Exploiting Frame Aggregation to Enhance Access Point Selection

Nour El Houda Bouzouita, Anthony Busson, Hervé Rivano

HAL Id: hal-03362285
https://inria.hal.science/hal-03362285

Submitted on 1 Oct 2021

Exploiting Frame Aggregation to Enhance Access Point Selection

Nour El Houda Bouzouita
Univ Lyon, UCBL, EnsL, CNRS, Inria,
LIP UMR 5668
Lyon, France
nour-el-houda.bouzouita@ens-lyon.fr

Anthony Busson
Univ Lyon, UCBL, EnsL, CNRS, Inria,
LIP UMR 5668
Lyon, France
anthony.busson@ens-lyon.fr

Herve Rivano
Univ Lyon, INSA Lyon Inria,
CITI EA 3720
Villeurbanne, France
herve.rivano@inria.fr

ABSTRACT

IEEE 802.11 (known as Wi-Fi) has emerged as a vital wireless network access technology for mobile devices. By providing the potential for high connectivity speeds, this technology has led to a dramatic rise in the number of access points (APs). In such environments, mobile devices have the choice to join several Wi-Fi networks. Despite its importance to user Quality of Experience (QoE), the AP selection is still trivial since it focuses at best on the received signal strength if not only the user's history. Crucial metrics that capture the overall dynamics of the AP load condition, such as the network load, are not taken into account. In this paper, we propose to use the Busy Time Fraction (BTF) as a metric to choose the best AP to attach to. The BTF level of a given channel is inferred based on the frame aggregation scheme proposed since the 802.11n standard. In this regard, we propose an analytical model based on a Discrete-Time Markov chain that discerns the theoretical DownLink aggregation levels for probe traffic concurrent to cross traffic. We validate the accuracy of our proposed approach against ns-3 simulations under several scenarios.

CCS CONCEPTS

• Networks → Wireless local area networks; Network performance modeling.

KEYWORDS

IEEE 802.11 networks, aggregated MAC protocol data unit (A-MPDU), Performance evaluation, Network load, Markov chain

1 INTRODUCTION

As one of the cornerstones for delivering flexibility and high speeds, Wireless-Fidelity (Wi-Fi) or IEEE 802.11 has become the most used

technology for accessing the Internet. Nearly, all companies, universities, and homes use Wi-Fi to satisfy their connectivity needs, which has led to the explosion of wireless data usage and the colossal rise of Access Points (APs), smartphones, and various mobile devices. According to [2], Internet users represented 3.9 billion in 2018, and they will exceed 5.3 billion by 2023, while the number of Wi-Fi hotspots will grow four-fold from 169 million hotspots in 2018 to nearly 628 million public Wi-Fi hotspots by 2023. With this huge proliferation, selecting the AP that offers the required performance is still a striking concern and a critical ongoing challenge.

In dense environments, a device seeking connectivity must choose among all the APs that are within its radio range. This association is initiated by the device's operating system based on the best Received Signal Strength Indicator (RSSI) or the user's history or preferences. RSSI or equivalent metrics may help to estimate the physical transmission rate that will be used. But, these simple criteria do not take into account the load of the channels, APs, or any information relative to the network performance imposed by the other stations and the current traffic.

In contrast to these metrics, the network load seems to be an excellent fit since it conditions the throughput and gives a broader view of the network state independently of the transmission conditions of the device. This latter can be expressed in many ways. In this work, we consider the Busy Time Fraction (BTF), defined as the fraction of time the wireless medium is sensed busy due to successful or unsuccessful transmissions. It captures concurrent transmissions, constituting the AP load, as well as inter-network interference.

This paper presents a method allowing a device to estimate the BTF of the channels used by the AP in its vicinity. We consider an active probing approach to estimate this quantity between two devices connected to the same AP. These two devices are not the only recipients of these measures, but they are supposed to be made available in a public database or through a crowd-sensing application in which participant mobile devices measure and share the load of their nearby networks. By building collective knowledge, other devices using the same application will be able to choose their AP as a function of the load and produce their own measures.

The classical channel load estimation tools based on time dispersion cannot be used in Wi-Fi [3, 21]. Indeed, in the most modern Wi-Fi versions (802.11n, ac, ax, and the most recent roll-up [7]), a frame aggregation scheme has been introduced. This scheme allows an AP or a station (STA) to send several frames through a unique access to the medium and thus making inter-packet arrival null. Hopefully, the load of the channel conditions the aggregation level. In this paper, we propose a Markov chain that describes the aggregation level, defined as the mean number of sub-frames within

the same aggregated frame, as a function of the load. A device can then compare the actual aggregation level to the analytical model to estimate the network load. Our model is compared to ns-3 simulations [20] allowing us to consider realistic scenarios. For the considered scenarios, we show that the mean aggregation level is an accurate metric to estimate the load of an AP.

The paper is organized as follows. Section 2 describes in detail the background of this work. Section 3 presents our proposed analytical model to estimate the DownLink frame aggregation levels of the probe traffic. Section 4 is dedicated to the performance evaluation and numerical results. Papers that relate to the studied problem are discussed in Section 5. We conclude this paper in Section 6.

2 BACKGROUND

Before plunging into the details of our proposed system, we present several key background concepts. First, we give an overview of the IEEE 802.11 frame aggregation scheme in Section 2.1. We then give a brief description of how the AP queue handles packets under the frame aggregation.

2.1 Frame aggregation overview

Ever since their first introduction in 1997, IEEE 802.11 network has considerably enhanced its throughput by coming with new technological advances to improve the physical (PHY) and the Medium Access Control (MAC) layers. At the MAC layer, one of the significant steps forward was the introduction of the frame aggregation scheme by the 802.11n standard. Since then, this latter has become the default mechanism of sending data on modern Wi-Fi (802.11n, ac, ax).

The general concept of frame aggregation is to allow several packets, that are destined to the same destination, to be clustered and sent out as one aggregated frame, and thus reducing the access time (Distributed Inter Frame Space (DIFS), Back-off, etc.) and the PHY overhead as all the aggregated packets share one PHY header. Frame aggregation can be achieved at two levels: at the Aggregate Mac Protocol Data Unit (A-MPDU) and the Aggregate MAC Service Data Unit. Since A-MSDU is rarely implemented, we focus on the A-MPDU aggregation.

Figure 1 depicts the structure of an aggregated frame. The A-MPDU aggregation consists in sending several sub-frames (A-MPDU) with a common Wi-Fi physical header (PHY Header). Each sub-frame is composed of an MPDU delimiter, followed by the MPDU, consisting of its own MAC header, MAC payload (a packet or any layer 3 Protocol Data Unit), and Frame Check Sequence (FCS). Padding bytes may be appended at the end.
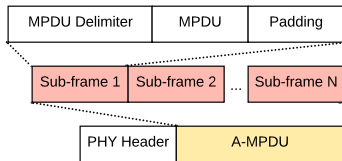

Figure 1: Illustration of A-MPDU frame aggregation.

2.2 AP queuing system under frame aggregation

Without frame aggregation and for a given enhanced distributed channel access (EDCA) class (an extension of the Distributed Coordination Function scheme that includes Quality of Service traffic prioritization for data packets), the AP follows a First In First Out (FIFO) per-packet scheduling method. With frame aggregation, the aggregated transmissions act as a batch scheduler which alters the timing characteristics of received packets. To explain this approach, we show an example in Figure 2. It consists of an AP queue with two interleaved sequences of packets sent to two different clients, A and B, on a Wi-Fi link. Ai defines the ith packet targeted to client A while Bi is the ith packet destined to client B. In the case depicted in Figure 2, the packet at the head of the AP queue is addressed to A. All A's packets (within the limit of the maximum number of aggregated sub-frames) will be aggregated and sent out as an A-MPDU. Then, an aggregated frame with B's packets is sent.
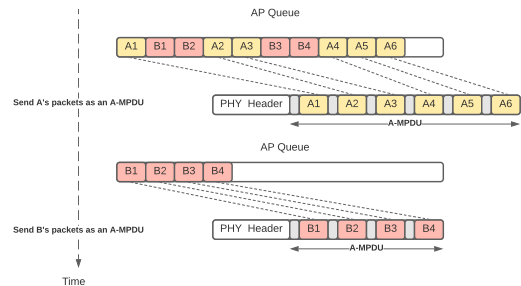
Figure 2: An Example of scheduling under frame aggregation.

3 SYSTEM DESCRIPTION

We consider a general Wireless Local Area Network (WLAN) based on IEEE 802.11 Distributed Coordination Function (DCF). Figure 3 depicts the scenario considered in our model. An AP and three stations operate on the same Wi-Fi channel. The AP is connected to a distribution system (a wired network) which is not modeled here. Each station or AP has its own physical transmission rate. It is a combination of the following parameters: the modulation type and the coding rate which is defined by the Modulation and Coding Scheme index (MCS), the number of spatial streams, the channel width, and the guard interval length.
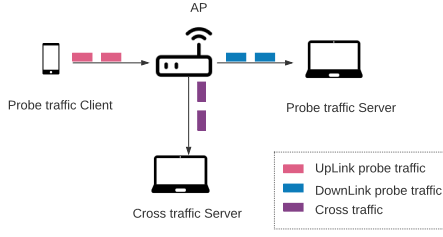
Figure 3: Scenario considered in the model. The probe traffic client and server estimate the current load. The cross traffic server aims to represent the current load of the system.

In our model, we distinguish three types of traffics:

- The UpLink (UL) probe traffic sent from the probe traffic client to the AP. It is a Constant Bit Rate (CBR) traffic, sent at regular interval d_p.
- The DownLink (DL) probe traffic forwarded to the probe server.
- The cross traffic coming from the distribution system (not represented in the figure) and sent to the cross traffic server. It is modeled by a constant bit rate source sending packets at regular interval d_c and managed by a unique queue.

The AP and the probe traffic client buffers are assumed to have a finite size. More precisely, we assume that when an aggregated frame is sent for a given destination, the buffer becomes empty for this destination. This assumption simplifies drastically the model and its computation. However, the simulation results show that it does not introduce a significant error. Table 1 summarizes the principal notation used in the paper.

Table 1: Principal notation

Parameter (unit)	Definition
T_{DIFS} (μs)	Distributed Inter Frame Space duration
T_{SIFS} (μs)	Short Inter Frame Space duration
T_{PHY} (μs)	The preamble and PHY header duration
FCS ($Bytes$)	Frame Check Sequence
$T_{BlockACK}$ (μs)	Required time to send the block acknowledgment
$T_{backoff}$ (μs)	Average backoff time
T_{slot} (μs)	Slot time
CW_{min}	Minimum size of the contention window
$P_{(i,j,k,u)(l,m,q,v)}$	Transition probability from state (i,j,k,u) to state (l,m,q,v)
$f(k)$ (μs)	Required time to send k DL probe aggregated packets
$g(k)$ (μs)	Required time to send k UL probe aggregated packets
$h(k)$ (μs)	Required time to send k cross traffic sub-frames
d_c (μs)	Inter-arrival time of cross traffic sub-frames
d_p (μs)	Inter-arrival time of probe traffic sub-frames
$AMPDU_{AP}$	Maximum A-MPDU size for the AP
$AMPDU_P$	Maximum A-MPDU size for the probe traffic client

In this section, we derive a mathematical model that estimates the frame aggregation levels of the DL probe traffic in this scenario. This frame aggregation level is a function of the cross traffic. The model aims to be compared with the aggregation levels as measured

by the probe traffic server to deduce the cross traffic load (load of the AP or channel).

The frame aggregation level of the DL probe traffic depends on the competition for medium access. In particular, when the cross traffic increases, the DL probe traffic has to wait longer. More packets may hence accumulate in the queue of the AP and be aggregated when the next DL probe transmission occurs.

The proposed analytical model is a Discrete-Time Markov Chain (DTMC). We consider the Markov chain defined as $(X_n, Y_n, Z_n, S_n)_{n \geq 0}$. X_n describes the number of probe sub-frames contained in the AP queue at the moment of the nth frame transmission departure. The possible states are $\{0, .., AMPDU_{AP}\}$. Y_n defines the number of cross traffic sub-frames in the AP queue. The possible states are $\{0, .., AMPDU_{AP}\}$. Z_n represents the number of packets at the probe traffic client. The possible states are $\{0, .., AMPDU_P\}$. S_n describes the nth transmission. It takes three possible values $\{APC, APP, SP\}$. APP (Access Point Probe) is a DL transmission of probe traffic from the AP. APC (Access Point Cross) denotes a cross traffic transmission from the AP to the cross server, while SP (Station Probe) corresponds to a UL transmission from the probe traffic client.

3.1 Transition probabilities

Having defined the set of possible states for each process, we shall now derive the transition probabilities. The transition probabilities are denoted $P_{(i,j,k,u)(l,m,q,v)}$, and represent the probability to go from state $(X_n, Y_n, Z_n, S_n) = (i, j, k, u)$ to state $(X_{n+1}, Y_{n+1}, Z_{n+1}, S_{n+1}) = (l, m, q, v)$.

The transition probabilities depend on the time between two successive transmissions. As both UL probe and UL cross traffics are deterministic, this time sets the number of packets that arrived in the AP buffer and the probing client buffer between two transmissions and thus the number of frames that will be sent in the next aggregated frame. We therefore analyze the events that may occur between two successive transmissions.

The next stage of our modeling is to decide when a transition from one state to another is allowed and to compute its probability. Note that impossible transitions have zero probability and that non-null transition probabilities are computed assuming that all concerned nodes are equally likely to access the channel. For ease of illustration, we categorize the state transitions into the following three classes.

Class I:Transition from state APP. For the sake of clarity, we recourse to Figure 4a to describe the possible transition probabilities and their computations. It illustrates an example of possible events between two successive transmissions when the current state is APP (an aggregated frame containing probe packets is transmitted by the AP). In this example, at time $t = 0$, we start from state (i, j, k, APP) (i.e $(3,2,2,APP)$ in Figure 4a). The AP sends the i corresponding probe frames currently in its buffer ($i = 3$ in this example), aggregating them in a single A-MPDU using its current MCS. Its transmission duration is denoted $f(i)$ which is given in Appendix A. Note that the AP has a single buffer that contains at once the probe and the cross traffics but in order to clarify the explanation of the model, we distinguish between them in Figure 4a.

(a) Transition from APP
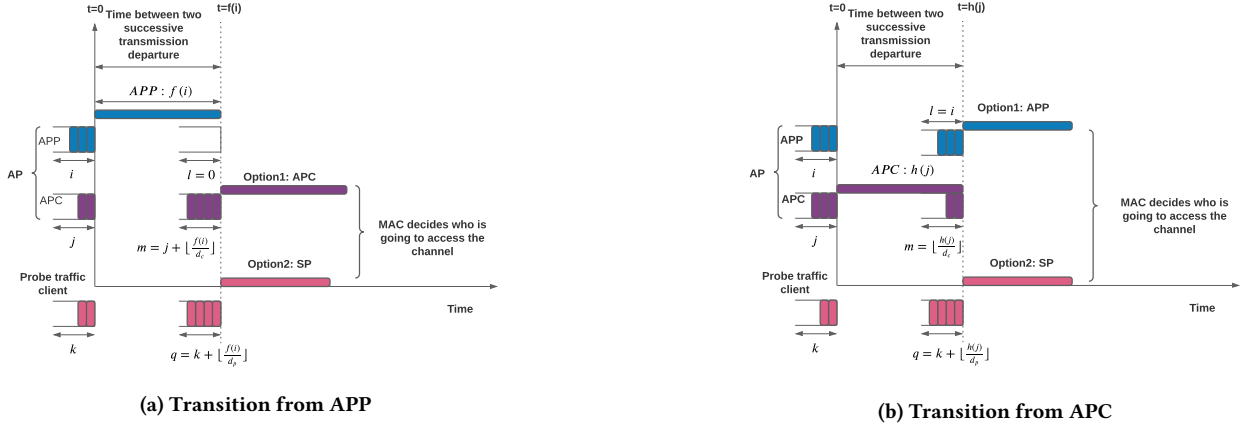
(b) Transition from APC

Figure 4: Possible transitions

During this transmission, the probe traffic client buffer receives $\lfloor \frac{f(i)}{d_p} \rfloor$ packets and the AP buffer receives $\lfloor \frac{f(i)}{d_c} \rfloor$ cross packets. As a result, at the end of the APP state, the probe traffic client buffer and the AP will contain respectively, $k + \lfloor \frac{f(i)}{d_p} \rfloor$ of probe packets and $j + \lfloor \frac{f(i)}{d_c} \rfloor$ of cross traffic packets ((4,4) in the Figure 4a).

At the end of this transmission, the buffer of the AP does not have probe frames to send ($X_{n+1} = 0$ almost surely), and another APP transmission is impossible ($S_{n+1} \neq APP$ almost surely). So from this state, only two transitions are allowed: to APC or SP with $X_{n+1} = l = 0$. It can occur only if $Y_{n+1} = m > 0$ and $Z_{n+1} = q > 0$ respectively.

We derive the non-null transition probabilities as follows. If the AP gains access to the channel, it will send the cross traffic currently in its buffer. For $m > 0$, the next transmission will be APC with probability:

$$P_{(i,j,k,APP)(0,m,q,APC)} = P\Big(S_{n+1} = APC | S_n = APP, Z_{n+1} = q,$$
$$Y_{n+1} = m > 0\Big) \cdot \mathbb{1}_{k + \lfloor \frac{f(i)}{d_p} \rfloor = q} \cdot \mathbb{1}_{j + \lfloor \frac{f(i)}{d_c} \rfloor = m}$$

where $\mathbb{1}_{condition}$ is the indicator function that equals to 1 if *condition* is true and 0 otherwise. $P\Big(S_{n+1} = APC | S_n = APP, Z_{n+1} = q, Y_{n+1} = m > 0\Big)$ denotes the probability that the event APC will occur given that the event APP has already occurred. In the interest of brevity, we postpone the computation of such probabilities to Appendix B.

Now, if the probe station gains access, the next event will be SP and the transition probability from APP to SP for $q > 0$ is given by:

$$P_{(i,j,k,APP)(l,m,q,SP)} = P\Big(S_{n+1} = SP | S_n = APP, Y_{n+1} = m,$$
$$Z_{n+1} = q > 0\Big) \cdot \mathbb{1}_{k + \lfloor \frac{f(i)}{d_p} \rfloor = q} \cdot \mathbb{1}_{j + \lfloor \frac{f(i)}{d_c} \rfloor = m}$$

Class II: Transition from state APC. Once again, when presenting the transition probabilities of this class, we resort to Figure 4b. It depicts a timeline of feasible events between two successive transmission departures when we start from the state APC.

The AP sends the j cross traffic frames currently in its buffer as an A-MPDU using the transmission rate associated to the cross traffic server. Its transmission duration is denoted $h(j)$. During the period $h(j)$, the probe traffic client buffer receives $\lfloor \frac{h(j)}{d_p} \rfloor$ packets and the AP buffer receives $\lfloor \frac{h(j)}{d_c} \rfloor$ cross traffic packets. At the end of the APC state, the probe traffic client buffer and the AP will thus contain respectively, $k + \lfloor \frac{h(j)}{d_p} \rfloor$ of probe packets and $\lfloor \frac{h(j)}{d_c} \rfloor$ of cross traffic packets.

Conversely to the previous class of transitions where only two possible transitions are allowed from the state APP, there are here three possible transitions under some conditions. S_{n+1} can be APP, APC, or SP.

First, we suppose that there will be another APC, the transition from APC to APC is deemed possible if and only if $X_n = i = 0$ and $Y_n = l = 0$: it is impossible to have two successive APC transmissions if $i > 0$ or $l > 0$ due to the AP queuing system detailed in Section 2.2. The transition probability from APC to APC with $Y_{n+1} = m > 0$ is:

$$P_{(0,j,k,APC)(0,m,q,APC)} = P\Big(S_{n+1} = APC | S_n = APC, X_n = X_{n+1} = 0,$$
$$Y_{n+1} = m > 0, Z_{n+1} = q\Big) \cdot \mathbb{1}_{\lfloor \frac{h(j)}{d_c} \rfloor = m} \cdot \mathbb{1}_{k + \lfloor \frac{h(j)}{d_p} \rfloor = q}$$

Now if we suppose that the AP gains the medium access to transmit the DL probe flow, the transition probability from APC to APP is possible only if $X_n = i = X_{n+1} = l$ since the AP cannot receive probe frames in its buffer during the transmission of the cross traffic and $X_{n+1} = l > 0$. The corresponding probability is given by:

$$P_{(i,j,k,APC)(l,m,q,APP)} = P\Big(S_{n+1} = APP | S_n = APC, X_{n+1} = X_n$$
$$= l > 0, Z_{n+1} = q\Big) \cdot \mathbb{1}_{\lfloor \frac{h(j)}{d_c} \rfloor = m} \cdot \mathbb{1}_{k + \lfloor \frac{h(j)}{d_p} \rfloor = q}$$

If we assume that the probe traffic client gains the medium access, the transition from APC to SP is also allowed only if $X_{n+1} = l = X_n = i$ and $Z_{n+1} = q > 0$. We have:

$$P_{(i,j,k,APC)(l,m,q,SP)} = P\Big(S_{n+1} = SP | S_n = APC, Z_{n+1} = q > 0,$$
$$X_{n+1} = X_n = l, Y_{n+1} = m\Big) \cdot \mathbb{1}_{\lfloor \frac{h(j)}{d_c} \rfloor = m} \cdot \mathbb{1}_{k + \lfloor \frac{h(j)}{d_p} \rfloor = q}$$

Class III:Transition from state SP. In order to derive the transition probabilities of the last class, we apply the same principle. Due to space limitations, we do not describe the details of the derived transition probabilities.

From the SP state, there are three possible transitions APP, APC, or SP depending on the competition for the channel resource. First, if the next transition is APP with $X_{n+1} = l > 0$, the transition probability is given by:

$$P_{(i,j,k,SP)(l,m,q,APP)} = P\Big(S_{n+1} = APP | S_n = SP, X_{n+1} = l > 0,$$
$$Z_{n+1} = q, Y_{n+1} = m\Big) \cdot \mathbb{1}_{l=i+k} \cdot \mathbb{1}_{j + \lfloor \frac{g(k)}{d_c} \rfloor = m} \cdot \mathbb{1}_{\lfloor \frac{g(k)}{d_p} \rfloor = q}$$

Second, if the next transition is APC with $Y_{n+1} = m > 0$, the transition probability is defined as:

$$P_{(i,j,k,SP)(l,m,q,APC)} = P\Big(S_{n+1} = APC | S_n = SP, Y_{n+1} = m > 0,$$
$$Z_{n+1} = q, X_{n+1} = l\Big) \cdot \mathbb{1}_{l=i+k} \cdot \mathbb{1}_{j + \lfloor \frac{g(k)}{d_c} \rfloor = m} \cdot \mathbb{1}_{\lfloor \frac{g(k)}{d_p} \rfloor = q}$$

Finally, if the next transition is SP with $Z_{n+1} = q > 0$, transition probability is thus formulated as:

$$P_{(i,j,k,SP)(l,m,q,SP)} = P\Big(S_{n+1} = SP | S_n = SP, Z_{n+1} = q > 0,$$
$$X_{n+1} = l, Y_{n+1} = m\Big) \cdot \mathbb{1}_{l=i+k} \cdot \mathbb{1}_{j + \lfloor \frac{g(k)}{d_c} \rfloor = m} \cdot \mathbb{1}_{\lfloor \frac{g(k)}{d_p} \rfloor = q}$$

3.2 Stationary probability

So far, we have evaluated all the transition probabilities. The last phase consists in deriving the stationary probability and computing the frame aggregation levels. Let us remind that a Markov chain is irreducible if and only if every state can be reached by any other state through one or several transitions. As our Markov chain is irreducible and has a finite number of states, hence it exists a unique stationary distribution. We solve this Markov chain through a numerical method. We compute the stationary probabilities and we denote by π the vector containing the corresponding values. The mean DL probe traffic aggregation levels, denoted *MeanAgg* are then given by: $MeanAgg = \sum_{n=1}^{AMPDU_{AP}} n \cdot \pi_n$.

4 PERFORMANCE EVALUATION

In this section, we assess our model's accuracy by comparing its outcomes with those delivered by the discrete-event simulator ns-3 (version 3.30).

4.1 Simulation settings

The simulation setup consisted of the one described in Figure 3 where a probe traffic client associated with an AP transmits User Datagram Protocol (UDP) flow to the probe server. Another server station is associated with the AP to receive a UDP cross traffic. The cross traffic generation rate is varied over a wide range of values covering saturated and unsaturated network conditions. The simulated cross traffic is set in order to have six different BTFs: 0, 0.125, 0.25, 0.375, 0.5, and 0.625. All the nodes use the 802.11n standard and operate on the 2.4 GHz frequency band. We provide three different scenarios to evaluate the performance of our model. The different scenarios share the same parameters but differ in the transmission rate used by each node.

4.2 Evaluation

We start by examining the performance details of our model when the same MCS index (MCS 15 with a physical transmission rate of 144.4 Mbps) is used for the probe traffic client and the AP. The maximum number of aggregated sub-frames $AMPDU_{AP} = AMPDU_P = 36$. Figure 5 shows the mean aggregation level for the probe traffic as a function of the probe packet gap for the model and simulations. The probe packet gap gradually varies from $50\mu s$ to $1000\mu s$.

According to these results, it appears that the model performs well since it follows closely the pattern of the ns-3 simulations for the three levels of the network loads. It is also the case for the three other BTFs (0, 0.125, and 0.625) that are not shown by lack of space. We can observe that the aggregation levels depend on loads of the network. Indeed, when the BTF increases, the probe traffic has to wait longer, and more packets are received between two successive probe transmissions.

Figure 6 shows the model and ns-3 simulations outcomes when we use different MCS for the probe traffic client and the AP. Figure 6a plots the results when we use the MCS15 (144.4 Mbps) for the AP to send the DL probe traffic and the cross traffic and the MCS11 (57.8 Mbps) for the probing client. Figure 6b provides the corresponding results when we use MCS15 for the AP and MCS13 (115.6 Mbps) for the probe traffic client. The results show that considering different MCS for the AP and the probe traffic client does not impact the accuracy of our approach.

Figure 7 illustrates the mean probe aggregation levels as a function of the probe packet gap for all the levels of BTF used in our work for the case where all the nodes (AP and stations) use the same transmission rate (144.4Mbps). These results nicely highlight the fact that varying the level of the network load significantly affects the probe aggregation levels. Consequently, the results are sufficiently separated to be used to infer the load level.
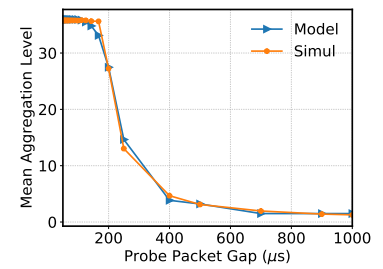
Inspired by the correlation between the aggregation level and the network load, we propose the following method that aims to infer the network load (BTF). For doing so, we send a batch of probe streams with an increasing inter-packet arrival from a client to a server. Upon receiving the probe packets, the server computes
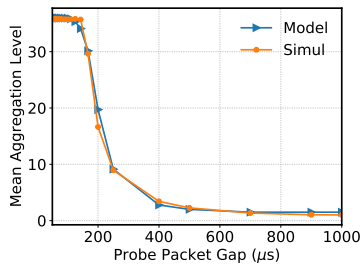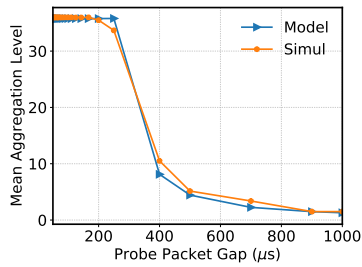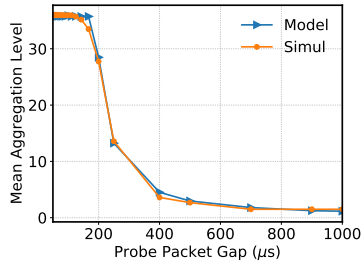
(a) BTF=0.25
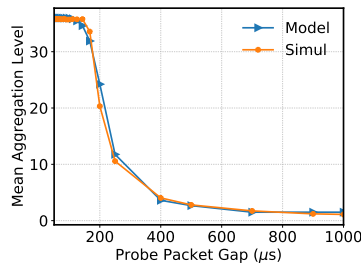
(b) BTF=0.375

(c) BTF=0.5

Figure 5: Mean aggregation level versus Busy Time Fraction

(a) BTF= 0.375, MCS11 for the probe traffic client

(b) BTF= 0.375, MCS13 for the probe traffic client

Figure 6: Mean aggregation level versus Busy Time Fraction

Figure 7: Mean Aggregation level for all BTFs for the simulations.

Table 2: BTF estimations for ns-3 simulations where the MCS15 (144.4 Mbps) is used for all the nodes

Ground Truth BTF	Estimated BTF					
	0	0.125	0.25	0.375	0.5	0.625
0	S/M					
0.125		S/M				
0.25			S/M			
0.375				S/M		
0.5					S/M	
0.625						S/M

Table 3: BTF estimations for ns-3 simulations where different MCS (MCS15 and MCS13) are used for the AP and the probe traffic client

Ground Truth BTF	Estimated BTF					
	0	0.125	0.25	0.375	0.5	0.625
0	S/M					
0.125		S	M			
0.25			S/M			
0.375				S/M		
0.5					S/M	
0.625						S/M

the mean aggregation levels. By relying on the model's results, the server then computes the error between the measured and theoretical (derived by the model) aggregation levels. The estimated BTF is the one that minimizes this error. Table 2 and Table 3 present the estimated BTFs for the ns-3 simulations when we use the same physical transmission rate (MCS15 with 144.4 Mbps) for all the nodes or different transmission rates (MCS15 with 144.4 Mbps for the AP and MCS13 with 115.6 Mbps for the probe traffic client) respectively. The 'ground truth column' gives the real value of the BTF. 'S' indicates the BTF used in simulations and 'M' presents the value returned by our approach. This approach finds the good result when the two letters ('S'/'M') are in the same box.

Results presented in Table 2 show that the method finds the good results for all the levels of BTF. We note that our model takes into account the time to send block ACK requests while computing the frame transmission duration. These requests mean that a response
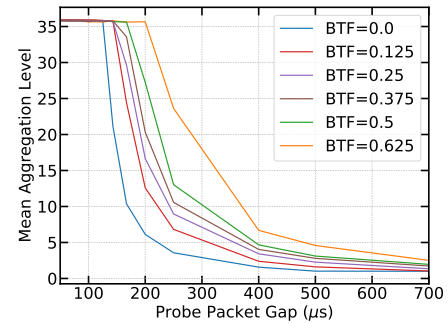
is requested upon transmission of a frame whose sequence number is distant at least by a given threshold multiplied by the transmit window size from the starting sequence number of the transmit window. Table 3 shows that the method finds the good results for the five BTFs 0, 0.25, 0.375, 0.5, and 0.625 and slightly overestimates the BTF for the case 0.125. This error is only of 0.125 (approximately 10%) and is due to the fact that we neglect some protocol aspects, such as the beacons sent by the AP. We obtain similar results as Table 3 for the other case where MCS11 is used for the probe traffic client. However, by lack of space, we do not present its outcomes in this paper.

Overall, the results show that the proposed analytical model accurately captures the mean aggregation level across different network settings and predicts with reasonable accuracy (sometimes with a shift of one BTF level) the BTF level.

5 RELATED WORK

Performance evaluation of Wi-Fi networks has been intensely studied in the literature to offer relevant information and methods to help its configuration and improvement. In this paper, we focus on three categories of work: Available bandwidth estimation, crowd-sensing of wireless networks, and Analytical models.

Available bandwidth estimation. Conversely to achievable throughput estimation tools, there exists another pivotal metric for estimating the overall network performance which is the available bandwidth. Active Available bandwidth measurements that involve the usage of probing packets can be generally divided into two categories: Packet Rate Model (PRM) and Probe Gap Model (PGM). The works based on PRM, such as TOPP [13], pathload [8], pathChirp [16] and DietTOPP [9], focus on comparing the probe packet and received packet rates. They inject increasing probe traffic in the network path and try to find the point where link congestion starts occurring to discern the available bandwidth. These methods differ according to the probing rate settlement.

Conversely to PRM, PGM techniques (e.g., Spruce [22], Initial Gap Increase/Packet Transmission Rate (IGI/PTR) [6], Delphi [15]...) leverage the dispersion in time between two received packets to infer the cross traffic. All these approaches were primitively focused on wired networks and unfortunately failed with the complexity of CSMA/CA of Wi-Fi networks. To copy with this latter, more complex schemes were developed, such as Exact [18], IdleGap [11] and WBest [12]. These techniques intensely depend on the fact that the received packet rate or packet gap should be an effective indicator to reveal the cross traffic. However, the dissimilar packet scheduling system under frame aggregation, proposed by modern Wi-Fi versions, prevents the use of these techniques as inter-packet arrival becomes null when they were embedded in the same aggregated frame. For monitoring the network conditions of these new versions of Wi-Fi, WBest+ [3] and AIWC [21] use the characteristics of frame aggregation to discern the available bandwidth by relying on a threshold-based method: if an inter-arrival time is less than a given threshold, the two packets are considered as aggregated. In [1], the mean aggregation level is compared to the one of an analytical model to estimate the network load. The network scenario used in these papers consists of a server with a wired Internet connection that sends [3, 21] or receives [1] probing traffic along the network path to/from a client with a last-hop wireless connection that assumed to be the bottleneck. The scenario considered here is different as both the probe traffic generator and the receiver are associated with the same Wi-Fi AP. The behavior of the mean aggregation is then much more complex as it may be different on the uplink and downlink links due to their different parameters (e.g., MCS). Moreover, the probe traffic accesses twice the channel that decreases the level of probe traffic for which the network becomes saturated and the size of the aggregated frames. While evaluating the probe aggregation levels of the UL probe traffic, the model presented in [1] assumes that the connection between the AP and the server is ideal. It is therefore modeled as if the server were implemented on the AP. However, the model presented here appraises the aggregation levels of the DL probe traffic by considering that the server is embedded on a second wireless device which is more realistic and practical.

Crowd-sensing. Alternatively, some techniques focus on crowd-sensing applications where mobile devices become active sensors to actively estimate the end-to-end network performance. The authors of [17] proposed MCNet, a tool based on active smartphone measurements to estimate the user-perceived performance in enterprise wireless networks. Another work based on channel scans (spectrum sensing) was also proposed in [19]. It is based on passive measurements while our work investigate the effectiveness of active ones. The authors in [4] proposed a measurement crowd-sensing study in the city of Edinburgh to characterize urban Wi-Fi. It revealed several problems in Wi-Fi deployments in public spaces. However, none of these works involves the network load while crowd-sensing the network performance.

Analytical models. Wi-Fi Frame aggregation has attracted a lot of attention due to the significant efficiency enhancements brought by this feature. Consequently, several frame aggregation analytical models have been studied in recent years. [5, 14] predict throughput while incorporating frame aggregation under saturated conditions. A Discrete-Time Markov Chain (DTMC) that considers the frame aggregation has been proposed to show its performance under various aspects under unsaturated conditions [10]. Unfortunately, these models are not applicable or extensible to our work since they consider dissimilar scenarios and traffic patterns.

6 CONCLUSION

In this paper, we have proposed an analytical model that predicts the DownLink probe aggregation levels concurrent to cross traffic. Based on that, we proposed an approach to estimate the expected network load via the Busy Time Fraction. This information could be then used by a crowd-sourcing application, in which participant devices measure and share the load of their surrounding networks. They would build a collective knowledge so that when a device arrives in some area and wants to connect to a Wi-Fi network, it could choose the less loaded AP. Our model was validated through ns-3 simulations across a variety of scenarios.

Overall, the model estimates with a satisfactory degree of precision the aggregation levels. We showed that this frame aggregation can effectively capture the AP Busy Time Fraction with at most 10% of errors. Our future efforts include addressing some of the open

challenges in network load estimation by an unmodified mobile device.

REFERENCES

[1] N. E. H. Bouzouita, A. Busson, and H. Rivano. 2020. Analytical study of frame aggregation level to infer IEEE 802.11 network load. In *International Wireless Communications and Mobile Computing (IWCMC)*. 952–957. https://doi.org/10.1109/IWCMC48107.2020.9148448

[2] Cisco. 2020. Annual Internet Report (2018–2023) White Paper.

[3] A. Farshad, M. Lee, M. K. Marina, and F. Garcia. 2014. On the impact of 802.11n frame aggregation on end-to-end available bandwidth estimation. In *IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 108–116. https://doi.org/10.1109/SAHCN.2014.6990333

[4] A. Farshad, M. K. Marina, and F. Garcia. 2014. Urban WiFi characterization via mobile crowdsensing. In *IEEE Network Operations and Management Symposium (NOMS)*. 1–9. https://doi.org/10.1109/NOMS.2014.6838233

[5] N. Hajlaoui, I. Jabri, and M. Ben Jemaa. 2018. An accurate two dimensional Markov chain model for IEEE 802.11n DCF. *Wireless Networks* 24 (2018), 1019–1031. https://doi.org/10.1007/s11276-016-1383-z

[6] N. Hu and P. Steenkiste. 2003. Evaluation and characterization of available bandwidth probing techniques. *IEEE Journal on Selected Areas in Communications* 21, 6 (2003), 879–894.

[7] IEEE. 2020. Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)* (2020), 1–4379. https://doi.org/10.1109/IEEESTD.2021.9363693

[8] M. Jain and C. Dovrolis. 2002. Pathload: A Measurement Tool for End-to-End Available Bandwidth. In *Passive and Active Measurement Workshop (PAM)*. 14–25.

[9] A. Johnsson, M. Bjorkman, and B. Melander. 2006. An Analysis of Active End-to-end Bandwidth Measurements in Wireless Networks. In *IEEE/IFIP Workshop on End-to-End Monitoring Techniques and Services*. 74–81. https://doi.org/10.1109/E2EMON.2006.1651282

[10] B. S. Kim, H. Y. Hwang, and D. K. Sung. 2008. Effect of Frame Aggregation on the Throughput Performance of IEEE 802.11n. In *IEEE Wireless Communications and Networking Conference (WCNC)*. 1740–1744. https://doi.org/10.1109/WCNC.2008.310

[11] H. K. Lee, V. Hall, K. H. Yum, K. I. Kim, and E. J. Kim. 2006. Bandwidth Estimation in Wireless Lans for Multimedia Streaming Services. In *IEEE International Conference on Multimedia and Expo (ICME)*. 1181–1184. https://doi.org/10.1109/ICME.2006.262747

[12] M. Li, M. Claypool, and R. Kinicki. 2008. WBest: A bandwidth estimation tool for IEEE 802.11 wireless networks. In *2008 33rd IEEE Conference on Local Computer Networks (LCN)*. 374–381. https://doi.org/10.1109/LCN.2008.4664193

[13] B. Melander, M. Bjorkman, and P. Gunningberg. 2000. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *IEEE Global Telecommunications Conference (Globecom)*, Vol. 1. 415–420. https://doi.org/10.1109/GLOCOM.2000.892039

[14] N. Mohammad and S. Muhammad. 2012. Modeling and Analyzing MAC Frame Aggregation Techniques in 802.11n Using Bi-dimensional Markovian Model, Vol. 293. 408–419. https://doi.org/10.1007/978-3-642-30507-8_35

[15] V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks, and R. Baraniuk. 2000. Multifractal cross-traffic estimation. In *ITC Specialist Seminar on IP Traffic Measurement, Modeling, and Management*. 15–1.

[16] V. Ribeiro, R. Riedi, J. Navrátil, and L. Cottrell. 2003. PathChirp: Efficient Available Bandwidth Estimation for Network Paths. In *Passive and Active Measurement Workshop (PAM)*. 1–11. https://doi.org/10.2172/813038

[17] S. Rosen, S. Lee, J. Lee, P. Congdon, Z. M. Mao, and K. Burden. 2014. MCNet: Crowdsourcing wireless performance measurements through the eyes of mobile devices. *IEEE Communications Magazine* 52, 10 (2014), 86–91. https://doi.org/10.1109/MCOM.2014.6917407

[18] S. Shah, K. Chen, and K. Nahrstedt. 2003. Available bandwidth estimation in IEEE 802.11-based wireless networks. *First ISMA/CAIDA Workshop on Bandwidth Estimation (BEst)* (01 2003).

[19] J. Shi, L. Meng, A. Striegel, C Qiao, D. Koutsonikolas, and G. Challen. 2016. A walk on the client side: Monitoring enterprise Wifi networks using smartphone channel scans. In *IEEE International Conference on Computer Communications (IEEE INFOCOM)*. 1–9. https://doi.org/10.1109/INFOCOM.2016.7524453

[20] The Network Simulator. 2021. ns-3. https://www.nsnam.org/

[21] L. Song and A. Striegel. 2017. Leveraging Frame Aggregation for Estimating WiFi Available Bandwidth. In *IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 1–9. https://doi.org/10.1109/SAHCN.2017.7964908

[22] J. Strauss, D. Katabi, and F. Kaashoek. 2003. A Measurement Study of Available Bandwidth Estimation Tools. In *ACM SIGCOMM Conference on Internet Measurement (IMC)*. 39–44. https://doi.org/10.1145/948205.948211

A FRAME TRANSMISSION DURATION

In this paper, the frame transmission duration takes into account the data packet transmission itself and all protocol overheads (backoff, SIFS, DIFS and ACK or block ACK transmission). The transmission duration of l sub-frames is expressed as :

$$f(l) = T_{DIFS} + T_{backoff} + T_{PHY} + T_{SIFS} + T_{Block\ ACK} +$$
$$\frac{1}{freq} \times T_{Block\ ACK\ Request}$$
$$+ \frac{(MPDU\ Delimiter + MacHeader + Payload + FCS) \times 8 \times l}{Physical\ transmission\ rate}$$

Where freq denotes the frequency of sending of block ACK requests. $T_{backoff}$ denotes the mean duration of the backoff period for frames without retransmissions. It is given by:

$$T_{backoff} = \frac{CW_{min}}{2} \times T_{slot}$$

$T_{BlockACK}$ and T_{ACK} denote respectively the time required to send the block acknowledgment frame and the acknowledgment frame. They also count their physical header.

Note that the only difference between the functions $f(.)$, $h(.)$ and $g(.)$ is the physical transmission rate and the packet size that depend on the local properties of the nodes.

B ADDITIONAL PROBABILITIES

$$\mathbb{P}\Big(S_{n+1} = APC | S_n = APP, Z_{n+1} = q, Y_{n+1} = m > 0\Big) = \frac{1}{2}\mathbb{1}_{q>0} + \mathbb{1}_{q=0}$$

$$\mathbb{P}\Big(S_{n+1} = SP | S_n = APP, Y_{n+1} = m, Z_{n+1} = q > 0\Big) = \frac{1}{2}\mathbb{1}_{m>0} + \mathbb{1}_{m=0}$$

$$\mathbb{P}\Big(S_{n+1} = APP | S_n = APC, X_{n+1} = X_n = l > 0, Z_{n+1} = q\Big) = \frac{1}{2}\mathbb{1}_{q>0} + \mathbb{1}_{q=0}$$

$$\mathbb{P}\Big(S_{n+1} = APC | S_n = APC, X_n = X_{n+1} = 0, Y_{n+1} = m > 0, Zn + 1 = q\Big) =$$
$$\frac{1}{2}\mathbb{1}_{q>0} + \mathbb{1}_{q=0}$$

$$\mathbb{P}\Big(S_{n+1} = SP | S_n = APC, Z_{n+1} = q > 0, X_{n+1} = X_n = l, Y_{n+1} = m\Big) =$$
$$\frac{1}{2}\mathbb{1}_{l+m>0} + \mathbb{1}_{l+m=0}$$

$$\mathbb{P}\Big(S_{n+1} = APP | S_n = SP, X_{n+1} = l > 0, Z_{n+1} = q, Y_{n+1} = l\Big) = \frac{1}{2} \cdot \frac{1}{2}\mathbb{1}_{q>0,m>0} +$$
$$\frac{1}{2}\mathbb{1}_{q>0,m=0} + \frac{1}{2}\mathbb{1}_{q=0,m>0} + \mathbb{1}_{q=0,m=0}$$

$$\mathbb{P}\Big(S_{n+1} = APC | S_n = SP, Y_{n+1} = m > 0, Z_{n+1} = q, X_{n+1} = l\Big) = \frac{1}{2} \cdot \frac{1}{2}\mathbb{1}_{q>0,l>0} +$$
$$\frac{1}{2}\mathbb{1}_{q>0,l=0} + \frac{1}{2}\mathbb{1}_{q=0,l>0} + \mathbb{1}_{q=0,l=0}$$

$$\mathbb{P}\Big(S_{n+1} = SP | S_n = SP, Z_{n+1} = q > 0, X_{n+1} = l, Y_{n+1} = m\Big) = \frac{1}{2}\mathbb{1}_{m+l>0} + \mathbb{1}_{m+l=0}$$