*, Joshua Stern*
* Jack Baskin School of Engineering, UC Santa Cruz, Santa Cruz, CA 95064
† Xerox PARC, Palo Alto, CA 94304
jj@soe.ucsc.edu, {joastern, spkthomp}@ucsc.edu

***Abstract*—Prior work on persistent and non-persistent transmission strategies of CSMA and CSMA/CD indicated that no persistence provides better performance; however, this result applies only to a specific approach to persistence. We introduce time-based persistence in which a node with a packet to send that finds the channel busy persists for a limited amount of time, and provide a simple unifying analysis of the impact of time-based persistence in channel-access protocols that use carrier sensing and operate in ad-hoc wireless networks. We focus on CSMA with priority acknowledgments (ACK) and CSMA with collision detection (CSMA/CD) and ACKs. Our analysis takes into account the effect that receive-to-transmit turnaround times have on performance, and shows that CSMA and CSMA/CD with time-based persistence can attain the same throughput values relative to a non-persistent strategy.**

I. Introduction

Carrier sense multiple access (CSMA) [9] is widely used for the sharing of radio channels in ad-hoc networks, and many approaches have been proposed and implemented to improve on the performance of CSMA to either cope with hidden-terminals (e.g., CSMA with collision avoidance or CSMA/CA [3]) or make better use of the channel (e.g., CSMA with collision detection or CSMA/CD [10]). However, very few of the approaches consider transmission policies other than non-persistence in the presence of carrier, which is due to the added complexity in analyzing transmission policies dependent on the present state of the channel.

Section II reviews prior work on channel-access schemes that address transmission policies in which nodes with packets to be sent persist attempting to transmit data packets or signaling packets based on current channel conditions. The most widely known treatment of persistent transmission policies in channel-access protocols based on carrier sensing dates back to the seminal work by Kleinrock and Tobagi [9], which introduced the concept of carrier sensing itself. Kleinrock and Tobagi defined p-persistence as a transmission policy in which a node decides to transmit with probability $1/p$ if the channel is found to be idle or after carrier is finally detected to be down after the node with a packet to transmit detects carrier. The 1-persistent case is widely known and was used in Ethernet [10] and renders much lower throughput that a non-persistent transmission policy used in the same channel-access protocol.

The two main contributions of this paper are: (a) the description of a persistent-transmission policy for contention-based channel-access protocols based on carrier sensing that can render similar performance than the traditional non-persistent transmission policy; and (b) the presentation of a unifying approach for the analysis of protocols based on time-based persistent-transmission discipline.

Section III presents our approach to persistent transmissions in the context of CSMA and CSMA/CD. The persistent-transmission policy we introduce is based on a *persistence time* after carrier is detected by a node with a packet to transmit, and is applicable to any contention-based channel-access protocol that uses carrier sensing. For the case of CSMA/CD, we assume that self-interference cancelation (SIC) approaches at the physical layer enable collision detection over a wireless channel in real time. While the technology for SIC is still evolving, recent results [8] indicate that collision detection at the medium-access control layer may be feasible soon. In each protocol, a node with a packet to send that finds the channel idle simply transmits as in the non-persistent version. However, a node with a packet to send that finds the channel busy persists for a finite period of time.

Section IV presents the throughput of non-persistent CSMA with priority ACKs and CSMA/CD, and Section V obtains the throughput of the time-persistent variants of the same two channel-access methods as a function of the persistence time ρ. In a nutshell, the model we use defines a simple three-state embedded Markov chain for each channel-access protocol based on the state of the channel when a new arrival occurs. In addition, in contrast to all prior modeling work reported to date, we address the use of priority acknowledgements (ACK) as part of the protocols and the impact that the receive-to-transmit turnaround times have on the performance of the protocols. These aspects of our analysis are important, because ACKs are a necessity in any practical channel-access protocol operating in a wireless network, and turnaround latencies of today's half-duplex radios are not negligible compared to the propagation delays of ad-hoc wireless networks.

Section VI provides numerical results for scenarios in which CSMA and CSMA/CD are applicable. The results of our model using practical values of relevant parameters show that using a time-based persistence policy is a viable approach to contention-based channel-access protocols based on carrier sensing compared to a non-persistent policy. Section VII presents our conclusions and summarizes new research avenues enabled by our results.

II. Related Work

Tobagi and Kleinrock introduced CSMA [9] and were the first to address the limitations of a non-persistent transmission strategy in the context of carrier sensing. In a non-persistent channel-access protocol based on carrier sensing, a node that detects carrier when it has a packet to transmit backs off immediately long enough for the current transmission to succeed. This over-reaction to channel utilization renders low throughput values when the channel load is light, because some of the few nodes with traffic to send may wait too long before attempting to transmit again if they find the channel busy in their first attempt. For this reason, Kleinrock and Tobagi introduced the notion of a p-persistence transmission policy, with which a node with a packet to transmit that finds the channel idle transmits with probability $1/p$ and a node with a packet to transmit that finds the channel bust waits until the channel becomes idle again and the transmits with probability $1/p$. This approach captures the essence of the most basic persistence policy, which is 1-persistence, in which a node with a packet to send that finds the channel busy simply waits for the end of carrier and transmits after that occurs. This is the approach taken in the original CSMA/CD scheme used in Ethernet, which works well at light loads.

The major limitations with the approach to persistence introduced by Kleinrock and Tobagi are that transmission opportunities may be wasted when nodes detect the channel idle, and too many collisions may be caused when nodes detect the channel busy. The first problem stems from the fact that nodes determine whether or not to transmit based on a probability value that is independent of whether the channel is found to be idle or busy. The second problem results from the fact that all nodes that have packets to send and detect a busy channel wait until the channel becomes idle again to make a transmission decision with the same probability value.

Surprisingly, no concerted effort has been mounted on the analysis of the impact that different persistence transmission policies have on the performance of channel-access schemes. To the best of our knowledge, most of the very few subsequent studies addressing persistence in channel-access methods using carrier sensing assumed the same approach to persistence defined by Kleinrock and Tobagi [11], [12], [13]. The notion of using limited persistence in channel-access protocols based on collision avoidance methods was introduced in [5]. In a nutshell, nodes that find the channel idle transmit their packets, and a node with a packet to send that finds the channel busy waits for a limited amount of time proportional to the time it takes to send a request-to-send (RTS) packet. If the channel continues to be idle after that time, the node backs off; otherwise, the node transmits. The limitation with the limited-persistence approach [5] in practice is that it favors nodes with local packet arrivals closer to the end of the current busy period over nodes with local packet arrivals that occur earlier in the busy period, which results in longer channel-access delays.

From the modeling perspective, none of the previous analysis of channel-access methods with persistence take into account the impact of turnaround times, and the analysis of p-persistent CSMA [9] does not take into account the use of acknowledgments and relies on the unrealistic assumption that an ideal channel is used to transmit acknowledgments in zero time without interference.

III. Time-Based Persistence in Channel-Access Protocols Using Carrier Sensing

A. Basic Approach for Time Persistence

The approach we propose for time persistence requires that nodes monitor the channel for the presence of carrier continuously; however, this is also the case in any channel-access protocol based on carrier sensing and virtual carrier sensing. We denote by ρ the persistence time period a node allows when it has a packet to send and the channel is busy. The local time when the node detects carrier is denoted by T_c and the local time when the node receives a local packet to send is denoted by T_p. A node that has obtained a new local packet to send carries out the following steps as part of the channel-access protocol:

1) If the channel is idle, transition to transmit mode and transmit the packet
2) If the channel is busy (node is in REMOTE state) then:
   (a) Compute $TD = T_p - T_c$;
   (b) Enter BACK-OFF state if $TD \geq \rho$
   (c) Enter PERSIST state if $TD < \rho$

We describe how CSMA with priority ACKs and CSMA/CD with ACKs can be modified to account for time persistence. The variables T_p and T_c are maintained separately from the state machine shown for the channel-access protocol, and TD is computed as soon as a local packet is ready for transmission. Time T_c is reset when the channel becomes idle again, taking into account the fact that priority ACKs follow a successful data packet and a virtual carrier of $\omega + \tau$ (turnaround time plus a maximum propagation delay) must be observed. Time T_p and TD are reset when the node transitions out of the REMOTE state. Once in the PERSIST state, a node transmits either a data packet or a request-to-send (RTS) packet as soon as the channel becomes idle.

B. Time-Persistent CSMA with Priority ACKs

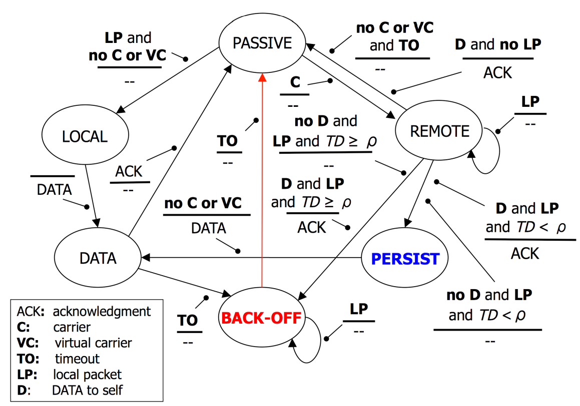

Fig. 1. Operation of time-persistent CSMA with priority ACKs

Fig. 1 illustrates the operation of time-persistent CSMA with priority ACKs using a state machine. Our description assumes that the channel-access protocol receives at most one packet to send at any given time, and does not receive a new packet until it is done processing the current packet.

Time-persistent CSMA operates just like non-persistent CSMA when a node has a packet to transmit and the channel is idle. The main difference between the transmission policies of non-persistent CSMA and time-persistent CSMA is the addition of the PERSIST state. The LOCAL state is used to emphasize the need for a radio to transition from listen to transmit mode, which involves a latency of ω seconds during which the node is unable to listen to the channel.

A node that receives a local packet to send and detects no carrier or virtual carrier transitions to the LOCAL state and transmits its packet once its radio is in the transmit mode. The node then transitions to the DATA state and waits for an ACK from the receiver. The node transitions to the PASSIVE state if an ACK is received or to the BACK-OFF state if the ACK is not received within a timeout period in order to schedule a retransmission.

A node that detects carrier or virtual carrier while in the PASSIVE state transitions to the REMOTE state and remembers the value of T_c (the local time when carrier was detected). If the node has no local packet (LP) to send and carrier goes down without receiving a data packet intended for itself (indicated as "no DATA to self" in the state machine), the node simply goes back to the PASSIVE state silently. If the node receives a data packet for itself from a transmitter (shown as "DATA to self" in the state machine) and the node has no local packet to send, it sends an ACK and goes to the PASSIVE state. The node remembers the arrival of a local packet to send and the local time when that occurs (T_p) while it waits to decode an ongoing transmission while in the REMOTE state.

A node in the REMOTE state with a local packet to send that decodes a remote packet sent to itself sends an ACK to the sender and transitions to the PERSIST state if $TD < \rho$, or transitions to the BACK-OFF state if $TD \geq \rho$. Similarly, a node with a packet to end in the REMOTE state that cannot decode a remote data packet for itself transitions to the PERSIST state if $TD < \rho$, or to the BACK-OFF state if $TD \geq \rho$.

Once in the PERSIST state, the node waits until the end of carrier or virtual carrier is detected, transmits its data packet, and transitions to the DATA state. A node in the BACK-OFF state computes a random back-off time after which it transitions to the PASSIVE state and attempt to transmit as needed if there is no carrier or virtual carrier detected.

C. *Time-Persistent CSMA/CD*

The state machine needed to represent the operation of time-persistent CSMA/CD is almost the same as the one for time-persistent CSMA. The only difference is that a node transmitting a data packet uses collision detection while sending its packet. If a a collision is detected the node aborts its transmission and transitions to the BACK-OFF state.

Recent developments on technologies for self-interference cancellation at the physical layer are likely to enable practical collision detection in wireless networks soon.

IV. Throughput of Non-Persistent Protocols

A. *Model and Assumptions*

We assume the same traffic model first introduced by Kleinrock and Tobagi [9] to analyze CSMA with priority ACKs and CSMA/CD. This model is only an approximation of the real case; however, our analysis provides a good baseline for the comparison of the various channel-access protocols and the relative benefits of the joint use of collision avoidance and detection compared to other techniques.

There is a large number of stations that constitute a Poisson source sending data packets to the the channel with an aggregate mean generation rate of λ packets per unit time. We assume the use of priority acknowledgments (ACK) in all protocols, because they are needed in practice to account for transmission errors not due to multiple-access interference.

Each node is assumed to have at most one data packet to sent at any time, which results from the MAC layer having to submit one packet for transmission before accepting the next packet. The hardware is assumed to require a fixed turn-around time of ω seconds to transition from receive-to-transmit or transmit-to-receive mode for any given transmission to the channel. According to the parameters assumed in IEEE 802.11 DCF, this value may be comparable to or larger than the propagation delay τ in a wireless local area network (WLAN).

The transmission time of a data packet is δ and the transmission time for an ACK is α. For the case of CSMA/CD, it is assumed that the time it takes for a node to detect collision with its own transmission and send a jamming bit sequence lasts η seconds, where $\eta \ll \gamma$, given that it is simply the time needed to identify the difference between the transmission of a node and the signal if decodes after SIC, plus the transmission of a short bit sequence that has to be larger than the error-checking field of a packet (e.g., 48 bits).

We assume that, when a node has to retransmit a data packet it does so after a random retransmission delay that, on the average, is much larger than the time needed for a successful transaction between a transmitter and a receiver and such that all transmissions of data packets can be assumed to be independent of one another. The channel is assumed to introduce no errors, so multiple access interference (MAI) is the only source of errors. Nodes are assumed to detect carrier and collisions perfectly. To further simplify the problem, we assume that two or more transmissions that overlap in time in the channel must all be retransmitted (i.e., there is no power capture by any transmission), and that any packet propagates to all nodes in exactly τ seconds.

The protocols are assumed to operate in steady state, with no possibility of collapse, and hence the average channel utilization of the channel is given by [9]

$$S = \frac{\overline{U}}{\overline{B} + \overline{I}}. \tag{1}$$

where $\overline{B}$ is the expected duration of a busy period, defined to be a period of time during which the channel is being utilized; $\overline{I}$ is the expected duration of an idle period, defined as the time interval between two consecutive busy periods; and $\overline{U}$ is the average time during a busy period that the channel is used for transmitting user data successfully. The throughput S is simply the percentage of an average system cycle in time that the system is used to transmit data successfully, where an average system cycle is the average time that the system takes to go from the start of one idle period to the start of the next idle period.

B. CSMA with Priority ACKs

The original throughput results for non-persistent CSMA by Kleinrock and Tobagi [9] assume an ideal secondary channel over which ACKs are sent in 0 time. We consider the throughput of non-persistent CSMA with priority ACKs. Figure 2 illustrates the transmission periods in non-persistent CSMA with priority ACKs.

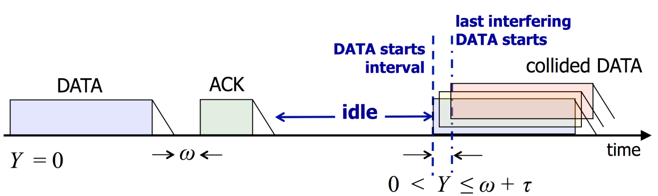

Fig. 2. Transmission periods in non-persistent CSMA with priority ACKs

Theorem 1: The throughput of non-persistent CSMA with priority ACKs is

$$S_{CSMA} = \frac{\delta}{\omega + \alpha + \tau + \frac{1}{\lambda} + e^{\lambda(\omega+\tau)}(\delta + \omega + 2\tau)} \quad (2)$$

Proof: The proof is presented in [6] and is similar to the proof in [14]. The key difference is the increase in the vulnerability period of a data packet caused by the turnaround times of length ω during which nodes are deaf. □

C. CSMA/CD

We obtain the throughput of non-persistent CSMA/CD assuming SIC is used to consider the use of priority ACKs (i.e., passive nodes defer until an ACK is transmitted after a successful data packet). Because nodes can listen to the channel while they transmit, nodes do not incur turnaround latencies as in CSMA. Fig. 3 illustrates the transmission periods for non-persistent CSMA/CD.

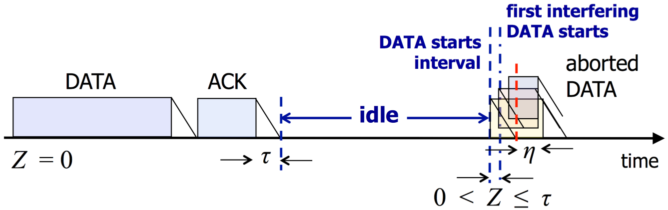

Fig. 3. Transmission periods in non-persistent CSMA/CD

Theorem 2: The throughput of non-persistent CSMA/CD is

$$S_{CSMA/CD} = \frac{\delta}{\delta + \alpha - \eta - \frac{1}{\lambda} + e^{\lambda\tau}(\eta + 2\tau + \frac{2}{\lambda})} \quad (3)$$

Proof: If a data packet does not collide with another transmission, the receiver sends an ACK without contention, and this takes $\delta + \alpha + 2\tau$ seconds. Given that each transmitter uses carrier sensing before transmitting a data packet, the probability that this occurs is simply $P_S = e^{-\lambda\tau}$.

On the other hand, if a data packet collides with others, then all the transmitters involved in the collision interval detect the collision, abort their transmissions, and send a jamming bit sequence. By assumption, the time needed to transmit the jamming bit sequence is η seconds. All nodes that interfere with the data packet that starts the collision interval receive the carrier from the first data packet in the interval in τ seconds after the start of the interval, and take η seconds to transmit a jamming pattern. Furthermore, those transmissions take τ seconds to propagate to all nodes.

On the other hand, as illustrated in Fig. 3, the node that started the collision interval detects a collision τ seconds after the first interfering data packet starts. Accordingly, the length of a collision interval is given by $Z + \tau + \eta + \tau$, where Z is a random variable representing the time between the arrival of the data packet that starts the collision interval and the arrival of the *first* data packet that causes a collision. A collision interval occurs with probability $1 - P_S = 1 - e^{-\lambda\tau}$.

The random variable Z varies from 0 to τ, and $Z = 0$ occurs when a data packet is successful. This is the case because it is not possible to have two or more arrivals of data packets into the channel exactly at the same time under the assumption that packet arrivals are Poisson distributed. Accordingly, the average length of a busy period equals

$$\begin{aligned}\overline{B} &= \overline{Z} + (1 - e^{-\lambda\tau})(\eta + 2\tau) + e^{-\lambda\tau}(\delta + \alpha + 2\tau) \\ &= \overline{Z} + \eta + 2\tau + e^{-\lambda\tau}(\delta + \alpha - \eta) \quad (4)\end{aligned}$$

For Z to last more than z seconds, it must be the case that no arrival occurs in the first z seconds of a collision interval, that is, $P(Z > z) = P\{\text{no arrivals in } [0, z]\} = e^{-\lambda z}$. Therefore, the cumulative distribution function of Z is

$$F_Z(z) = P(Z \leq z) = 1 - P(Z > z) = 1 - e^{-\lambda z} \quad (5)$$

Given $F_Z(z)$ and the fact that Z assumes non-negative values, the mean of Z can be computed as follows:

$$\overline{Z} = \int_0^\infty (1 - F_Z(t))dt = \int_0^\tau e^{-\lambda t}dt = \frac{1}{\lambda}\left(1 - e^{-\lambda\tau}\right) \quad (6)$$

Substituting $\overline{Z}$ in Eq. (4) we have

$$\overline{B} = e^{-\lambda\tau}\left(\delta + \alpha - \eta - \frac{1}{\lambda}\right) + \eta + 2\tau + \frac{1}{\lambda} \quad (7)$$

The average length of an idle period $\overline{I}$ in CSMA/CD is simply the average inter-arrival time of data packets into the channel, which equals $1/\lambda$. The average time period used to transmit useful data $\overline{U}$ is simply the useful portion of a successful busy period, i.e., $\overline{U} = \delta P_S = \delta e^{-\lambda\tau}$. Substituting the values of $\overline{U}$, $\overline{B}$, and $\overline{I}$ into Eq. (1) we obtain Eq. (3). □

V. Throughput of Time-Persistent Channel-Access Protocols

A. Model and Assumptions

To analyze the performance of channel-access protocols with time persistence we need to provide more granularity in the description of throughput, given that he utilization of the channel consists of *transmission periods* that can be classified based on the number of transmissions at the beginning of a transmission period. The length of time from the instant a node detects carrier to the present time during which the channel is still busy is denoted by T, and the persistence time is denoted by ρ. The rest of the assumptions are the same as in the previous section.

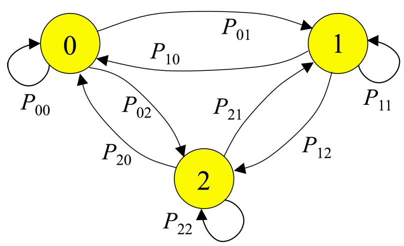

Fig. 4. Markov chain for time-persistence channel-access protocols

If we view channel utilization in terms of transmission periods, it follows from the operation of time-persistent channel-access protocols that the type of the next transmission period depends on the arrivals that take place during the persistent time ρ of the current transmission period.

If no arrivals occur during the ρ seconds during which nodes persist after detecting carrier in the current transmission period, then an idle transmission period follows. We call this TP_0 (transmission period of type 0) because no transmissions at the beginning of the next transmission period.

If one arrival occurs during the ρ seconds of persistence of the current transmission period, then the next transmission period has a single arrival, and similarly the next arrival after an idle period starts a transmission period with a single arrival. We call a transmission period that starts with a single transmission a TP_1 (transmission period of type 1).

If two or more arrivals occur during the persistent time ρ of the current transmission period, then the next transmission period starts with two or more transmissions. We call this type of transmission period a TP_2 (transmission period of type 2).

Given that the type of the next transmission period in the channel depends on the number of arrivals in the current transmission period, we can define the three-state Markov chain shown in Fig. 4. The chain then has one state for each type of transmission period that is possible under the time-persistence transmission policy we assume.

Sohraby et al. [11] also used a three-state Markov chain formulation to analyze the throughput of 1-persistent CSMA and 1-persistent CSMA/CD with no ACKs. Our model generalizes their earlier formulation.

We denote by π_i $(i = 0, 1, 2)$ the stationary probability of being in state i, i.e., that the system is in a type-i transmission period. The transition probability from state i to state j is denoted by P_{ij}. The average time spend in state i is denoted by T_i. Given our assumption that the protocols operate in steady state, we have a homogeneous Markov chain and we can define the throughput of the network to be the percentage of time in an average cycle that the channel is used to transmit data successfully, which is

$$S = \frac{\pi_1 \overline{U}}{\pi_0 T_0 + \pi_1 T_1 + \pi_2 T_2} \tag{8}$$

We can use the facts that the channel must be in one state at every instant and the channel must transition from one state to another state including itself with probability 1, plus the balance equations for state 1 and state 2 to express the state probabilities as functions of the transition probabilities. From the Markov state diagram in Fig. 4 we have the following four equations

$$\begin{gathered} \pi_1(P_{12} + P_{10}) = \pi_2 P_{21} + \pi_0 P_{01}; \\ \pi_0(P_{01} + P_{02}) = \pi_1 P_{10} + \pi_2 P_{20}; \\ \pi_0 + \pi_1 + \pi_2 = 1; \;\; P_{10} + P_{11} + P_{12} = 1 \end{gathered} \tag{9}$$

Because arrivals are Poisson, there can be no more than one arrival at any instant and hence $P_{02} = 0$. On the other hand, because the system is in equilibrium, there must be an arrival within a finite time once the channel is idle, and we have $P_{01} = 1$. In addition, the type of the next transition period that occurs is independent of whether the current transmission period is of type 1 or type 2, because is only a function of the number of arrivals during the persistence time; therefore,

$$P_{1j} = P_{2j} \qquad\qquad j = 0,\ 1,\ 2 \tag{10}$$

The state probabilities can then be obtained from the values of P_{01}, P_{02} and Eqs. (9) and (10) as functions of P_{10} and P_{11}:

$$\pi_0 = \frac{P_{10}}{1 + P_{10}};\ \pi_1 = \frac{P_{10} + P_{11}}{1 + P_{10}};\ \pi_2 = \frac{1 - P_{10} - P_{11}}{1 + P_{10}} \tag{11}$$

Making use of the fact that $P_{10} + P_{11} = 1 - P_{12}$ in the previous three equations we obtain

$$\pi_0 = \frac{P_{10}}{1 + P_{10}};\ \pi_1 = \frac{1 - P_{12}}{1 + P_{10}};\ \pi_2 = \frac{P_{12}}{1 + P_{10}} \tag{12}$$

Substituting Eqs. (12) in Eq. (8) we obtain the following expression of S as a function of transition probabilities P_{10} and P_{12}, $\overline{U}$, and the average times of each transmission period:

$$S = \frac{(1 - P_{12})\overline{U}}{P_{10} T_0 + (1 - P_{12}) T_1 + P_{12} T_2} \tag{13}$$

The length of T_0 in all time-persistent protocols is simply the average inter-arrival time of packets as in non-persistent protocols, that is, $T_0 = 1/\lambda$. P_{10} equals the probability that no arrivals occur during the ρ seconds of persistence during a TP_1 and hence $P_{10} = e^{-\lambda\rho}$. On the other hand, P_{12} equals the probability that two or more arrivals occur during the ρ seconds of persistence during a TP_1, which equals the probability of the complement of the event that zero or one arrivals occur in ρ seconds. Therefore, $P_{12} = 1 - (e^{-\lambda\rho} + \lambda\rho e^{-\lambda\rho})$.

Substituting the values of T_0, P_{10} and P_{12} in Eq. (13) we obtain

$$S = \frac{(1+\lambda\rho)\overline{U}}{\frac{1}{\lambda} + (1+\lambda\rho)T_1 + (e^{\lambda\rho} - (1+\lambda\rho))T_2} \tag{14}$$

The following sections obtain the values of T_1, T_2, and $\overline{U}$ for CSMA and CSMA/CD.

B. Time-Persistent CSMA with Priority ACKs

Figure 5 illustrates the transmission periods that may occur in time-persistent CSMA with priority ACKs. The figure illustrates a sequence of transmission periods and their lengths, which are indicated by the numbers 0, 1, and 2. As the figure shows, only a TP_1 can be successful.

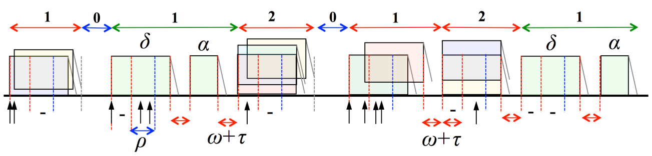

Fig. 5. Transmission periods in time-persistent CSMA with priority ACKs

Theorem 3: The throughput of time-persistent CSMA with priority ACKs is

$$S_{TCS} = \frac{\delta}{\frac{1}{1+\lambda\rho}e^{\lambda v}\left[\frac{1}{\lambda} + e^{\lambda\rho}(\delta+\omega+2\tau-\frac{1}{\lambda}(1-e^{-\lambda v}))\right] + C}$$

$$\text{where } \; v = \omega+\tau \; \text{ and } \; C = \omega+\alpha+\tau \tag{15}$$

Proof: We observe that new arrivals can occur with the first $\tau+\omega$ seconds of a transmission period of type 1 or type 2, because it takes τ seconds for the start of the first transmissions to propagate to all nodes, and a given node that perceives the channel being idle within a window of time of up to ω second before the first transmission of a transmission period and at most τ seconds after the start of the first transmission can collide with the first transmission. Accordingly, the actual length of a transmission period of type 1 or 2 is a function of the time between the first and the last transmission in the transmission period, which is a random variable Y that can assume values between 0 and $\tau+\omega$.

If the time period between the start of the the first and the last data packets in a collision interval equals y seconds, then there are no more packet arrivals in the remaining time of the vulnerability period of the first packet of the collision interval, i.e., $\omega+\tau-y$ seconds. Accordingly, $P(Y \leq y) = F_Y(y) = e^{-\lambda(\omega+\tau-y)}$. Therefore, given that Y assumes only non-negative values, the average value of Y equals

$$\overline{Y} = \int_0^{\infty}(1-F_Y(t))dt = \int_0^{\omega+\tau}\left(1-e^{-\lambda(\omega+\tau-t)}\right)dt$$
$$= \omega+\tau-\frac{1-e^{-\lambda(\omega+\tau)}}{\lambda} \tag{16}$$

A transmission period of type 2 starts with two or more transmissions; therefore, no success can occur in it and hence it must consist of overlapping packets that cannot be decoded by the intended receivers. Accordingly, the average length of a transmission period of type 2 equals $\overline{Y}+\delta+\tau$, and substituting the value of $\overline{Y}$ in this expression we have

$$T_2 = \delta+\omega+2\tau-\frac{1-e^{-\lambda(\omega+\tau)}}{\lambda} \tag{17}$$

A transmission period of type 1 succeeds if no arrivals occur during the vulnerability period of the first transmission that starts the transmission period, which occurs with probability $e^{-\lambda(\tau+\omega)}$. If successful, the transmission period includes an ACK, and otherwise it consists of overlapping data packets as in a transmission period of type 2. Given that arrivals are assumed to be Poisson distributed, there can be no more than one arrival at any instant, which means that $Y = 0$ occurs when the transmission period succeeds, because the first and the last transmission in the period are the same. Therefore,

$$T_1 = T_2 + e^{-\lambda(\tau+\omega)}(\omega+\alpha+\tau) \tag{18}$$

Substituting Eq. (17) in Eq. (18) we obtain

$$T_1 = \delta+\omega+2\tau-\frac{1}{\lambda}+e^{-\lambda(\omega+\tau)}\left(\omega+\alpha+\tau+\frac{1}{\lambda}\right) \tag{19}$$

The value of $\overline{U}$ is simply the average time in a transmission period of type 1 dedicated to successful data. Given that a successful transmission period of type 1 occurs with probability $e^{-\lambda(\tau+\omega)}$, we have $\overline{U} = \delta e^{-\lambda(\tau+\omega)}$.

Substituting the values for T_1, T_2, and $\overline{U}$ in Eq. (14) we obtain Eq. (15). $\square$

It is important to note that making $\rho = 0$ in Eq. (15) results in the same throughput for non-persistent CSMA with priority ACKs stated in Eq. (2). This validates our Markov-chain formulation and should not be surprising. The state machine for time-based persistent CSMA is the same as for non-persistent CSMA when the persistence time $\rho = 0$, i.e., when there is no persistence.

C. Time-Persistent CSMA/CD

Figure 6 illustrates the transmission periods that may occur in time-persistent CSMA with priority ACKs. The figure illustrates a sequence of transmission periods and their lengths, which are indicated by the numbers 0, 1, and 2.

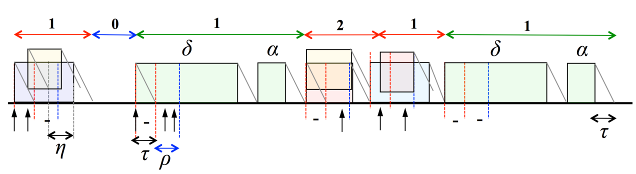

Fig. 6. Transmission periods in time-persistent CSMA/CD

Theorem 4: The throughput of time-persistent CSMA/CD is

$$S_{TCD} = \frac{\delta}{\frac{1}{1+\lambda\rho}e^{\lambda\tau}\left[\frac{1}{\lambda} + e^{\lambda\rho}(\eta+2\tau+\frac{1}{\lambda}(1-e^{-\lambda\tau}))\right] + D} \tag{20}$$

where $D = \delta+\alpha-\eta$.

Proof: As in non-persistent CSMA/CD, nodes do not incur receive-to-transmit turnaround latencies because they can listen and transmit to the channel concurrently. Given that the propagation delay in the network is τ seconds, the first packet in a transmission period of type 1 or 2 is vulnerable to interference for τ seconds, and hence the probability that a data packet is sent without interference is $e^{-\lambda\tau}$.

A transmission period of type 2 always consists of a collision interval that commences with two or more transmissions. Nodes detect the carrier from the transmissions that start the transmission period after τ seconds, and the nodes that start the transmission period with their own transmissions detect the carrier from the first interfering packet τ seconds after it starts. Accordingly, the length of a transmission period of type 2 is the same as the length of a collision interval obtained in Theorem 2, which is $Z+\eta+2\tau$, where Z is a random variable representing the time between the arrival of the data packet that starts the collision interval and the arrival of the first data packet that causes a collision. Using the value of $\overline{Z}$ in Eq.(6) we have

$$T_2 = \frac{1}{\lambda}\left(1 - e^{-\lambda\tau}\right) + \eta + 2\tau \tag{21}$$

A transmission period of type 1 consists of a collision interval if there is an arrival within τ seconds from the start of the transmission period, and consists of a data packet and an ACK if no arrivals occur within that time. Accordingly, the length of T_1 is the same as the average length of a busy period in non-persistent CSMA/CD, and from Eq. (7) we have

$$T_1 = e^{-\lambda\tau}\left(\delta + \alpha - \eta - \frac{1}{\lambda}\right) + \eta + 2\tau + \frac{1}{\lambda} \tag{22}$$

From Eq. (21), we can also express T_1 in terms of T_2 as follows: $T_1 = e^{-\lambda\tau}\left(\delta + \alpha - \eta\right) + T_2$.

The average time period used to transmit useful data is the same as in non-persistent CSMA/CD, $\overline{U} = \delta e^{-\lambda\tau}$.

Substituting the values for T_1, T_2, and $\overline{U}$ in Eq. (14) we obtain Eq. (20). □

We observe that, as it should be expected, making $\rho = 0$ in Eq. (20) results in the same throughput expression provided in Eq. (3) for non-persistent CSMA/CD.

VI. Performance Comparison

A. Modeling Assumptions

The throughput attained by a channel-access protocol is a function of the physical layer and medium-access control (MAC) layer. However, for the channel-access protocols we consider, the physical-layer overhead is roughly the same for all the MAC protocols. For simplicity, we do not consider the PHY-level overhead in our comparison. The results could then be interpreted by assuming that either the actual throughput attained by the protocols would be reduced by roughly the same amount, or the normalized length of data packets and ACKs takes into account the length of physical-layer headers.

We assume a channel data rate of 1 Mbps even though higher data rates are common today; this is done just for simplicity. We assume MAC-level lengths of signaling packets similar to those used in IEEE 802.11 DCF. For simplicity, however, we assume that an ACK is 40 bytes.

We assume that the time needed to detect collisions and send a jamming signal (η) in CSMA/CAD is roughly the duration of a jamming signal in CSMA/CD, or 48-bit time. We also assume that ω is 20μs, similar to the recommendations for IEEE 802.11 DCF.

We normalize the results to the length of a data packet by making $\delta/\delta = 1$, $G = \lambda \times \delta$, and $a = \tau/\delta$; and by using the normalized value of each other variable, which equals its ratio with δ (e.g., the normalized ACK length is α/δ).

B. Numerical Results

We compare the throughput of time-based persistent CSMA with priority ACKs and CSMA/CD with their non-persistent counterparts. As we have shown, the throughput of time-persistent CSMA equals the throughputs of non-persistent CSMA for $\rho = 0$, and the same applies to CSMA/CD. Accordingly, we simply use Eq. (15) for CSMA with priority ACKs and Eq. (20) CSMA/CD to present the throughput (S) versus the offered load (G) attained by CSMA and CSMA/CD for different values of ρ.

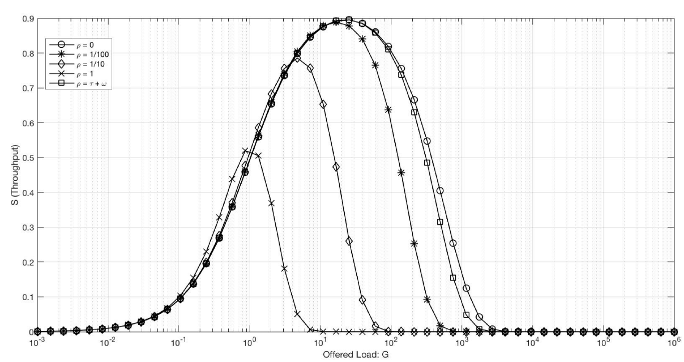

Fig. 7. S vs. G for time-based persistent CSMA with priority ACKs

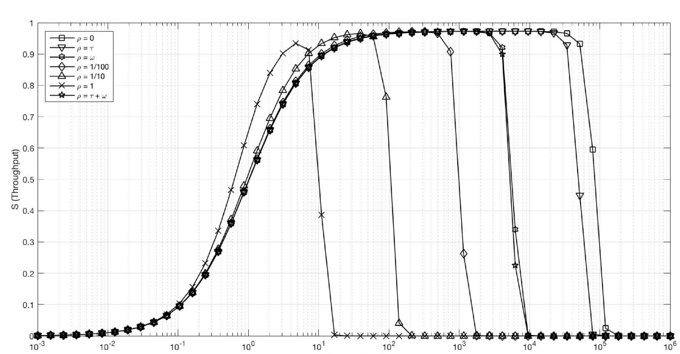

Fig. 8. S vs. G for time-based persistent CSMA/CD with priority ACKs

We present results for a local-area scenario that highlights the performance of the protocols when latencies are very short and signaling overhead is small relative to the time needed to transmit data packets. Physical distances are around 500 meters, and the duration of a data packet is 1500 bytes, which is an average-length IP packet and takes 0.012s to transmit at 1 Mbps. We use a normalized propagation delay of $a = 1 \times 10^{-4}$.

Figures 7 and 8 show the results for CSMA and CSMA/CD, respectively. With $\rho/\delta = 1$, time-based persistent CSMA and CSMA/CD correspond to their 1-persistent counterparts analyzed in the past by Kleinrock and Tobagi [9] and Sohraby et al. [11], with the only difference being the use of priority ACKs as part of the channel-access protocols. Similarly, $\rho/\delta = 0$ corresponds to the results for non-persistent CSMA with priority ACKs and CSMA/CD with ACKs.

It is clear from Figs. 7 and 8 that smaller values of ρ lead to higher throughput values. However, it can also be observed that relatively large values of ρ lead to higher throughput at light loads. This result points out the need to further improve on the basic time-based persistence transmission policy we have introduced. More specifically, a node should use large values of ρ at light loads and $\rho = 0$ once the channel is perceived as being congested. The resulting time and state-dependent protocols are the subject of future work, but is important to note that their design can be based on a state-aware extension to the proposed time persistence strategy, and their analysis for the case of a fully-connected network can be based on the same Markov-chain formulation we have presented.

VII. Conclusions and Future Work

We introduced time-based persistence in the context of channel-access protocols based on carrier sensing. With time-based persistence, a node with a packet to send that finds the channel busy determines whether the time of arrival of its local packet took place no later than ρ seconds (the persistence interval) from the time when the node started to detect carrier. If the time difference between the time for carrier detect and the time of the local arrival are smaller than ρ, then the node transmits its packet as soon as the channel becomes idle again.

We introduced a simple unifying analysis of the impact of time-based persistence in channel-access protocols that use carrier sensing and focused on CSMA with priority ACKs and CSMA with collision detection (CSMA/CD). Our model can be viewed as a generalization of the approach first described by Sohraby et al. [11] for the analysis of 1-persistent CSMA and CSMA/CD. Our analysis takes into account the effect that receive-to-transmit turnaround times have on performance, and the use of ACKs in the channel-access protocols.

The results of our analysis shows that time-based persistence can be as efficient as non-persistence, and our Markov-chain model was shown to provide the same results than the traditional model of non-persistent CSMA and CSMA/CD by making $\rho = 0$. Making $\rho/\delta = 1$ results in the traditional 1-persistent instantiations of CSMA and CSMA/CD.

As we have pointed out, our results enable the design and analysis of versions of CSMA and CSMA/CD in which a node uses different values of ρ (persistence interval) depending on the perceived state of the channel.

Our analysis of time-based transmission persistence strategies can be applied to CSMA/CA protocols, such as those examined in [5], as well as channel-access protocols that rely on carrier sensing and collision resolution (e.g., [7]). Furthermore, it is likely that the same Markov chain we have used can be adopted to analyze channel-access protocols that do not use carrier sensing and have persistence only after the correct reception of data packets [3].

Our future work in this area focuses on necessary changes to our analytical model in order to study time-based persistence in wireless network with hidden terminals [4], which calls for the adoption of more approximate models [1], [15].

Acknowledgments

This work was supported in part by the Jack Baskin Chair of Computer Engineering at UC Santa Cruz.

References

[1] M. Carvalho and J. J. Garcia-Luna-Aceves, “A scalable model for channel access protocols in multihop ad hoc networks,” *Proc. ACM MobiCom 2004*.

[2] K.-C. Chen, “Medium Access Control of Wireless LANs for Mobile Computing,” *IEEE Network*, vol. 8, no. 5, pp. 50–63, 1994.

[3] C.L. Fullmer and J.J. Garcia-Luna-Aceves, “Floor Acquisition Multiple Access (FAMA) for Packet-Radio Networks,” *Proc. ACM SIGCOMM ’95*, 1995.

[4] C. Fullmer and J. J. Garcia-Luna-Aceves, “Solutions to Hidden Terminal Problems in Wireless Networks,” *Proc. ACM SIGCOMM 1997*.

[5] J.J. Garcia-Luna-Aceves and A. Tzamaloukas, “The Effect of Exerting Adequate Persistence in Collision Avoidance Protocols,” *Proc. IEEE MoMuC ‘99*, Nov. 1999.

[6] J.J. Garcia-Luna-Aceves, “Carrier Resolution Multiple Access,” *Proc. ACM PE-WASUN 2017*, Miami, FL., Nov. 21–25, 2017.

[7] R. Garces and J. J. Garcia-Luna-Aceves, “Floor Acquisition Multiple Access with Collision Resolution,” *Proc. ACM MobiCom 1996*.

[8] M. Jainy et al., “Practical, Real-Time, Full Duplex Wireless,” *Proc. ACM MobiCom ‘11*, 2011.

[9] L. Kleinrock and F. A. Tobagi, “Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics,” *IEEE Trans. Commun.*, 1975.

[10] R. M. Metcalfe and D. R. Boggs, “ETHERNET: Distributed packet switching for local computer networks,” *CACM*, vol. 19, no. 7, pp. 395 – 403, 1976.

[11] K. Sohraby et al., “Comments on Throughput Analysis for Persistent CSMA systems, *IEEE Trans. Commun.* January 1987.

[12] H. Takagi and L. Kleinrock, “Throughput Analysis for Persistent CSMA Systems,” *IEEE Trans. Commun*, July 1985.

[13] F. A. Tobagi and V.B. Hunt, “Performance Analysis of Carrier Sense Multiple-Access with Collision Detection,” *Comput. Networks*, Oct./Nov.1980.

[14] F. Tobagi and L. Kleinrock, “The Effect of Acknowledgment Traffic on the Capacity of Packet-Switched Radio Channels,” *IEEE Trans. Commun.*, 1978.

[15] Y. Wang and J. J. Garcia-Luna-Aceves, “Collision Avoidance in Multi-hop Ad Hoc Networks,” *Proc. IEEE MASCOTS 2002*.