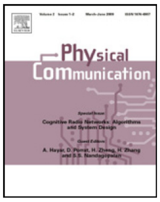
Physical Communication 28 (2018) 69–77

Contents lists available at ScienceDirect

Physical Communication

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



Full length article

# Cross-layer design benchmark for throughput maximization with fairness and delay constraints in DCF systems

Zihao You *, I-Tai Lu

*NYU Wireless, Department of ECE, NYU Tandon School of Engineering, 2 Metrotech Center, NY, 11201, USA*

ARTICLE INFO

*Article history:*
Received 1 May 2017
Received in revised form 12 March 2018
Accepted 13 March 2018
Available online 26 March 2018

*Keywords:*
Cross-layer
DCF
IEEE 802.11
Throughput
Fairness
Delay

ABSTRACT

This paper proposes a cross-layer design for the stations (STAs) in a distributed coordination function (DCF) network. By selecting the modulation scheme, coding scheme, and packet length of each STA in the network, the design aims to maximize the total throughput of all the STAs and satisfy the minimum throughput requirement or delay requirement of each STA, thus addressing the fairness and delay issues. The proposed scheme applies to the system where each STA employs a contention based channel access mechanism, Furthermore, unlike the existing optimization schemes, it takes two important factors, changeable data rate and changeable packet error rate (PER), into consideration. Using an existing Markov chain model to predict the performance of the STAs, we propose an approach that updates the selection of each STA sequentially, thus avoiding the large complexity from the exhaustive search. Many issues are discussed based on the numerical results, including how the approximations in our design affect the processing time and result of the design, how the change of one STA affects the performance of the other STAs, how the minimum throughput constraints affect the fairness and total throughput, how to select these constraints to satisfy the delay requirements, etc.

## 1. Introduction

With the explosive increase of wireless communication techniques, devices, and applications, the support of high data rate and quality of service (QoS) requirements for one or multiple users in a wireless communication network has drawn significant attentions. This topic has been studied extensively at either the physical (PHY) layer or the medium access control (MAC) layer. But, the studies should really be done by taking both layers into considerations since the two layers are intimately coupled. Among the relatively rare cross-layer designs, a single user scenario is considered in [1–3], where a packet could be dropped either at the transmitter due to the finite size of buffer, or at the receiver due to the bad channel condition. In [2], an adaptive modulation and coding (AMC) scheme is proposed to minimize the packet drop rate. In [3], the packet retransmission is considered, and the packet error rates (PERs) of different retransmission indices can be different. A downlink orthogonal frequency division multiple access (OFDMA) system is considered in [4–7]. A resource allocation scheme is proposed in [4] to design the power allocation and user allocation of each subcarrier to maximize the system capacity and satisfy the power constraints, QoS constraints, etc. In [5–7], the downlink OFDMA system is considered as the second users (SUs) in a cognitive radio (CR) system. Therefore, the interference suffered by the primary users (PUs) becomes the new constraints in the optimization problem. In [8], a multiple-input multiple-output (MIMO) single user system is considered. The data rate is adjusted to minimize the data loss from both capacity outage and buffer overflow.

All of the techniques above optimize a desired parameter of a single user (either in a single user scenario, or as a transmitter in a downlink scenario). Furthermore, that user is in a contention free environment, though there could be some co-existing PUs. In reality, however, there are some systems where multiple users need to compete for the channel access. An example could be a distributed coordination function (DCF) or hybrid coordination function (HCF) network, where every station (STA) may work in a carrier sense multiple access with collision avoidance (CSMA/CA) protocol and a binary exponential backoff procedure. In that case, optimal schemes in a single user scenario would not be optimal any more. For example, minimizing the PER may not maximize the throughput, which is affected by the collision probability as well. Furthermore, the choice of any STA affects the performance of every STA in the network, which makes any optimization problem for these systems highly nonlinear and complicated. Since DCF is one of the main MAC protocols that are employed by the IEEE 802.11 standards, the performance of a DCF network has been widely analyzed. A foundational two-dimensional Markov chain model is proposed by Bianchi in [9]. Based on this model, several follow-up models have been proposed in [10–14]. The PER is considered

---

[*] Correspondence to: 84 Fox Pl, NY, 11801, USA.
*E-mail addresses:* zy282@nyu.edu (Z. You), itl211@nyu.edu (I.-T. Lu).

https://doi.org/10.1016/j.phycom.2018.03.005
1874-4907/

in [10,11]. The buffer of a finite size is considered in [12,13]. The concept of the frozen backoff counter is considered in [11,13,14]. The optimization schemes for DCF systems have been proposed, too. For example, the optimal setting of contention windows to achieve the maximum throughput is discussed in [9,15]. The backoff procedure was modified in [16]. The DCF packet transmission procedure was modified in [17]. The relation between the maximum throughput and the system fairness is analyzed in [18–21]. The DCF network throughput optimization schemes with fairness constraints have been proposed in [22,23]. Schemes with similar goals and constraints are also proposed for TXOP in [24] and for AP association in [25], respectively. [22–25] adjust the contention windows and the packet durations to achieve a tradeoff between the throughput and the fairness, which is defined in different ways. (The fairness definitions will be discussed in the next section.)

However, all of these optimization schemes above assume that the PER is the same for every STA in the system. (Actually, most of them assume that the PER is always zero). Since the PER is unchangeable, the data rate is also unchangeable. With the same PER and data rate assumption in the MAC layer model, there is no base for differential physical layer designs (such as modulation coding scheme (MCS) and/or packet length selection) for different STAs in different propagation conditions. Furthermore, all of those schemes above propose to adjust the size of the minimum contention window, which is usually fixed as it is defined in various 802.11 amendments, e.g., 31 in 802.11b and 15 in 802.11n/ac. Our goal is to optimize the network throughput without modifying the widely used DCF mechanism.

Therefore, we had proposed a DCF model in [26] to analyze the performance of all individual STAs with different PERs. Now, in this paper, with the model in [26] at hand, we propose a novel cross-layer design for the STAs in a DCF network. Based on the PHY layer information such as the channel conditions, decisions are made to achieve the goals in MAC layer. By choosing the MCS and the packet duration for each STA, we try to maximize the total throughput of all STAs. Meanwhile, the scheme also satisfies the minimum throughput or delay requirement of each STA, thus addressing the fairness and delay issues.

The main idea of the proposed design is that the PER of a STA changes as its MCS and/or packet duration change. Subsequently, the average collision probability and transmission time of the entire system change. Thus, we can change the MCS and/packet duration of a STA to change the total throughput and the minimum individual STA throughput, where the latter is related to the measures of fairness and delay. The design can be used to generate a performance benchmark for the STAs in the DCF network. As WIFI technologies continue to flourish rapidly [27], such a cross-layer design can be useful in various system performance analyses and optimizations for DCF networks.

The outline of this paper is given below. Section 2 lays the system description and the preliminary formulation foundation of the Markov chain model in [26]. Section 3 presents our proposed design. Numerical results and discussions are shown in Section 4. Conclusion is made in Section 5.

## 2. System description and preliminaries

### 2.1. System description

We consider a system where there are $N$ STAs contending for the channel access. The system works in DCF protocol. Only one STA is allowed to transmit at a time. For example, the system could be a basic service set (BSS) consisting of one access point (AP) and several users. It is assumed that every STA has infinite information bits to transmit. In other words, followed by a successful packet transmission, a STA always contends for the channel access to transmit its next packet. It is also assumed that every STA can hear each other so that there is no hidden node in the system. $R$ denotes the maximum retransmission limit. $CW_j$ $(j = 0, 1, \ldots, R)$ denotes the contention window in the $j$th retransmission. The number of information bits in a packet from the $k$th STA is denoted as $B_k$. The duration of a packet (including both the header and the data) is denoted as $T_{P_k}$. The probability that a packet from the $k$th STA and the $i$th retransmission attempt suffers error is denoted as $p_{e_{i,k}}$.

### 2.2. Throughput, fairness, and delay

We propose a scheme that deals with three issues, which are the throughput, fairness, and delay. The throughput of a STA is defined as the number of information bits that are successfully transmitted (without collision or error) from that STA per second. The total throughput of the network is the sum of the throughput of every STA. (Note that the instantaneous throughput is a random variable and the average throughput is a deterministic value. Without specific clarifications, the throughput considered here is in the sense of ensemble average.) Our scheme aims to maximize the total throughput. Fairness has been defined in different ways. For example, the temporal fairness is defined in [18]. The proportional fairness is defined in [19]. And the max–min fairness for WLAN system is discussed in [21]. In this paper, fairness is assessed by the minimum throughput of all the STAs in the network, which usually has the worst channel conditions. Our scheme can ensure that the throughputs of those STAs satisfy a pre-determined constraint. The delay is the time that is required to successfully transmit a data stream of certain length. Due to the random backoff generation mechanism, the delay is also a random variable. Most optimization schemes work on the average delay. In addition to the average delay, our scheme can let a desired proportion of the delays smaller than a pre-determined constraint by setting the throughput constraints for the STAs. These three goals above are achieved by designing the MCS and packet duration for every STA, denoted as $\{MCS_k\}$ and $\left\{T_{P_k}\right\}$. So, once $\{MCS_k\}$ and $\left\{T_{P_k}\right\}$ are chosen, the PER, $\left\{p_{e_{i,k}}\right\}$, is also fixed, where $i$ is the retransmission index, and $k$ is the STA index. The mapping from $\{MCS_k\}$ and $\left\{T_{P_k}\right\}$ to $\left\{p_{e_{i,k}}\right\}$ is assumed to be known by the designer. In order to calculate the throughput of a STA in a DCF network, we employ the Markov chain model in [26], which will be described in the next subsection.

### 2.3. Markov chains

A model has been proposed in [26] to evaluate the MAC performance of each STA in the system when the PER $p_{e_{i,k}}$ is known. The contention process of each STA is modeled by a two-dimensional ergodic Markov chain. The state of a STA indicates the retransmission stage and the backoff counter of that STA at the current *model slot* index. The model slot used in this model corresponds to one of the following:

1. an idle backoff slot, where no STA transmits,
2. a time interval including one or more consecutive successful transmissions and a final extra backoff slot,
3. a time interval including one or more consecutive successful transmissions, followed by a transmission which suffers error, plus an Extended Interframe Space (EIFS) and an extra backoff slot,
4. a time interval including a collision or an error transmission, followed by an EIFS, plus an extra backoff slot.

Let $r_k(t)$ and $c_k(t)$ represent the value of the retransmission stage and the value of the backoff counter stage of the $k$th STA at the model slot index $t$.

The bi-dimensional random contention process of a STA, e.g., the $k$th STA, is shown in Fig. 1, where the arrows connecting