Achieving Near Maximum Throughput in IEEE 802.11 WLANs with Contention Tone

Juki Wirawan Tantra, *Student Member, IEEE,* and Chuan Heng Foh, *Member, IEEE*

***Abstract*— Future *wireless local area networks* (WLANs) promise bit rates higher than 100 Mbps. Previous research by Xiao *et al.* reported that the current IEEE 802.11 *medium access control* (MAC) protocol does not scale well to high bit rate channels. In this letter, we propose an enhancement that uses *contention-tone* transmitted on a separate narrow band signaling channel. The proposed contention tone mechanism avoids more than 96% of transmission collisions, hence achieving near to the theoretical maximum throughput of a WLAN MAC protocol.**

***Index Terms*— MAC protocol, wireless LAN, IEEE 802.11.**

I. Introduction

The IEEE 802.11 standards currently specify *wireless local area networks* (WLANs) with bit rates up to 54 Mbps. The ongoing development of the new IEEE 802.11n standard promises bit rates over 100 Mbps. With these bit rates, it is expected that an access point will serve a large number of stations for scenarios such as large conference halls or lecture theaters. However, the throughput of the IEEE 802.11 WLANs at these high bit rates is much lower than their specified bit rates because of the *medium access control* (MAC) and physical overheads [1]. Moreover, collision overheads in the IEEE 802.11 *distributed coordination function* (DCF) could cause performance degradation with a high number of stations [2].

Previous researches have highlighted the benefits of out-of-band signaling [3]–[5] to improve system performances. Recently, Yang and Vaidya propose an implicit pipelining scheme that mimics the benefits of out-of-band signaling [6]. In this letter, we propose a method that efficiently avoids collisions among stations and make the channel assignment work conserving even under heavy load. The protocol makes use of *contention-tone* (CT) on a separate narrow band signaling channel to resolve the station contention concurrently during an ongoing frame transmission.[1]

In this design, the busy tone (BT) [3] may be used as the CT. However, instead of using the tone to indicate busy in BT-based protocols, we use the tone for channel assignment. With our design, the proposed *CT protocol* (CTP) is capable of resolving more than 96% of the contentions within an ongoing frame transmission period; hence, the station that is assigned to access the medium in the next sequence can transmit its frame immediately after the ongoing transmission. This efficient channel assignment allows CTP to operate close to the theoretical maximum throughput of a WLAN MAC protocol.

Accepted
IEEE Communications Leters 2006

[1]Similar to [7], a special tone is used to resolve collisions during an ongoing data transmission. However, the proposed protocol presented here further utilizes the out-of-band signaling technique.

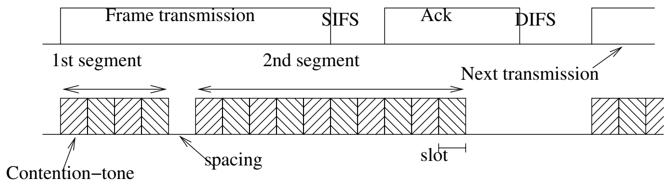

Fig. 1. Illustration of contention-tone protocol.

II. Contention-Tone Protocol

First, we define a tone as a slot length transmission burst identifiable by all stations. The proposed protocol uses a separate signaling channel with a separate circuitry for the transmission and detection of CTs; hence, a station can perform the CT contention resolution operation even when the station is currently transmitting data on the data channel.

CTP is implemented as follows. Stations access the medium using DCF basic access method [8]. When a station starts a data frame transmission on the data channel, it also starts a *CT contention procedure* on the signaling channel concurrently. The goal of the CT contention procedure is to distributively assign a *winner* among a group of contending stations. The duration of the CT contention procedure is always shorter than the DCF data transmission period, which is inclusive of all overheads such as the acknowledgment reply, so that a winner can be decided by the end of the data transmission period. An illustration of the CTP operation is shown in Fig. 1.

CTP divides the duration of the CT contention period into two consecutive time segments, which are called the *first segment* and the *second segment* for the earlier and the later time segments respectively. The first segment is accessible by stations which are currently transmitting data on the data channel. In most circumstances, there is only one station accessing this segment. However, if a transmission collision occurs on the data channel, the number of stations accessing the first segment is the number of collided stations on the data channel. The purpose of the first segment is to resolve such an unlikely but inevitable collision on the data channel.

The second segment is accessible for stations that do not participate in the data transmission but has a backlogged data frame in their local buffer. CT contention procedures for the two segments are similar. As shown in Fig. 1, each time segment is subdivided into several mini-slots named *contention slots*. The duration of a contention slot must be at least the sum of a CT detection time and the listen-to-transmit turnaround time. The number of contention slots in each time segment is a protocol parameter. Typically, the second segment consists of more contention slots due to the fact that the number of stations involved in the second segment is usually higher than that in the first segment, and with more contention slots, the probability of generating a single winner is higher. We shall describe in details the CT contention procedure as follows.

In CT contention procedure, the time segment starts with a CT that must be transmitted by all the involved stations. This CT signifies the existence of the segment. For each contention slot after the first CT, each involved station probabilistically transmits a CT. A station can either transmit or listen to the signaling channel, but not performing both during a contention slot. When listening to the channel, if a station detects a CT, it loses the contention and stops transmitting CTs for the rest of the segment period. As the procedure progresses, more stations will lose the contention leaving fewer stations transmitting CTs towards the end of the segment period. The winner of the contention shall be the last station that transmits a CT.

Obviously, it is possible that more than one winner is resulted by the end of a time segment. According to CTP, where winners transmit immediately after the ongoing transmission, a data transmission collision occurs if the number of winners is more than one. However, with a proper protocol design in practical situations, such a collision can be rare. In the next section, we show that it is possible to design a practical system such that the success rate to obtain a single winner is more than 96%.

With each segment produces a winner, the priority is given to the winner of the second segment if the second segment of the CTP exists. Hence, the winner of the second segment will start its transmission immediately after the current transmission is completed (successfully or not). If the second segment does not exist, the winner of the first segment will be the ultimate winner. It is possible that this winner is also the sender of the data transmission during the CT contention. If its data transmission is successful and it has no more data to transmit, it simply turns itself idle and let the medium be free for future access from any station generating new data frames. With this CT contention procedure, the data channel can be fully utilized as long as there is a station with a data frame to transmit.

III. Success Rate of CTP

The performance of CTP depends on the efficiency of the CT contention procedure. We first derive the probability that a single winner is resulted from the CT contention procedure. This probability determines the number of successful transmissions from the total number of possible transmissions, i.e. the success rate of the system. Let N denote the total number of stations in the network and K_j denote a random variable of the number of stations contending on the j-th contention slot. Here, we assume that the stations are in saturation condition [2], i.e. they always have frames to transmit. The number of stations contending on the j-th slot depends only on the number of stations contending on the previous slot, i.e. $(j-1)$-th slot. This relationship forms a recursive equation which defines the probability density of K_j conditioned on K_{j-1}:

$$Pr\{K_j = k_j | K_{j-1} = k_{j-1}\} = \tag{1}$$
$$\begin{cases} \binom{k_{j-1}}{k_j}\theta^{k_j}(1-\theta)^{k_{j-1}-k_j}, & \text{if } k_j = 1, 2, \cdots, k_{j-1}-1 \\ \theta^{k_{j-1}} + (1-\theta)^{k_{j-1}}, & \text{if } k_j = k_{j-1}, \end{cases}$$

where θ is the slot access probability, i.e. the probability to transmit a CT on a contention slot. Unconditioning (1), we get a recursive equation describing the probability density of K_j:

$$Pr\{K_j = k_j\} = \left[\theta^{k_j} + (1-\theta)^{k_j}\right] Pr\{K_{j-1} = k_j\} + \sum_{m=k_j+1}^{n} \binom{m}{k_j}\theta^{k_j}(1-\theta)^{m-k_j} Pr\{K_{j-1} = m\}. \tag{2}$$

The probability of successfully resolving contentions among N stations, $P_S(N)$, is the probability that the number of contending stations on the last slot is one, i.e. $Pr\{K_\omega = 1\}$:

$$P_S(N) = Pr\{K_{\omega-1} = 1\} + \sum_{m=2}^{N} m\theta(1-\theta)^{m-1} Pr\{K_{\omega-1} = m\}. \tag{3}$$

Given the number of stations, the number of contention slots, and the slot access probability, we can compute the probability of a successful transmission resulted from the CT contention using (3). We found that with nine contention slots and slot access probability between 0.3 and 0.4, we can achieve over 96% success rate for up to 100 stations. With three contention slots and slot access probability of 0.35, we can successfully resolve more than 70% of the contentions among up to 20 stations. Non-uniform access probability can yield better success rate. However, in favor of simplicity, we implement uniform access probability, since the uniform probability approach is good enough to drive the protocol to operate at close to the theoretical maximum throughput.

With CTP, we expect that most of the contentions occur in the second segment; the second segment CT contention usually involves high number of stations, hence, we allocate more contention slots to the second segment, specifically 9 contention slots which gives 96% success rate. Moreover, priority is given to the winner of the second segment which gives incentive to promote the second segment's success rate.

On the other hand, we expect that the number of participating stations in the first segment is small. These stations may be those accessing an idle medium at the same time or those claiming to be winners in the previous CT contention. Either scenario typically involves a small number of stations. Thus, we allocate only 3 contention slots to the first segment.

Adding the first compulsory CT, we design 4 contention slots for the first segment and 10 contention slots for the second segment. For protocol robustness, we enforce an empty mini-slot between the first and the second segments as a guard time. This results in 15 slots for the duration of the CT contention period. Considering the IEEE 802.11a standard [8] which specifies a duration of 9 μs for the slot time, the

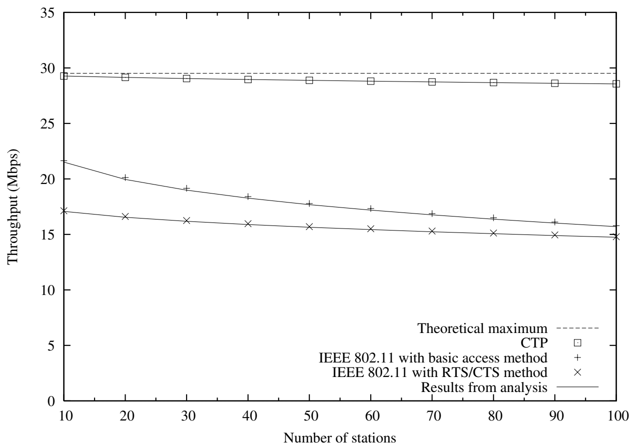

Fig. 2. Saturation throughput performance of the proposed CTP compared with the DCF standard. We use IEEE 802.11a parameters [8] with a bit rate of 54 Mbps and a frame size of 1000 bytes. We use a basic bit rate of 6 Mbps for the purpose of frame header transmissions.

duration of the CT contention period will be 135 μs. This duration is still shorter than a DCF data frame transmission period even for a zero sized data frame.[2]

IV. Throughput Performance of CTP

Wireless communications impose high synchronization overhead at the physical layer on every transmitted frame. Besides, the necessity of an explicit positive acknowledgment further reduces the throughput at the MAC layer. With these considerations, let P be the payload size, the maximum achievable throughput of a WLAN MAC protocol is

$$S_{max} = \frac{E[P]}{T_H + E[P] + T_{SIFS} + T_{ACK} + T_{DIFS}}, \quad (4)$$

where T_H and T_{ACK} are the time durations required to transmit the header and acknowledgment respectively; T_{SIFS} and T_{DIFS} denote the inter-spacing spaces between transmitted frames [8].

The throughput of CTP can be easily computed as follows. From (3), we determine the success rate for the channel assignment. Precisely, it is the probability that a data transmission is free from a collision. Assuming N saturated stations, for a successful data transmission, the success rate for the next channel assignment is $P_S(N-1)$ as there are $N-1$ participants in the second segment of the CT contention. If the CT contention results in c winners where $c > 1$, a collision occurs and the success rate of the following CT contention will be $P_S(N-c)$. Our numerical results suggest that $P_S(N-c)$ is not sensitive to c for our design.[3] Hence, we assume that regardless of the outcome of the data frame transmission, the success rate for the next channel assignment is always $P_S(N-1)$. With this approximation, the throughput of CTP can be determined by

$$\begin{aligned} S_{CTP} &= P_S(N-1) \cdot S_{max} \\ &= \frac{P_S(N-1)E[P]}{T_H + E[P] + T_{SIFS} + T_{ACK} + T_{DIFS}}. \end{aligned} \quad (5)$$

Fig. 2 shows the saturation throughput performance of CTP compared with the DCF protocols. Symbols represent simulations results, whereas solid lines represent analytical results (we use the equations in [9] for DCF). We also plot the theoretical maximum throughput of the MAC protocol in the figure. The figure shows that CTP can achieve throughput results that are very close to the theoretical maximum. The performance advantage is obvious when it is compared with the existing DCF standard protocols. For comparison, with 50 stations, CTP offers 61.7% additional goodput than that of the DCF basic access method.

V. Conclusion

In this letter, we proposed a contention-tone protocol, which uses a separate narrow band signaling channel to resolve contentions among stations. Since the contention resolution occurs concurrently with the data transmission period, the channel assignment exhibits work conserving characteristic. This allows the protocol to operate at near maximum throughput. Our analysis confirmed that CTP resolves over 96% of the contentions successfully. Further analysis and simulation provide evidence of the performance advantage of CTP over the existing DCF standard. Combining DCF basic access method as the method to access an idle medium with our proposed CT contention to resolve collisions, CTP offers not only low transmission delay in light load, but also high efficiency in heavy load.

[2] A DCF data frame transmission period includes the overheads of header and an acknowledgment sent with the basic bit rate. For future high bit rate WLAN implementation, we have the option of either enforcing frame padding for very short data frames or reducing the CT contention slots. The former option is favorable as it will have imperceptible impact on the system performance.

[3] Except when = which is a very rare event. This event occurs when all stations claim to be the winner of a CT contention. Even when this event occurs, it is expected that the recurrence of = is unlikely due to the CT contention procedure in the first segment, and hence normal operation resumes immediately.

References

[1] Y. Xiao and J. Rosdahl, "Throughput and delay limits of IEEE 802.11," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 355–357, Aug. 2002.

[2] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[3] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels–part II: the hidden terminal problem in carrier sense multiple-access and the busy-tone solution," *IEEE Trans. Commun.*, vol. 23, no. 12, pp. 1417–1433, Dec. 1975.

[4] J. Angelopoulos and T. Orfanoudakis, "An ATM-friendly MAC for traffic concentration in HFC systems," *Computer Commun.*, vol. 21, no. 6, pp. 516–529, May 1998.

[5] J. W. Tantra, C. H. Foh, and B. S. Lee, "An efficient scheduling scheme for high speed IEEE 802.11 WLANs," in *Proc. IEEE VTC 2003*, vol. 4, pp. 2589–2593.

[6] X. Yang and N. H. Vaidya, "A wireless MAC protocol using implicit pipelining," *IEEE Trans. Mobile Comput.*, vol. 5, no. 3, pp. 258–273, May-June 2006.

[7] C. H. Foh and M. Zukerman, "CSMA with reservations by interruptions (CSMA/RI): a novel approach to reduce collisions in CSMA/CD," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 9, pp. 1572–1580, Sept. 2000.

[8] IEEE Std. 802.11a, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: high-speed physical layer in the 5 GHz band," Sept. 1999.

[9] C. H. Foh and J. W. Tantra, "Comments on IEEE 802.11 saturation throughput analysis with freezing of backoff counters," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 130–132, Feb. 2005.