Computer Networks 127 (2017) 31–42

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

A CSMA-based MAC protocol for WLANs with automatic synchronization capability to provide hard quality of service guarantees

Yaw-Wen Kuo*, Jane-Hwa Huang

Department of Electrical Engineering, National Chi Nan University, Nantou, Taiwan, ROC

ARTICLE INFO

Article history:
Received 15 April 2016
Revised 12 July 2017
Accepted 17 July 2017
Available online 27 July 2017

Keywords:
Wireless LAN
Quality of service
Medium access control protocol

ABSTRACT

The carrier sensing multiple access with collision avoidance (CSMA/CA) protocol is a widely-adopted MAC protocol in the current wireless networks, but the quality of service (QoS) cannot be guaranteed due to random access. Investigations reveal that the collision avoidance mechanism which relies on the binary exponential backoff scheme is the root cause of QoS issue. Therefore, this paper first proposes a CSMA with automatic synchronization (CSMA/AS) MAC protocol to mitigate the collision problem caused by random access. By CSMA/AS, all the stations can be synchronized and then served in a round-robin fashion without contention collisions. Even if a new station joins, the wireless network can also quickly converge and go back to the synchronized state. The simulation results show that the proposed CSMA/AS protocol can fully mitigate the issues caused by random access, such as the severe contention collisions and large delay variation. In addition, this paper demonstrates how to provide hard QoS guarantees, such as fairness, rate guarantee, and delay guarantee, which cannot be achieved by the existing CSMA-based protocols. Because CSMA/CA does not rely on any additional control message, the implementation complexity of CSMA/AS is similar to that of legacy CSMA/CA protocols.

1. Introduction

The increasing prevalence of real-time applications has created a strong demand for quality of service (QoS) support in network infrastructures. A centralized contention-free medium access control (MAC) protocol [1–5] has frequently been considered the only method capable of serving real-time applications with hard QoS guarantees; however, the implementation of this protocol is highly complex because it requires the dynamic allocation of system resources in real time. In the IEEE 802.11 wireless local area network (WLAN) standard [1], the centralized contention-free MAC protocol was considered optional, and only a simple reference scheduler was presented. In practice, all commercial products are currently based on a contention MAC protocol. The primary MAC protocol of IEEE 802.11 [1] is called the distributed coordination function (DCF) based on the carrier sensing multiple access with collision avoidance (CSMA/CA) protocol with the slotted binary exponential backoff (BEB) scheme. The performance of the DCF protocol is extensively analyzed and studied in the literature. The works in [6–10] were based on the two-dimensional Markov chain model for computing performance metrics such as throughput, average access delay, or distribution of access delay. Several previous studies [11–19] have investigated the performance of nonsaturated networks under the assumption of homogeneous traffic. To enhance the performance of the DCF protocol, various algorithms [20–23] have been developed to adaptively adjust the initial backoff window size based on the estimated number of active stations or to control the amount of traffic entering the network by employing a shaper at each station. Collectively, these results indicate that QoS support is poor because random access generates large variations in packet delay. The enhanced distributed channel access (EDCA) protocol [1] is aimed at improving QoS guarantee in WLANs by assigning unique parameters to individual traffic classes. The EDCA protocol can provide only service differentiation without QoS guarantees, because it is also based on the CSMA/CA protocol for channel contention. Hard QoS guarantees cannot be provided to real-time applications with a strict delay bound by a random access MAC.

Packet delay plays a crucial role in the service quality of real-time applications, but the definition of delay depends on the studied scenarios. In previous studies on saturated WLANs [6–10], the delay of a successfully transmitted packet has been defined as the time interval from the moment the packet reaches the head of its queue to the moment it is transmitted to the destination (i.e., ac-

* Corresponding author.
E-mail address: ywkuo@ncnu.edu.tw (Y.-W. Kuo).

http://dx.doi.org/10.1016/j.comnet.2017.07.007
1389-1286/

cess delay). In studies investigating nonsaturated WLANs [11–19], wherein packets were generated according to a particular model, the delay has been defined as the interval from the moment when the packet is generated to the moment it is transmitted to the destination (i.e., end-to-end delay). Because of random access, a station can be infinitely interrupted whenever other stations are active. Consequently, only the statistics of delay are studied in the literature.

The problem of poor QoS capability arises from the BEB-based collision avoidance feature. The fundamental goal of BEB is to reduce collision probability under heavy loads by doubling the size of the contention window of collided stations. Under this design, various stations have equal share of the wireless channel in the long run, but packets from a particular station are not fairly treated. The doubled contention window forces collided packets to wait for long time, thus resulting in a lengthy delay. By contrast, non-colliding packets can be transmitted quickly. Consequently, the delay variation of a station is typically substantial, which is unfavorable for real-time applications. Our previous research [24,25] started with parameter optimization with delay constraint under the DCF protocol, and we proposed the fixed contention window backoff (FCWB) protocol to achieve a better performance. Because we believe that further improving the delay performance of the DCF protocol is difficult, we focused on modifying it. In [26], a novel protocol called the delay contention DCF (DC-DCF) protocol was proposed to assign higher priority to collided stations by adding several additional backoff slots in the first transmission attempt of a new head of line packet. The simulation results showed that the DC-DCF protocol possesses a traffic-shaping feature that was considered favorable because it provides traffic isolation between stations.

Basically, DC-DCF can be considered as the general form of the CSMA/CA protocol. The performance and behavior differ a lot with parameter configuration. In this paper, we first focused on optimizing the parameters of the DC-DCF protocol for a saturated WLAN based on the Markov chain model. The most critical result is that the optimal initial contention window size for the DC-DCF protocol should have a value of 1, which differs considerably from the values typically used in the DCF protocol. Subsequently, we conducted several simulations to validate the optimization, and the simulation results differed from what we anticipated. Through careful observation, a fundamental assumption of the Markov chain model was identified as invalid for the DC-DCF protocol with the above configuration. Nevertheless, we found that saturated WLANs become synchronized quickly when the value of the initial contention window is set to 1. Through careful adjustment of the backoff mechanism to operate under nonsaturated conditions, an appropriate version called CSMA/AS was proposed to form a synchronized WLAN without collisions. Consequently, a minimum service rate is guaranteed for each active station, and the bound of end-to-end delay can then be derived for a given type of traffic in order to provide hard QoS guarantees. The proposed CSMA/AS makes the traditional contention-based protocol able to provide a station level guarantee. We believe that it is cost-effective than Point coordination function (PCF) [1] which requires a customized scheduler.

This remainder of this paper is organized as follows. Section 2 provides a brief overview of the DCF, FCWB, and DC-DCF protocols. Section 3 describes our method for designing the proposed CSMA/AS protocol. The properties of CSMA/AS protocol are derived and explained in Section 4. Simulation results demonstrating the advantages of the proposed protocol are given in Section 5. Finally, Section 6 presents the conclusions of this study.

2. Preliminaries

To clarify the concepts presented in this paper, this section briefly introduces the fundamental concepts of WLANs in the MAC layer.

2.1. DCF protocol

When a station transmits a packet, it must first detect the wireless medium. If the medium is busy, it defers transmission until the medium is in idle. When the medium becomes idle, the source station can initiate a backoff operation only after an additional idle time interval, which is referred to as the DCF interframe space (DIFS). The backoff counter, which has a uniformly selected initial value, is decreased by one after an idle slot time, and is frozen when the source station detects that the medium is busy. When the backoff counter reaches zero, the source station starts transmitting the packet. When the backoff counters at multiple stations reach zero at the same time, a collision occurs when packets are transmitted simultaneously. When the destined station successfully receives the packet, it transmits a positive acknowledgment (ACK) to the source station after a time interval, which is known as the short interframe space (SIFS). After the source station receives the positive ACK, the transmission is successfully completed. If the source station does not receive the positive ACK, it schedules a retransmission, and the backoff operation restarts.

For each transmission attempt, the initial backoff count value is uniformly selected from $[0, W_i - 1]$, where W_i is the current contention window size, and i denotes the backoff stage (i.e., the number of failed transmissions for a given packet). Initially, $i = 0$ for each packet, and this value is increased by one when transmission failure is detected. The contention window size W_i at stage i is controlled according to the BEB scheme. At the first transmission attempt, W_0 is equal to the minimum contention window size, which is denoted as W. When a station detects a failed transmission, it doubles W_i until a maximum value $W_{\max} = 2^m W$ is reached, as shown in (1). Subsequently, W_i remains constant until the packet is successfully transmitted or dropped. The contention window size W_i is defined as

$$W_i = \begin{cases} 2^i W, 0 \le i \le m, \\ 2^m W, m < i \le R, \end{cases} \tag{1}$$

where R is the retry limit, and m is the maximum number of times that W_i can be doubled. When a station fails to transmit the packet at backoff stage R, it drops the packet and initiates a new transmission for the head-of-queue packet where $W_0 = W$.

To avoid the hidden node problem, a four-way handshake protocol, called the request to send (RTS)/clear to send (CTS) access mechanism, is typically used. When the backoff counter reaches zero, the source station transmits an RTS packet instead of the original data packet. After the destined station receives the RTS packet, it delays transmission of the CTS packet according to the SIFS. After successfully receiving a CTS packet, the source station delays transmission of the data packet according to the SIFS. If no CTS packet is received, then the source station must schedule a retransmission. RTS and CTS packets also contain a network allocation vector (NAV) field to notify other stations how long the source station requires to complete the transmission.

2.2. Transmission opportunity (TXOP)

By adopting the DCF protocol, a station can transmit a packet only after completing the backoff operation. Hence, when the stations perform backoff operations, there exist idle periods in the wireless medium, resulting in high protocol overhead. The IEEE