1-2009

Minimizing Spatial and Time Reservation With Collision-Aware DCF in Mobile Ad Hoc Networks

Lubo Song
Telephony Software Development Technology Division, lubo.song@palm.com

Chansu Yu
Cleveland State University, c.yu91@csuohio.edu

Follow this and additional works at: http://engagedscholarship.csuohio.edu/enece_facpub

Part of the Digital Communications and Networking Commons

Publisher's Statement

NOTICE: this is the author's version of a work that was accepted for publication in Ad Hoc Networks. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Ad Hoc Networks, 7, 1, (01-01-2009); 10.1016/j.adhoc.2008.02.002

Original Citation

Repository Citation
http://engagedscholarship.csuohio.edu/enece_facpub/59

This Article is brought to you for free and open access by the Electrical Engineering & Computer Science Department at EngagedScholarship@CSU. It has been accepted for inclusion in Electrical Engineering & Computer Science Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.
Minimizing spatial and time reservation with Collision-Aware DCF in mobile ad hoc networks

Lubo Song\textsuperscript{a}, Chansu Yu\textsuperscript{b, *}

\textsuperscript{a}Telephony Software Development Technology Division, Palm, Inc., 950 W. Maude Avenue, Sunnyvale, CA 94085, United States
\textsuperscript{b}Department of Electrical and Computer Engineering, Cleveland State University, 2121 Euclid Avenue, Stilwell Hall 437, Cleveland, OH 44115, United States

1. Introduction

Wireless LANs (WLANs) based on IEEE 802.11 standards [1] are in great popularity in public as well as in residential areas offering new level of information accessibility and convenience of life. This trend
enables the 802.11 network interface card available at an affordable price making it feasible to equip additional features such as transmit power and transmit rate control [2]. However, WLAN hardware as well as the underlying 802.11 standards has been originally developed for single-hop wireless communications between an access point (AP) and end clients, which may not be appropriate for multihop communications. As various forms of multihop networks such as wireless mesh networks [3,4] and wireless sensor networks [5,6] emerge, it is critically important to reconsider design choices made for WLANs in the context of multihop environment. For example, Ramanathan recently discussed three key areas of reconsideration for mobile ad hoc networks (MANETs) [7]: Relay communication primitive in addition to transmit and receive primitives, path-centric medium access rather than hop-centric, and cooperative transport of data frames.

In this context, this paper proposes an efficient medium access control (MAC) mechanism optimized for multihop environment. More specifically, this paper claims that carrier sense (CS)-based MAC algorithm such as distributed coordination function (DCF) of IEEE 802.11 standard [1] does not efficiently utilize the spatial spectral resource in MANETs. According to the DCF, a node needs to defer its transmission if it observes the carrier signal above the pre-specified CS threshold [8]. Since there exists at most one communication in WLANs, it is beneficial to set the CS threshold as low as possible in order to preclude as many other communication attempts as possible and thus to protect the ongoing communication. This must be contrasted to multihop environment, the essential idea of which is to support multiple concurrent communications for the efficient use of the limited spectral resource. A low CS threshold is not desirable because it disallows more communications than necessary and degrades the network performance. On the other hand, a high CS threshold may cause more collisions by encouraging more concurrent communications. Finding a tradeoff between the spatial reuse and collision has been an active research area in recent years [9–13].

Our goal in this paper is to develop a distributed MAC algorithm, called Collision-Aware DCF (CAD), where each node adjusts the CS threshold of its radio just enough to protect the ongoing communication. It essentially reserves the smallest spatial area and thus allows other nodes to initiate the transmission of their pending frames. In other words, CAD provides more transmission opportunities than DCF by eliminating unnecessary transmission deferrals and thus helps increase the spatial spectral utilization leading to a higher network throughput. For instance, when the sender and the receiver are sufficiently near with each other, their communication would be successful even in the presence of interference because signal-to-noise-interference ratio (SINR) at the receiver is high anyway [14]. Therefore, any additional communication in the neighborhood would not be a problem even though both transmitters are within each other’s carrier sense area. In CAD, each transmitter estimates the spatial reservation requirement for its data transfer and shares this information with its neighbors by piggybacking it in control and data frames. It makes it possible for the neighbors to speculate on whether their data transfer would interfere or be interfered by ongoing communication. Contributions of this paper are three-fold:

- First, this paper highlights the importance of efficient use of spatial resource in multihop environment. It proposes CAD where each communication expenses the smallest necessary spatial area for the least amount of time and thus the overall network throughput can be maximized.
- Second, CAD is a cooperative scheme in the sense that each node makes its deferment decision based not only on its own requirements but also on the collective information from nodes in the neighborhood. Considering the distributed nature of multihop environment, cooperation among the nodes is a necessity in optimizing the network performance. We hope this work facilitates future research on MANETs in this direction.
- Third, another unique feature of CAD is that the reservation requirements are embedded in the physical (PHY) header (called physical layer convergence protocol (PLCP) header in 802.11 terminologies) rather than in the MAC header as explained in Section 4. Since the PLCP header is usually transmitted at the basic, lowest rate, this guarantees that the information is propagated to a larger group of neighbors leading to a more efficient use of the spatial resource. To the best of the authors’ knowledge, there has been no such study in the literature.

A caveat is that the CAD protocol requires incompatible changes to the DCF because of the additional fields in the PLCP header. However, as discussed
earlier, we believe that multihop networking demands new designs in many respects and hope the proposed changes in this paper would be considered for adopting in future WLAN standards for MANETs. The rest of the paper is organized as follows. Section 2 summarizes techniques that improve the spatial reusability in multihop ad hoc networks. Section 3 presents system model including signal propagation model and IEEE 802.11 DCF. The proposed algorithm, CAD, is introduced in Section 4. Section 5 demonstrates the performance benefits of CAD via ns-2 simulation. Finally Section 6 draws conclusions and describes the future work of this study.

2. Related work

In MANETs, spatial area is one of the valuable resources. Improving the spatial reusability takes more importance as wireless bandwidth is critically limited, particularly in unlicensed ISM bands. This section summarizes recent work addressing this issue based on transmit power control (TPC), transmit rate control (TRC), directional antenna control (DAC) and carrier sense control (CSC) techniques.

2.1. A transmit power, transmit rate and directional antenna control

The idea of TPC is to apply the lowest necessary transmit power that can maintain the communication between the sender and the receiver while consuming the least energy [15–17]. For example, two control packets, request-to-send (RTS) and clear-to-send (CTS), can be exploited to detect the link quality and thus to determine the optimal power level for transmitting data packets. Data transmission uses the minimum necessary transmit power. Therefore, it saves power as well as reduces spatial footprint. However, RTS and CTS packets use the maximum power for ensuring the correct operation of collision avoidance mechanism and thus, they are unable to enhance the spatial utilization [17]. On the contrary, smallest common power (COMPOW) [18] and power-stepped protocol (PSP) [17] use the same radio power for both data and control packets but adaptively changes it depending on node and/or traffic density. A disadvantage of these TPC protocols is that they incur an additional overhead to compute the optimal transmit power level, which may not be appropriate in dynamic, mobile environment.

On the other hand, DAC is an explicit way of improving spatial reusability because signal propagation and the corresponding interference are limited to a certain direction with directional antennas [19–22]. Unfortunately, the well-know hidden terminal problem [23] becomes even more serious because more neighbors are ignorant or “deaf” to the ongoing communication. Alternatively, TRC exploits the physical layer multi-rate capability to make data transfers more robust to interference. For example, a receiver measures the channel quality based on the RTS it received and then informs to the sender an appropriate data rate so that the channel can always be utilized at the highest feasible data rate. According to Shannon, a certain data rate is always achievable even if the SINR is not high. More concurrent data transfers can be enabled leading to more spatial reuse. Shepard studied the theoretic bounds of the network throughput, assuming that the transmit rate is arbitrarily adjustable [24]. Prabhakar et al. proposed an energy-efficient communication schedule that takes the TRC capability into account [25]. And Sadeghi et al. proposed an opportunistic media access scheme that better exploits the channel via TRC and channel quality information [26].

2.2. Carrier sense control

CSC has been considered recently in the literature based on the assumption that the CS threshold is tunable within the detect sensitivity of the hardware [9]. A higher CS threshold encourages more concurrent transmissions but at the cost of more collisions. On the other hand, a lower CS threshold reduces the collision probability but it requires a larger spatial footprint and prevents simultaneous transmissions from occurring, potentially limiting the network throughput. Obviously there is a trade-off between high spatial reuse and increased chances of collisions [9].

Fuemmeler, et al. studied the collision prevention conditions in this context and concluded that the product of transmit power and CS threshold should be kept to a fixed constant [10]. Yang and Vaidya [11] and Zhai and Fang [12] considered the CSC adaptation while taking the influence of MAC overhead and transmit rate into consideration. These studies focused on analytic models for obtaining aggregate throughput. In comparison, our focus in this paper is to develop a CSC algorithm that realizes the optimal adaptation. Zhu, et al. used an analytical model to determine the optimal CS threshold and
proposed a distributed algorithm, called adaptive physical carrier sensing (APCS), that dynamically adjusts the CS threshold in 802.11 mesh networks [13]. In APCS, ideally all the nodes use the same optimal CS threshold. The main reason behind the network-wide fixed CS threshold is fairness. Using the same CS threshold can give the nodes an equal chance to transmit their packets. This scheme was improved recently by adding receive sensitivity adaptation [27]. The node’s radio would not be locked onto the first signal and thus becomes available to receive any late-arriving signals [27]. It is considered a receiver technique in the sense that a node is allowed not to receive a signal even though it is stronger than the CS threshold. In contrast, our approach in this paper is considered a transmitter technique because a node is allowed to transmit its pending frame even in the presence of carrier signal.

3. System model

CAD is an efficient MAC scheme that makes enough but no more than necessary spatial and time reservation. Before explaining the CAD protocol in Section 4, this section describes the system model assumed throughout this paper and time-spatial spectral resource reserved in DCF based on the model. Section 3.1 introduces the signal propagation and reception model, which is used to estimate those reservation requirements. Section 3.2 explains DCF and its time-spatial region reserved by the physical carrier sensing (PCS) and virtual carrier sensing (VCS) mechanisms. Section 3.3 discusses the optimal necessary reservation, which is then contrasted with that of DCF.

3.1. A propagation and reception model

Signal propagation in wireless channel is affected by path loss, shadowing and multiple-path fading. This paper assumes an open area environment where path loss due to communication distance is the most important. According to the corresponding two-ray ground reflection model [28], the receive power \( P_r(d) \) at distance \( d \) is given by

\[
P_r(d) = P_t G_t G_r \left( \frac{h_t h_r}{d^4 L} \right)^2
\]

where \( P_t \) is the transmit power, \( G_t \) and \( G_r \) are the antenna gains of the transmitter and the receiver, respectively, \( h_t \) and \( h_r \) are their antenna heights, and \( L \) is the system loss.

### Table 1

2.4 GHz Orinoco 11b Client PC Card Specification (Nominal output power of 15 dBm) [29].

<table>
<thead>
<tr>
<th>Transmit rate</th>
<th>1 Mbps</th>
<th>2 Mbps</th>
<th>5.5 Mbps</th>
<th>11 Mbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receive sensitivity</td>
<td>-94 dBm</td>
<td>-91 dBm</td>
<td>-87 dBm</td>
<td>-82 dBm</td>
</tr>
<tr>
<td>Range (open area)</td>
<td>550 m</td>
<td>400 m</td>
<td>270 m</td>
<td>160 m</td>
</tr>
</tbody>
</table>

To successfully receive a transmission the following two conditions have to be satisfied. First, the receiver must be within the transmission range of the sender. In other words, the receive power must be equal or larger than the receive sensitivity. Given a radio hardware and the transmit power level, the receive sensitivity is mostly affected by transmit rate. Table 1 shows their relationship of the 2.4 GHz Orinoco 11b Client PC Card [29]. Second, the receive power must be strong enough to overcome the interference of the noise and interference. This condition is described by the following SINR model.

\[
\text{SINR} = \frac{P_r}{N_0 + \Sigma I} \geq Z_0
\]

where \( N_0 \) is the background noise, \( \Sigma I \) is the interference from all other simultaneous transmissions, and \( Z_0 \) is the minimum required SINR ratio, commonly called capture ratio. The SINR model translates that even if more than one signal overlaps at the receiver, one of them could survive if it is much stronger than the sum of the others. This is called the capture effect [14] and it suggests that a careful coordination could make multiple communications be simultaneously successful even though they interfere with each other. The CAD protocol proposed in this paper is such a coordination algorithm where nodes do not have to reserve a large space dictated by the CS threshold.

3.2. Time and spatial reservation in DCF

Distributed coordination function (DCF) is a MAC layer protocol defined in IEEE 802.11 standard [1] and is used as a baseline MAC throughout this paper. DCF adopts both PCS and VCS mechanisms to effectively avoid collisions. The PCS is implemented based on the clear channel assessment (CCA) function, which is one of service primitives supported by IEEE 802.11 PHY as shown in Table 2. More specifically, PHY-CCA.indicate informs by the PHY to the MAC that the medium is sensed busy and the MAC believes so until it is informed...
otherwise. Even after PHY–CCA.indicate informs that the medium is free, a node defers a little longer (extended interframe space or EIFS period of time). This is to protect the reception of an ACK frame. As for spatial reservation in PCS, nodes in the proximity of the transmitter would sense the carrier signal and defer. The corresponding range is called carrier sense range (CR) in this paper. On the other hand, VCS is achieved by using RTS and CTS control frames. Upon overhearing an RTS or a CTS, a node considers the medium busy and holds its communication during the time period specified in the Duration/ID field of the control frames. The corresponding range from the transmitter is called transmission range (TR).

In summary, time reservation in DCF is based on the Duration/ID field in a MAC header and the EIFS mechanism. For the former, any neighbor who receives or overhears the MPDU would honor the reservation by withholding its data transmission for the specified amount of time. The latter is implemented by using the CCA function mentioned above. On the other hand, the spatial reservation is not explicitly specified in DCF and is simply mandated for those who receive the frame. For example, consider that a frame is transmitted. Node A receives the frame successfully and it would defer based on the Duration/ID field in the frame. Node B, however, senses the frame but does not successfully receive it (checksum error). Node C does not even sense the frame. We can say that the spatial area around node A and node B has been reserved based on the VCS and PCS, respectively. But, the spatial reservation does not cover the area around node C.

Fig. 1 shows the time-spatial reservation using PCS and VCS for the communication between a sender (s) and a receiver (r). Assuming that radio propagation is mostly affected by path loss, TR and CR translate to circular regions around the transmitter as shown in the figure. And they denote spatial reservations based on VCS and PCS, respectively. Dotted rectangles denote the time-spatial reservation due to PCS and solid rectangles denote that due to VCS. As shown in the figure, this reservation changes over time. Fig. 1a depicts the case of sender–receiver distance of 200 m and Fig. 1b depicts the case of 100 m. TR and CR are assumed to be 250 m and 550 m, respectively. From the figure, we made the following important observations.

![Fig. 1. Time-spatial reservation using PCS and VCS in DCF.](image-url)
Spatial reservation during the DATA transmission, which is usually the longest among the 4-way handshaking process, can be denoted as \((\text{CR}_s \cup \text{CR}_r)\).

- A shorter communication (Fig. 1b) reserves the similar time-spatial area as a longer communication (Fig. 1a) in DCF, which in fact is a source of inefficiency.

- While the PCS starts to be effective as soon as node \(s\) begins transmitting a frame, the VCS becomes effective since after the complete reception of the frame. This is because the Duration/ID field is embedded in the MAC header and can be used only after it is confirmed using the CRC checksum at the end of the MAC frame. This is another source of inefficiency in DCF as will be explained in more detail in the next section.

3.3. C interference range and optimal time and spatial reservation

We have seen the time-spatial reservation of the DCF in the previous subsection. It is a conservative reservation but is considered optimal in wireless LAN environment because no concurrent communication is supported. In multihop environment, this conservative approach diminishes the spatial reusability of the channel and it is required to reserve the smallest necessary area during the smallest necessary time period. This subsection discusses the optimal (minimum) reservation requirements based on the radio propagation model described in Section 3.1. For this purpose, we introduce interference range (IR) which is the minimum separating distance between the receiver and a potential interfering node such that the node does not cause collisions at the receiver. Unlike the fixed CR in DCF, IR varies depending on communication distance \((d)\) between the sender and the receiver, capture ratio \((Z_o)\), as well as the number of interferers \((k)\) and their locations. It is noted that the optimal reservation requirement is IR because nodes outside of IR does not cause collisions.

In the below, we first of all (i) derive the IR with different number of interferers \((1 \leq k \leq 6,\) Fig. 2 and Eq. (3)), (ii) show that IR does not vary much with \(k\) as long as each case represents the worst-case for the given \(k\) (Fig. 3), (iii) solve for the IR for the case of \(k = 1\) for simplicity, and (iv) show it on Fig. 1 to reveal the “over-reserved” time-spatial area in DCF (Fig. 4).

![Fig. 2. Worst-case interference scenarios with \(k\) (1-6) first-tier interferers \(D_1 = D - d, D_2 = D_1 + \sqrt{D^2 + d^2} - Dd, D_4 = D_3 = \sqrt{D^2 + d^2} + Dd, D_6 = D + d\).](image-url)
Let $D$ be the separating distance between the sender and an interferer. The optimal spatial reservation (IR) can be deduced using Eq. (2) introduced in Section 3.1, i.e.,

$$\text{SINR} = \frac{P_r(d)}{N_0 + \sum_{i=1}^{k} P_r(D_i)} \geq Z_0$$

(3)

where $D_i$ is the distance between the receiver and interferer $i$. Here, only the first-tier interferers ($1 \leq k \leq 6$) are considered [30] as in Fig. 2 because their influence is dominant compared to that of the second- and third-tier interferers. To obtain the conservative estimate of the optimal spatial reservation or $D_{\text{min}}$, each figure in Fig. 2 depicts the worst-case scenario for the given $k$, where the interferers are located to cause the maximum interference to the receiver.

- Now, given $d$, $k$, $Z_0$ and $N_0$, it is not difficult to find $D_{\text{min}}$ that satisfies the inequality (3). Fig. 3 plots $D_{\text{min}}$ versus the communication distance ($d$) with differing number of interferers ($k$) assuming that $N_0$ is ignorable. To our surprise, it is obvious from Fig. 3 that $d$ almost dominates the influence. For example, when $d$ is 150 m, the variation of $D_{\text{min}}$ with different $k$ is at most 7%. This is because the signal attenuates very quickly with distance and thus the topmost interferer in each of the six figures in Fig. 2 (with the shortest distance to the receiver) dominates the interference.

- Due to this fact, each node in CAD estimates the spatial reservation based on the assumption that $k = 1$ and the interferer is located in the direction of the receiver as in the first figure of Fig. 2. And, $\text{IR or } D_{\text{min}} = \sqrt[4]{Z_0} / d$ based on Eqs. (1) and (3). Considering the four-way handshaking process, the optimal spatial reservation to protect a communication is $(\text{IR}_r \cup \text{IR}_s)$. This must be contrasted to the spatial reservation in DCF, $(\text{CR}_r \cup \text{CR}_s)$, as explained in Section 3.2.

- The optimal time-spatial reservation for the communication between a sender ($s$) and a receiver ($r$) is shown in Fig. 4 assuming that $Z_0 = 10$ dB. In

![Fig. 3. $D_{\text{min}}$ versus $d$ with varying number of interferers ($k$).](image)

![Fig. 4. IR and over-reservation in DCF.](image)
the figure, IR_s and IR_r denote the optimal necessary spatial area when node s or r receives, respectively. In the bottom of the figure, solid rectangles show unnecessary time-spatial area that is reserved by the PCS and VCS in DCF. Over-reservation is more serious when the communication distance is short as evident in Fig. 4b.

The next section proposes CAD, where each node adjusts its CS threshold based on the reservation requirement which in turn based on distance information of the ongoing as well as its own communication to protect both.

4. Collision-Aware DCF

This section proposes a new MAC mechanism, called Collision-Aware DCF (CAD), which expenses the minimum time-spatial spectral resource that is required to successfully deliver frames. Section 4.1 explains how it is facilitated in CAD. Section 4.2 explains how it is implemented in the context of 802.11 standards. Section 4.3 introduces how to estimate the spatial and time reservations. Section 4.4 discusses how to make transmission decisions in the presence of ongoing communication and thus increase the communication concurrency.

4.1. Three design considerations in CAD

A basic idea of any carrier sensing MAC is to defer a node’s communication when the node senses an ongoing communication. It avoids collisions and thus improves the overall network performance. From a communicating pair’s point of view, they make a time and spatial reservation to protect their communication. A question is how to accurately estimate the reservation requirements and how to let their neighbors know the reservation. In DCF, time reservation is specified in the Duration/ID field in the MAC header of a MAC protocol data unit (MPDU) frame as described in Section 3.2. Spatial reservation in DCF is implicit but is far from optimal, particularly in multihop MANET environment.

The main goal of the proposed CAD protocol is to make an optimal time and spatial reservation to maximize the performance of multihop networks. Three key considerations in this regard are: (i) Explicit spatial reservation, (ii) reservation information in PHY rather than MAC frame (MPDU) header, and (iii) time-spatial reservation limited to the very next frame instead of the end of the frame exchange sequence (RTS, CTS, DATA and ACK). Note that these three considerations are specifically targeted at multihop networks. In wireless LANs, these issues are not critically important and thus largely ignored in the current design of 802.11 standards.

- **Explicit spatial reservation**: In MANETs, a prudent spatial reservation as well as time reservation should be considered because it maximizes the channel utilization by allowing more concurrent communications [31].

- **Reservation information in PHY rather than MAC frame header**: In DCF, the reservation information is embedded in the MAC frame header and is propagated to the transmitter’s neighborhood. Nodes outside of the transmitter’s transmission range would be ignorant of the reservation. Without knowing the reservation requirement, the easiest way to avoid collisions is to defer one’s transmission whenever doubtful. The EIFS mechanism, explained in Section 3.2, implements this in DCF. CAD eliminates the unnecessary deferment by propagating the reservation information to a wider group of neighbors by embedding the reservation information in the PHY header. Note that while the MAC header can be transmitted at a higher rate, the PHY header (or called PLCP header) is always transmitted at the lowest data rate or 1 Mbps in 802.11 or 802.11b standards and thus travels a farther range.

- **Time-spatial reservation limited to the very next frame instead of the end of the frame exchange sequence**: In DCF, a transmitter estimates the time for the completion of the frame exchange sequence (i.e., RTS, CTS, DATA and ACK), which constitutes the Duration/ID field of a MAC header. It is used to set network allocation vector (NAV) of the nodes in the transmitter’s neighborhood to make time reservation. In CAD, the spectral resource is reserved only for the current and the following frame. It is beneficial because of the followings. First, consider a communication that is proceeded by RTS and CTS control frames. When RTS–CTS exchange fails, neighbors would still defer during the estimated completion time of the communication assuming that the four-way handshaking is successful. In CAD, neighbors would defer until the estimated reception time of the next frame (CTS) eliminating the unnecessary deferment or spectral resource wastage. Second, in DCF, there
exists a potential fairness problem. Consider the same example as above where the RTS–CTS exchange fails. The failed transmitter will obtain another chance to retransmit an RTS immediately (after waiting a DIFS) because other nodes would defer their transmissions for a longer time. In CAD, neighbors do not defer for the period of an entire frame exchange sequence and therefore, the fairness problem is avoided. Third, as the four-way handshaking progresses, the transmitter and the receiver will have a more accurate estimate on the required spatial resource based on the feedback from the counterpart of the communication. In CAD, spatial reservation in RTS frame is made in a conservative way as in DCF. However, spatial reservation in CTS and DATA frame will be estimated smaller than in DCF leading to a higher utilization of the spectral resource.

4.2. PLCP frame format and handling reservation requirements

Fig. 5 shows the PLCP frame format in CAD, where spatial and time reservation requirements are embedded. They are referred to as REQ_SR and REQ_TR, respectively. As discussed in the previous subsection, the benefit of embedding the reservation information in the PLCP header is that it is propagated to a wider group of neighbors because the PLCP header is transmitted at the lowest data rate (1 Mbps in 802.11 and 802.11b). Data rate used for the PHY payload (MPDU) transmission is indicated by the 8-bit SIGNAL field of the PLCP header in Fig. 5.

Another benefit is based on the fact that the reservation information in the MAC header cannot be utilized while the MPDU is being received. This is because an MPDU frame is not considered legitimate until the whole packet is received and its CRC checksum is confirmed. When a node overhears a long DATA frame, it will be beneficial to know whether it can safely transmit its own frame even before the complete reception of the current frame. This is in fact accomplished in CAD by including the time and spatial reservation information in the PLCP header. The PLCP header is received and confirmed for its integrity even before a node starts receiving the PHY payload (MPDU frame). This small change makes the CAD algorithm much simpler because it is possible in CAD for each node to be responsible for the current frame as well as the subsequent frames. It is more effective when RTS/CTS control frames are not used. On the other hand, CAD is disadvantageous as it increases the frame size. When the communication channel is bad and is characterized by a high bit error rate (BER), each transmission in CAD could be more subjective to corruption than in DCF. However, the frame size increase is not large enough to make a significant impact on network performance. For example, for a 512-byte PHY payload (MPDU), the PHY frame size increases from 4288 to 4320 bits, which is just a 0.7% increase.

Frame transmit and receive procedures in CAD are very similar to those in 802.11 standards with little changes. When a transmitter’s MAC requests a frame transmission to its PHY, a set of parameters are summarized in TXVECTOR and passed to its PHY layer along with PHY_TXSTART.request [1].
PHY_TXSTART.request (TXVECTOR) PHY_DATA.request (DATA)

MAC ......

PHY

SYNC, SFD SIGNAL, SERVICE, LENGTH REQ_SR, REQ_TR CRC MPDU

CRC start CRC end Rate change start

(a) Transmit procedure (TXVECTOR includes REQ_SR and REQ_TR)

If (RSSI < REQ_SR and RSSI < REQ_SR0)
medium is considered idle (see Section IV.D)

PHY_CCA.indicate (BUSY) PHY_RXSTART.indicate (RXVECTOR)

MAC ......

PHY

SYNC, SFD SIGNAL, SERVICE, LENGTH REQ_SR, REQ_TR CRC MPDU

CRC start CRC end Rate change start

(b) Receive procedure (RXVECTOR includes REQ_SR and REQ_TR)

Fig. 6. Transmit and receive procedures in CAD.

4.3. Estimating spatial and time reservation requirements

A main question in CAD is to accurately estimate the reservation requirements at a transmitter. Estimation of REQ_TR is based on frame length, frame type and data rate. Unlike 802.11, it is until the completion of the current and the next frame in the four-way handshaking process. For example, time reservation for a CTS frame is estimated as (SIFS + DATA + SIFS + ACK) and (CTS + SIFS + DATA) in DCF and CAD, respectively. Here, CTS, DATA and ACK denote the time transmit the corresponding frames. SIFS stands for short interframe space. Note that CTS is not included in the time reservation of DCF, which was explained in Section 4.2.

Regarding the spatial reservation requirements, we first of all compare that of DCF and CAD for two node pairs with communication distance of 100 m and 200 m as in Fig. 7. For brevity, the figure shows the spatial reservation during the DATA transmission only. It is clear from the figure that DCF reserves a much larger spatial area than CAD, which is more prominent when the communication distance is shorter as in Fig. 7b and d. For example, in Fig. 7d, when $d = 100$ m, the transmitter’s (node $s$) signal arriving at the receiver (node $r$) as shown in Fig. 6a. Note that the data rate used for MPDU transmission is indicated in the 8-bit SIGNAL field as discussed earlier. Fig. 6b shows the PLCP receive procedure. Upon detecting a coming signal, PHY_CCA.indicate (BUSY) will be issued to the MAC layer if the signal strength is higher than CS threshold. Then the PHY layer will begin searching for start frame delimiter (SFD) and start to receive a PLCP header. If a PLCP header is successfully received (CRC check passes), the PHY_RXSTART (RXVECTOR) will be issued to the MAC layer according to the 802.11. The RXVECTOR contains the information of SIGNAL field, SERVICE field, LENGTH field, RSSI, signal quality, and antenna used for receive.

In CAD, a transmitter estimates the reservation requirements, REQ_SR and REQ_TR, for protecting its communication and adds this information in the TXVECTOR as shown in Fig. 6a. The transmitter’s neighbors as well as the designated receiver obtain the reservation requirements and decide whether or not to defer. Similarly, in Fig. 6b, when a node receives a frame, it obtains information included in the PLCP header including the reservation requirements. They will be passed to the MAC in the form of RXVECTOR, which includes REQ_SR and REQ_TR.

Note that CTS is not included in the time reservation of DCF, which was explained in Section 4.2.
Fig. 7. Spatial reservation (shaded area) during the DATA transmission. (CR, CR, in DCF and IR, IR, in CAD. DCF reserves a much larger area than CAD leading to spectral resource wastage.)

r) is much stronger than necessary and thus is resilient to interference. This means that its neighbors could transmit their own frames without disrupting the s-r communication. In other words, nodes outside of the shaded area (reserved area) in Fig. 7d can concurrently transmit in CAD allowing a higher channel utilization than in DCF shown in Fig. 7b.

As discussed in Section 3.3, REQ_SR or the optimal spatial reservation is \( D_{\text{min}} = \sqrt{Z_0 / d} \). This is the maximum interference level that the receiver can tolerate. A key issue is how to translate any measurable quantity such as signal strength information (e.g., received signal strength indicator or RSSI) to communication distance assuming \( Z_0 \) is known. We assume in this paper that RSSI can be translated to distance information. Even though it does not hold, the basic idea of CAD can still be applicable because RSSI itself can be used as the spatial reservation requirement. For example, when a transmitter receives a CTS, it can measure the RSSI while receiving the frame. The higher the RSSI, they can tolerate a higher interference and thus their spatial reservation requirement is smaller. More specifically, when a node (say, node r) transmits a frame, it includes the spatial reservation requirement (REQ_SR) in its PHY frame header. It essentially says “any node within REQ_SR should defer”. In other words, any node that is separated by d from node r and \( r \leq d \) should defer. On the other hand, let us assume that the distance information is not available but RSSI is available for use in place of REQ_SR. The spatial reservation requirement in the PHY header says “any node that receives the signal with the signal strength of RSSI and \( 1 / \text{RSSI} \leq 1/\text{REQ_SR} \) should defer”.

4.4. Setting NAV and making transmission decisions

In DCF, a node decides to transmit its frame only when no carrier is sensed (PCS) and NAV is zero (VCS). For the latter, NAV is set to the Duration/ID field in the MPDU upon receiving a legitimate frame. For an erred frame, NAV is set to EIFS upon the end of the frame as explained in Section 3.2. Both mechanisms are to protect the ongoing communication.

Since CAD attempts to reserve the minimum necessary time and spatial resource, a care must be taken
not to interfere the ongoing communication as well as not to be interfered by the ongoing communication. Consider the case in Fig. 7d. A potential transmitter (say, node $A$) outside of the shaded area is free to start its communication, for example, with node $B$. Although node $A$ is sure that it does not cause any trouble to the $s$–$r$ communication, but it is not clear whether $A$’s communication is interfered by the $s$–$r$ communication or not.

This is clearer in Fig. 8. Node $A$’s transmission would not interfere node $r$’s communication in Fig. 8a and b because node $A$ is outside of node $r$’s spatial reservation requirement (REQ_SR). However, node $A$ does not always decides to transmit its own frame because it may be interfered by node $r$’s communication if it starts. It depends on node $A$’s spatial reservation requirement (REQ_SR0). In Fig. 7a, node $A$ needs a small REQ_SR0 and node $r$ is outside of the REQ_SR. Therefore, node $A$ decides to transmit. On the other hand, in Fig. 8b, node $A$ demands a larger REQ_SR0 and node $r$ is inside the shaded area. Node $A$, in this case, decides not to transmit because its communication could be interfered by node $r$’s communication.

In summary, the transmission decision is based on the answers for the following two questions: (i) Whether its communication is successful if it transmits concurrently with the ongoing data transfer. (ii) And, whether the ongoing communication is successful if it transmits. For the former question, the node compares the RSSI of the incoming signal with its own spatial requirement (REQ_SR0). (Here, we assume that RSSI and REQ_SR are converted to an equivalent measure, such as distance or signal strength.) In other words, the node defers if RSSI $\geq$ REQ_SR0 because the strength of the incoming signal exceeds the maximum interference level that its outgoing transmission can tolerate. For the latter, the node compares the RSSI of the incoming signal with REQ_SR of the ongoing transmission i.e., the node defers if RSSI $\geq$ REQ_SR because the ongoing communication would fail if the node transmits. This is based on the assumption that the link is symmetric; The RSSI of the incoming signal is equal to the RSSI (interference) that the node would cause to the ongoing transmission.

As a result, if RSSI $\geq$ REQ_SR0 or RSSI $\geq$ REQ_SR, the medium is considered busy and the node defers its transmission. In this case, the PHY will continue to receive MPDU but NAV is set to a new REQ_TR obtained from the incoming PLCP header. On the other hand, if RSSI $<$ REQ_SR0 and RSSI $<$ REQ_SR, the medium is considered idle. In this case, the PHY will issue PHY_CCA.indicate (IDLE) to the MAC layer so that the node can transmit its frame even though there is an ongoing communication as shown in Fig. 6b in Section 3.2. Therefore, CAD encourages more concurrent communications as long as they do not interfere with each other and thus increases the network throughput.

5. Performance evaluation

In order to evaluate the performance improvement of CAD, this section compares CAD with DCF based on ns-2 [32]. Section 5.1 explains the simulation environment and Section 5.2 presents simulation results with discussions.

5.1. Simulation environment

Our performance study is based on the simulation of 50 mobile nodes distributed on a $300 \times 1500$ m$^2$ area. Radio propagation model used is two-ray

![Fig. 8. Making a transmission decision. (Node r is currently communicating and the corresponding spatial reservation requirement is REQ_SR. Node A is a potential transmitter and it demands REQ_SR0.)](image-url)
ground propagation model as discussed in Section 3.1 and Eq. (1). Transmit range (TR) of 250 m and carrier sense range (CR) of 550 m is assumed. Capture ratio ($Z_0$) of 10 dB is used in our performance study. Regarding signal transmission and capture, we have extended ns-2 as follows.

- In the original implementation of the ns-2, the signal strength comparison is per-packet based. That is, the receiver compares the most-recently-arriving packet only with the one it is receiving. We modified then-2 to allow each node to track every incoming signal so that the effect of the additive interference can be simulated.
- Since this paper assumes that radio hardware uses different data rates for PLCP header and MPDU, we modified ns-2 to support multi-rate communication. The PLCP header is transmitted at 1 Mbps while the MPDU is transmitted at 2 Mbps.

The movement of the nodes is described by the random waypoint mobility model with the maximum speed of 5 m/s and with the pause time of 0–900 s. Ten to fifty CBR (constant bit rate) traffic is used to simulate the network traffic. Ad hoc on-demand distance vector (AODV) routing protocol [33] is used in our study to determine the routing path between the source and the destination. The simulation time is 900 s and each simulation scenario is tested with five runs to obtain the average performance measures.

5.2. Simulation results and discussion

This subsection presents simulation results comparing the performance of the proposed CAD with DCF and APCS [13]. Packet delivery ratio (PDR) and packet delay are used as primary performance metrics as shown in Fig. 9. The pause time varies between 0 and 900 s. Note that pause time of 900 s translates to a static network where nodes do not move because the simulation time is 900 s. On the other hand, pause time of 0 s corresponds to a constant moving scenario. Thirty CBR connections are simulated where source and destination nodes are chosen randomly among the 50 mobile nodes. Each traffic source generates three 1024-byte packets every second. As clearly seen in Fig. 9a and b, CAD outperforms DCF and APCS with a large margin. It achieves 16–19% higher PDR and 59–76% lower packet delay than DCF. Compared to APCS, it achieves 11–13% higher PDR and 32–67% lower delay.

The dramatic performance improvement of CAD over DCF is attributed to higher concurrency and yet its excellent capability in avoiding collisions. While nodes make transmission decisions depending on the carrier signal and the pre-determined CS threshold in DCF, the CAD protocol allows the nodes to make more intelligent decisions based on information from their neighbors. The reason why APCS does not perform well compared to DCF is that it basically is a conservative scheme as well. In APCS, a node determines its CS threshold by taking the minimum of its neighbors’ CS thresholds. When a node steps down its CS threshold, it is possible that all its neighbors step down leading to a suboptimal spatial reservation. CAD produces more communication opportunities but reduces collisions as evident in Fig. 9c and d, respectively. Note that we investigated the number of RTS transmissions and their collisions for this purpose because every routing or data packet is preceded by a RTS packet. Fig. 9c shows the number of RTS packets transmitted during the simulation. With CAD, nodes send 6–20% more RTS packets (communication opportunities) than with DCF, which is 0–12% than APCS. However, CAD results in 7–11% less collisions on RTS packets as shown in Fig. 9d. Compared to APCS, it is 9–14% less collisions. Note that Fig. 9c does not count the retransmitted RTS packets because we wanted to know the overall collision probability.

In order to understand how CAD improves the performance, we measured MAC layer control overhead as well as the packet queue size as shown in Table 3 and Fig. 10. One important advantage of CAD is short packet delay as in Fig. 9b. Our investigation shows that packet queuing delay is an important ingredient for this. Having more communication opportunities in CAD facilitates a mobile node to quickly offload pending packets and therefore, it helps keep its packet queue at each node as short as possible. In each of 900 s of simulation runs, we collected packet queue size every 10 s at each node and calculated the average statistics across all mobile nodes in the network. As shown in Table 3, each node has about 1.49–2.07 and 1.14–1.74 packets in its queue on the average with DCF and APCS, respectively, while it is 0.17–0.87
with CAD. This amounts as much as 90% and 80% reduction in packet queue size in comparison to DCF and APCS, respectively.

To observe the MAC control overhead, we measure the MAC layer control traffic (RTS, CTS and ACK) in DCF, APCS and CAD as in Fig. 10a. Although they generate about the same amount of MAC layer overhead, its detailed figures are quite different. For example, the number of RTS packets is almost the same in DCF, APCS and CAD but this is only true for the combined initial and retransmitted RTS packets. While CAD allows more number of initial RTS packets as already seen in Fig. 9c, it causes less RTS retransmissions than DCF and APCS. Similarly, the total number of ACK packets is similar among the three protocols. However, CAD results in more ACKs than DCF and APCS in response to DATA packets while it is exactly the opposite for ACKs in response to RREP packets as shown in Fig. 10b indicating that CAD uses more bandwidth for useful data transmission than DCF and APCS.

One important advantage of CAD is short packet delay as shown in Fig. 9b. Our investigation shows

Table 3
Average packet queue size

<table>
<thead>
<tr>
<th>Pause time (seconds)</th>
<th>DCF</th>
<th>APCS</th>
<th>CAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.07</td>
<td>1.74</td>
<td>0.87</td>
</tr>
<tr>
<td>100</td>
<td>1.71</td>
<td>1.14</td>
<td>0.17</td>
</tr>
<tr>
<td>300</td>
<td>1.88</td>
<td>1.27</td>
<td>0.40</td>
</tr>
<tr>
<td>600</td>
<td>2.06</td>
<td>1.38</td>
<td>0.30</td>
</tr>
<tr>
<td>900</td>
<td>1.49</td>
<td>1.57</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Fig. 9. Performance comparison with mobility.
that packet queuing delay is an important ingredient for this. Having more communication opportunities in CAD facilitates a mobile node to quickly offload pending packets and therefore, it helps keep its packet queue at each node as short as possible. In each of 900 s of simulation runs, we collected the packet queue size every 10 s at each node and calculated the average statistics across all mobile nodes in the network. As shown in Fig. 10c, each node has about 1.5–2.1 and 0.9–1.0 packets in its queue on the average with DCF and APCS, respectively, while it is 0.2–0.9 with CAD.

Another important measure in evaluating the CAD protocol is fairness. Since nodes may not defer their transmissions even though they sense the medium busy, it might cause unfairness among communication links. In order to compare the fairness of the MAC algorithms, we measured fairness index, defined as follows [34,35]:

$$F = \frac{\left(\sum_{i=1}^{N} y_i^2\right)^2}{N\sum_{i=1}^{N} y_i^2}$$

where $N$ is the number of connections and $y_i$ is the number of received packets for connection $i$. The value of this index ranges from 0 (completely unfair) to 1 (perfectly fair). According to our simulation result shown in Fig. 10d, CAD does not hurt the fairness of communications. On the contrary, it improves it in comparison to DCF and APCS. We think that off-loading of packets sooner than other protocols helps improve the fairness in CAD.

Fig. 11a and b depict the effect of traffic intensity in terms of the number of connections of CBR
traffic and node density. The number of connections varies from 10 to 50 in Fig. 11a, and the number of nodes changes from 75 to 175 in Fig. 11b. In both scenarios, the same data rate and packet size (three 1024-byte packets per second) are used. As shown in the figure, CAD consistently outperforms DCF and APCS although both of them suffer when the number of connections increases in Fig. 11a.

Recent experimental studies show that shortest (hop count) path does not always provide the best performance because it usually consists of longer hops, each of which is easily subjective to interference with a small SINR [4,36]. In order to see how CAD performs in a more realistic environment, a set of experiments has been conducted with the shadowing propagation model instead of the conventional two-ray ground propagation model introduced in Section 3.1. Shadowing is caused by the lack of visibility between two communicating nodes and it causes slow variations over the mean received power. The mean received power is calculated deterministically based on the communication distance. The randomness of channel is described by a log-normal random variable, the distribution function of which is Gaussian with zero mean and a specified standard deviation (SD).

Before presenting the simulation results, Fig. 12a shows how the radio channel behaves with the
shadowing model presenting the success ratio versus communication distance using ns-2. In case of SD of 0.0 dB, the shadowing model is equivalent to the deterministic two-ray ground model and thus the success ratio is 100% if the distance is less than 250 m, which is the transmission range. Otherwise, it is 0%. As SD increases, more communications fail even if the distance is less than 250 m, and more communications succeed even if the distance is longer than 250 m. When the communication distance is 200 m, the success ratio is 42% with SD of 10 dB. Less than a half of the transmission attempts can be successful even if the communication distance is shorter than the transmission range. Fig. 12b shows the effect of channel randomness in term of SD on the network performance. CAD consistently performs better than DCF and APCS in terms of PDR. However, the PDR margin between CAD and the other two decrease when SD is getting larger. This is because in CAD the REQ_SR calculation is deterministic and based on the communication distance only. It means that the REQ_SR estimation is not accurate in random channel. The more random the channel is, the more error the estimation contains. Correspondingly, the performance margin reduces. Efficient operation in the presence of randomness of the communication channel comprises one of our future works.

6. Conclusions and future work

This paper proposes Collision-Aware DCF (CAD) mechanism that encourages more concurrent transmissions but at the same time avoids collisions more efficiently. While the DCF avoids collisions based on a pre-determined carrier sense threshold (physical carrier sense or PCS) and advertise of the communication duration embedded in the MAC header (virtual carrier sense or VCS), both methods often fail to achieve the maximum achievable performance, particularly in multihop network environment. In CAD, each node estimates the range that it wishes to reserve for its data transfer (spatial reservation requirement) and the time duration (time reservation requirement) based on the communication distance or received signal strength. And, they are embedded in the PHY header of the transmitted packet so that a larger group of potential interferers become aware of it and ahead of time. Our simulation study based on ns-2 shows that CAD significantly improves the network performance in terms of packet delivery ratio and packet delay. It is observed that the benefit of CAD comes from more number of concurrent transmissions and smaller collision ratio, which in fact was the original goal of the CAD mechanism. CAD is designed to be compatible with TPC and TRC capability in the sense that estimation of the spatial reservation requirement can easily accommodate the transmit power and transmit rate information. This issue needs further study and is remained as a future work.

References

[16] J. Monks, J.-P. Ebert, W.-M.W. Hwu, A. Wolisz, Energy saving and capacity improvement potential of power control...


