DCF performance analysis of open- and closed-loop adaptive IEEE 802.11n networks

Gabriel Martorell, Felip Riera-Palou, and Guillem Femenias
Mobile Communications Group – University of the Balearic Islands (Spain)
Email: {gabriel.martorell,felip.riera,guillem.femenias}@uib.es

Abstract—This paper presents a MAC-PHY cross-layer semi-analytical model allowing the goodput performance evaluation of open- and closed-loop link adaptive schemes in WLAN environments based on IEEE 802.11n. Without loss of generality, this paper considers the well-known open-loop automatic rate fallback (ARF) algorithm and a closed-loop fast link adaptation (FLA) algorithm based on exponential effective SNR mapping (EESM). In contrast to previous works, the channel conditions are not assumed to be ideal and users are allowed to use different transmission modes. Results are presented demonstrating the accuracy of the proposed model. Moreover, results serve to show the superiority of closed-loop over open-loop techniques, specially when the number of stations contending for the medium grows.

I. INTRODUCTION

The IEEE standards committee has published the final version of IEEE 802.11n wireless local area network (WLAN) standard [1] as a new amendment of IEEE 802.11 [2]. Compared to previous WLAN standards, this new norm allows higher throughputs and the fulfillment of more stringent quality of service (QoS) requirements. This amendment specifies enhancements to the IEEE 802.11 physical layer (PHY) and the medium access control (MAC) sublayer, most notably, the use of multiple antennas at the transmitter and receiver (MIMO) and frame aggregation, respectively. Additionally, it incorporates a feedback control channel between transmitter and receiver that enables closed-loop adaptive mechanisms.

Adaptation plays a crucial role in facing the time varying nature of the wireless channel. Adaptive mechanisms allow the reconfiguration of system parameters in order to exploit the available instantaneous capacity while satisfying QoS constraints. One of the most widely used reconfiguration techniques is adaptive modulation and coding (AMC), which selects an appropriate modulation and coding scheme (MCS) in response to changes in the environment or system behaviour. AMC algorithms can be broadly categorized as open- or closed-loop, depending on whether an explicit feedback channel exists between receiver (Rx) and transmitter (Tx). Open-loop setups operate in an heuristic manner and their rate of adaptation tends to be slow with respect to channel changes, thus differentiating the fulfillment of QoS constraints. In contrast, closed-loop mechanisms track more accurately the channel behavior and their response is more reactive to rapid channel variations.

Most IEEE 802.11-based systems employ the distributed coordination function (DCF) at the MAC sublayer and adopt open-loop AMC policies such as automatic rate fallback (ARF) [3] and its variations (e.g. CARA [4], SARA [5]). Owing to its simplicity, ARF is by far the most popular algorithm in use. However, the DCF scheme does not differentiate between collisions and transmission failures caused by poor channel conditions. Consequently, when the system experiences a high collision probability, ARF tends to use the lowest transmission rate even if the channel conditions are favorable to use much higher transmission modes (see for example, [4], [6], [7], [8]). Other adaptive strategies have been proposed to solve the ARF problem, but they may require frame format modifications, different transmission schemes (RTS/CTS), or the use of channel quality indicators (e.g. signal strength indicator) and, in fact, none of them has achieved widespread use in current WLAN systems.

Recent studies have demonstrated that, in the context of IEEE 802.11n, the use of closed-loop techniques such as fast link adaptation (FLA) offers important benefits in terms of physical layer throughput [9], [10]. However, current literature does not explore how this improvement reflects on the MAC goodput of a FLA-based system. This paper presents an analytical framework that can be used to assess the goodput performance at the MAC layer of both, open- and closed-loop adaptive schemes in the context of a modern WLAN system based on MIMO-OFDM physical layer. Building upon the model introduced in [11], where a performance study of the DCF over IEEE 802.11 under ideal transmission conditions was presented, the analysis proposed in this paper goes one step beyond by also considering the effects of adaptive modulation and coding mechanisms, and completing the work of [12], where an initial consideration of transmission errors was presented. As in [11], it is assumed that no hidden nodes exist in the cell under consideration. Numerical results clearly indicate the superiority of FLA over ARF in terms of goodput for more than two active users per cell. Moreover, these results illustrate the strengths and weaknesses of each adaptive scheme, thus providing hints on how spectral efficiency can be maximised as a function of the operating conditions.

The rest of the paper is structured as follows. Section II describes the system model under consideration. Section III briefly reviews the two adaptive schemes covered in this work. Section IV presents the analytical framework used to analyze the system goodput. In Section V numerical results are presented comparing the performance of open- and closed-loop schemes. Finally, in Section VI, the main conclusions of this study are summarized.
II. SYSTEM OVERVIEW

A. Physical layer description

Our study focuses on the IEEE 802.11n standard [1], whose PHY layer is based on MIMO-OFDM. The MIMO component exploits multiple transmit and receive antennas to increase the system capacity or reliability, depending on the specific technique employed [13], e.g., space-time block coding (STBC), space division multiplexing (SDM), cyclic delay diversity (CDD) and/or combinations of them. At the transmitter side, information bits are first encoded with a rate \( R = \frac{1}{2} \) convolutional encoder with generator polynomials \([133, 171]\) and then punctured to one of the possible coding rates \( R_m \in \{1/2, 2/3, 3/4, 5/6\} \). Depending on the selected MIMO configuration, the resulting bits are demultiplexed into \( N_s \) spatial streams. For each stream, the coded bits are interleaved and then mapped to symbols from one of the allowed constellations (BPSK, QPSK, 16-QAM or 64-QAM). In accordance with the selected MIMO configuration, the symbols are then either STBC encoded or antenna mapped on the available \( N_T \) transmit antennas. The resulting symbols are finally supplied to a conventional OFDM modulator consisting of an IFFT and the addition of a guard interval. For simplicity of exhibition, this paper focuses on a \( 2 \times 2 \) MIMO system \((N_T = 2 \text{ and } N_R = 2)\), implying that MCSs with \( N_r = 1 \) and \( N_s = 2 \) spatial streams employ STBC [14] and SDM [15], respectively.

At the receiver side, Alamouti decoding or Minimum Mean Square Error (MMSE) detection is applied depending on whether STBC or SDM has been employed. In either case, the detector extracts soft information in the form of log-likelihood ratios (LLRs) that, after suitable de-interleaving/de-parsing, can be exploited by a soft Viterbi decoder [10].

B. MAC layer description

The IEEE 802.11 MAC specifies three different medium access control (MAC) mechanisms for WLANs, namely the contention-based distributed coordination function (DCF), the point coordination function (PCF) and the hybrid coordination function (HCF). The DCF is the mandatory MAC mechanism for the IEEE 802.11 standard [2]. It is a random access scheme based on the carrier sense multiple access with collision avoidance (CSMA/CA) protocol that incorporates a binary exponential backoff (BEB) algorithm to manage the retransmission of collided packets.

The basic access technique is the most extensively used access scheme in DCF [4] and the one central to the study presented in this contribution. For completeness it should be mentioned that there is another DCF access technique, namely RTS/CTS, which is of mandatory use when the packet length is larger than a given packet length threshold. For shorter packet lengths or high transmission rates, its overhead considerably affects the system performance.

At the MAC sublayer, the data information from the upper layer is treated as MAC service data units (MSDUs), which are conveniently fragmented if their size exceeds the fragmentation threshold, and subsequently converted to MAC protocol data units (MPDUs) through the MAC headers addition. At the physical layer, the MPDUs are treated as physical layer data units (PSDUs) and processed by the physical layer convergence procedure (PLCP) to form the PLCP protocol data units (PPDUs) prior to transmission. The 802.11n MPDU and PPDU frame formats are presented in Fig. 1.

C. DCF transmission time

According to the DCF operation, the elapsed time for the successful transmission of an \( L \)-bit MPDU using MCS \( m \) is

\[
T_s(m, L) = T_{Tr}(m, L) + T_{Prop} + T_{SIFS} + T_{ACK+HTC}(m) + T_{Prop} + T_{DIFS},
\]

where \( T_{SIFS} \) and \( T_{DIFS} \) are 802.11n time constants defined in [1] (see Table I), \( T_{Prop} \) is the propagation time, and \( T_{Tr}(m, L) \) is the elapsed time in the MPDU transmission that can be defined as

\[
T_{Tr}(m, L) = T_{Preamble} + N_{sym}(L)T_{sym},
\]

with

\[
N_{sym}(L) = m_{STBC} \left\lceil \frac{L + 22}{m_{STBC}N_{DBPS}(m)} \right\rceil,
\]

where \( m_{STBC} = 2 \) if STBC is used and \( m_{STBC} = 1 \) otherwise, \( N_{DBPS}(m) \) is the quantity of bits forming each OFDM symbol as defined by MCS \( m \), \([z]\) denotes the smallest integer greater than or equal to \( z \), and \( N_{sym}(L) \) is the number of OFDM symbols involved in the transmission of an \( L \)-bit packet. Similarly, the time required for the transmission of an ACK+HTC frame \(^1\) using PHY mode \( m \) is given by

\[
T_{ACK+HTC}(m) = T_{Preamble} + N_{sym}(20 \cdot 8)T_{sym}.
\]

A collision occurs when two or more stations transmit on the same slot, finishing when the longest transmission

\(^1\)Wrapper control frame that encapsulates the ACK and the HT control field required to feedback the MCS selection.
of the collided stations ends. That is, its duration depends on the MCS and MPDU length corresponding to the longest transmission, denoted by \( m^* \) and \( L^* \), respectively. Therefore, the collision duration can be mathematically expressed as

\[
T_c(m^*, L^*) = T_{Tr}(m^*, L^*) + T_{prop} + T_{DIFS}. \quad (5)
\]

The MPDU error transmission time \( T_e(m, L) \) is the elapsed time in a transmission that experiences errors without collisions, and can be computed as

\[
T_e(m, L) = T_{Tr}(m, L) + T_{prop} + T_{DIFS}. \quad (6)
\]

In this model, due to its negligible probability of occurrence, we have not considered the possibility of an error in the ACK transmission. The ACK transmission takes place under the same system conditions as the packet being acknowledged, e.g., using the same MCS and suffering similar channel conditions. However, its packet size is considerably smaller and therefore, its error probability can be considered negligible.

III. ADAPTIVE MODULATION AND CODING

A. Automatic rate fallback (ARF)

This algorithm adapts the transmission rate according to the number of consecutive transmission failures and successes, both reported by the ACK mechanism. The transmission rate is decreased after two consecutive transmission failures and increased after either ten consecutive successful packet transmissions or a timeout. This timeout counter is reset after a transmission rate change or after a transmission failure, in order to improve the system adaptation during long intervals of inactivity [3]. An acceptable timeout value lies in the range of 50-200ms [16]. We further note that following a rate increase, the next data transmission is deemed as a probing transmission for the new mode. If an ACK is not received for this probing packet, the system falls back to the previous data rate.

In order to implement ARF in IEEE 802.11n it is necessary to determine the available rates in the MCS set, denoted by \( \mathcal{M} \). In contrast to previous IEEE 802.11 standards, in 802.11n different MCSs \( \in \mathcal{M} \) can provide the same transmission rate, but only one of them can be used by the ARF algorithm. For this reason, the MCSs in \( \mathcal{M} \) are reordered according to their transmission rate [10]. For those rates that can be attained using either SDM or STBC, only the STBC MCS is kept as it can be shown to be more robust against channel variations [14]. The reordered/pruned MCS set will be denoted by \( \tilde{\mathcal{M}} \).

B. Fast link adaptation (FLA)

Fast link adaptation is a closed-loop technique that relies on the availability of the feedback channel from the receiver to the transmitter. The main idea behind FLA is that the receiver, thanks to an accurate knowledge of the channel response, can compute a reliable prediction of the error rate for all available MCSs and choose the one maximising the instantaneous throughput while satisfying QoS constraints in the form of outage packet error rate probability. The selected MCS can then be communicated to the transmitter via the feedback channel. In this work we assume the use of the methodology presented in [10], where link performance prediction for each MCS is based on the exponential effective SNR mapping (EESM) [17]. Using this approach, the EESM for a given MCS can be easily associated to PER using look-up tables that have been previously computed during an off-line calibration phase.

Note that in real scenarios there will be some latency between the reception of MCS selection feedback and its use in the forthcoming transmission. If this delay exceeds the channel coherence time, the provided feedback might prove useless. In order to counteract the effects of using expired MCS feedback, the station (STA) decreases the transmission data rate when the MCS feedback delay exceeds a fixed timeout. This timeout is set to a value close to the channel’s coherence time so that, prior to the timeout expiration, the channel response can be safely assumed to be similar to the one on which the MCS selection was made. The STA will decrease again the transmission data rate in subsequent packet retransmissions (if any) until the packet is successfully sent across.

IV. GOODPUT ANALYSIS

Following [11], the goodput analysis focuses on the saturation region, defined as the operation point where each STA has always new packets to transmit. The system saturation goodput \( S \) can be defined as

\[
S = \frac{E[\text{payload information in a slot}]}{E[\text{Duration of a slot}]}, \quad (7)
\]

where \( E[\cdot] \) denotes statistical expectation. The duration of a slot refers to the time interval between two consecutive backoff counter decrements.

In any given slot, one out of four events can occur: a successful packet transmission \( s \), an error packet transmission \( e \), a collision \( c \) or an idle slot \( i \). From the point of view of the BEB algorithm, error transmissions and collisions are undistinguishable. The conditional probability of the union of these events can be computed as

\[
p = 1 - (1 - \zeta_u)(1 - \tau)^{n-1}, \quad (8)
\]

where \( n \) is the number of active stations in the scenario, \( \zeta_u \) is the error transmission probability for the considered AMC algorithm, averaged on a per-user basis, and \( \tau \) is the stationary probability that a particular STA transmits in a given slot. This transmission probability can be obtained as,

\[
\tau = \frac{2(1 - 2p)}{(1 - 2p)(CW_{min} + 1) + pCW_{min}(1 - (2p)^{m_{max}})}, \quad (9)
\]

where \( CW_{min} \) is the minimum congestion window and \( m_{max} \) is the maximum backoff stage (see Table I). Notice that \( p \) and \( \tau \) can be obtained by solving the nonlinear system formed by eqs. (8) and (9).

Using \( \tau \), the probability that only one STA transmits on a given slot is

\[
P_s = n\tau(1 - \tau)^{n-1}. \quad (10)
\]

Furthermore, the probability that a given slot is idle (e.g. not used by any user) is given by

\[
P_i = (1 - \tau)^n. \quad (11)
\]
Observing all possible events, only the successful packet transmission increases the payload information, any other event leads to a goodput degradation. Consequently, generalizing the goodput expression in [12, eq. (50)] by taking into account multiple transmission modes, we obtain
\[
S = \frac{(1 - \zeta_u)P_s E[L_p]}{P_i \sigma + (1 - \zeta_i)P_i T_s^{(n,L)} + \zeta_i P_i T_s^{(n,L)} + (1 - P_s - P_i)T_c^{(n,L)}},
\]
where \(L_p = L - L_h\) is the packet length in bits, \(L_h\) is the MAC header length, \(\zeta_i\) denotes the average packet error probability for a given slot, and \(T_s^{(n,L)}, T_c^{(n,L)}\) and \(T_c^{(n,L)}\) represent the average elapsed time for successful, colliding and error transmissions, respectively. These elapsed times depend on the number of contending stations \(n\) and the MPDU length \(L\), and can be computed as
\[
T_x^{(n,L)} = \sum_{m=0}^{\lfloor M \rfloor} P_x^{MCS}(m,n)T_x(m,L),
\]
where \(|M|\) denotes the cardinality of set \(M\), and \(P_x^{MCS}(m,n)\) is the probability of using MCS \(m\) when the system is in event \(x \in \{s, c, e\}\). Notice that these probabilities are AMC dependent.

Probabilities \(P_x^{MCS}(m,n)\) can be determined either by simulation or analytically. In the analytical approach the statistical behaviour of the effective SNR must be characterized and then the switching thresholds between transmission modes should be determined. In this work, and for the sake of simplicity, the simulation-based approach has been followed deeming the resulting model to be considered as semi-analytical. Similarly, \(\zeta_i\) and \(\zeta_u\) are determined by simulation for each AMC.

V. NUMERICAL RESULTS

In order to validate our semi-analytical model, an IEEE 802.11n system-level Matlab simulator has been implemented using the link-level parameters derived in [10]. It should be stressed that this model is considerably more realistic than the one in [11] since it allows non-ideal channel behaviour and AMC mechanisms to be taken into account. In this paper we concentrate on the performance evaluation of the uplink scenario where, nevertheless, MAC control frame responses from access point (AP) to STA are also accounted for. Different scenarios have been generated by uniformly distributing \(n\) static users in a circular area of radius \(R_{max}\) centered around the AP and then determining the individual channel response from each user to the AP. To this end, the MIMO channel generation tool presented in [18], parameterized with each user’s distance to the AP, has been employed. The maximum radius \(R_{max}\) has been set to a value that ensures the avoidance of the hidden terminal problem and precludes the utilization of the no transmission mode (available in FLA).

The main protocol parameters used to obtain both analytical and simulation results are summarised in Table I. Furthermore, a constant packet payload size of \(L_p = 1500\) bytes has been used. The physical layer uses only the first 16 MCS modes of IEEE 802.11n (MCS0-MCS15), achieving rates up to 130 Mbps [1]. The ARF and FLA timeouts have been set to 60 ms and 20 ms, respectively. In order to obtain an accurate estimate of the average system performance, \(N_{sim} = 100\) simulation runs of duration \(t_{sim} = 22\) seconds have been generated for each value of \(n\).

Figures 2 and 3 present the overall goodput performance and the conditional probabilities \(p\) and \(\tau\) (see eqs. (8) and (9)), respectively. In both figures a very accurate match between the semi-analytical and simulated system performance metrics for both FLA- and ARF-based schemes can be appreciated. Noticeably, Fig. 2 illustrates the goodput advantage of FLA-based adaptation over the ARF approach. As the number of users in the system grows, the probability of collision increases leading to a very different behaviour of the two systems. While ARF quickly lowers its transmission mode even when the channel is in good condition, FLA is able to match the transmission rate to the most recent channel state information it has acquired. Thus, irrespective of the number of users in the system, a high transmission mode is used in FLA when the channel is favourable. Consequently, ARF experiences a drastic fall in throughput when 3 or more users are present in the network, whereas FLA exhibits a graceful degradation as more users enter the system. Remarkably, for more than 2 active users, FLA goodput more than trebles that of the ARF-based scheme.

Table II shows system parameters that have been obtained by means of simulation, including \(\zeta_u\), \(\zeta_i\), \(T_s^{(n,L)}\), \(T_c^{(n,L)}\) and \(T_c^{(n,L)}\), for both ARF and FLA adaptive schemes under
incorporate the effects of non-ideal channels (e.g., channel errors) and the availability of different transmission modes. This model allows an exhaustive comparison to be made in terms of goodput performance between two of the most popular adaptive strategies used in WLAN systems, namely, ARF (open-loop) and FLA (closed-loop). Results clearly show that as the number of users in the system grows, the FLA-based adaptation proves to be much more robust to collisions than ARF. This effect can be mostly noticed by the fact that whereas ARF-based schemes suffer a dramatic reduction in goodput for more than 2 users, the FLA-based strategy exhibits a very graceful degradation in goodput thanks to a more accurate rate selection in the presence of collisions. Notably, for most system loads, FLA adaptation offers a goodput that more than trebles that of ARF.

VI. CONCLUSIONS

This paper has introduced a semi-analytical model to validate the performance of open- and closed-loop link adaptation schemes in WLAN systems based on the DCF at the MAC sublayer and on multirate MIMO-OFDM at the physical layer. Unlike previous works in the area, this model is able to different network loads. Notice that in ARF the mean elapsed time grows with the number of users motivated by the more frequent use of the lower transmission modes as the number of collisions increases. This effect is illustrated in Fig. 4, where the probability of using the different transmission rates is shown for three system loads. Regarding FLA, and as it could be anticipated, the values of the mean elapsed times are almost unrelated to the number of users in the system. Only a slightly increasing trend can be observed in $T_s^{n,L}$ and $T_c^{n,L}$ that can be attributed to the mismatch between channel information and selected mode due to the larger feedback delays when the number of users increases.

REFERENCES