PERFORMANCE ANALYSIS OF WIRELESS LANS: AN INTEGRATED PACKET/FLOW LEVEL APPROACH

Remco Litjens*[1], Frank Roijers*,
Hans van den Berg*,†, Richard J. Boucherie†,‡ and Maria Fleuren•

* Knowledge Innovation Center, TNO Telecom, The Netherlands
† Department of Applied Mathematics, University of Twente, The Netherlands
‡ Centrum voor Wiskunde en Informatica, The Netherlands
• Business Innovation Center, TNO Telecom, The Netherlands

Abstract

In this paper we present an integrated packet/flow level modelling approach for analysing flow throughputs and transfer times in IEEE 802.11 WLANs. The packet level model captures the statistical characteristics of the transmission of individual packets at the MAC layer, while the flow level model takes into account the system dynamics due to the initiation and completion of data flow transfers. The latter model is a processor sharing type of queueing model reflecting the IEEE 802.11 MAC design principle of distributing the transmission capacity fairly among the active flows. The resulting integrated packet/flow level model is analytically tractable and yields a simple approximation for the throughput and flow transfer time. Extensive simulations show that the approximation is very accurate for a wide range of parameter settings. In addition, the simulation study confirms the attractive property following from our approximation that the expected flow transfer delay is insensitive to the flow size distribution (apart from its mean).

Keywords: Wireless LANs, IEEE 802.11, performance analysis, quality of service, processor sharing models.

AMS subject classificiations: primary: 90B18, 90B22; secondary 60K25.

1 INTRODUCTION

Wireless Local Area Networks (WLANs) are expected to play an important role in future everyday's communication, not only in the private domain but also for public use. In particular, they may fulfill the need for an additional public wireless access solution for data services in hot spots (e.g. train stations, airports, etc.), besides the access provided by mobile cellular networks such as GSM/GPRS and UMTS [10]. WLANs provide an interesting possibility to offer additional capacity and higher bandwidths to end-users without sacrificing the inherently scarce and expensive capacity of cellular networks. However, critical factors for successful introduction are security and performance, which applies in particular to deployment of WLANs in the public environment.

WLAN performance is largely determined by the maximum data rate at the physical layer and the MAC layer protocols defined by the IEEE 802.11 standards [13, 17]. The most widely employed WLAN MAC protocol is the Distributed Coordination Function (DCF). The DCF is a random access scheme based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), which uses random backoffs in order to manage packet retransmissions in case of a collision. If the DCF is used in its BASIC access mode, the total WLAN throughput decreases drastically when the number of active stations becomes larger, due to a rapidly increasing number of collisions. To overcome this throughput degradation the Request-To-Send/Clear-To-Send (RTS/CTS) mechanism has been proposed and standardised. Under RTS/CTS a station sends a small control packet in order to reserve the channel for transmission of a data packet. In particular, this solves the so-called hidden station problem,

[1]Corresponding author: TNO Telecom, P.O.Box 421, 2260 AK Leidschendam, The Netherlands, e-mail: R.Litjens@telecom.tno.nl, phone: +31 70 446 3419, fax: +31 70 446 3477.

which may lead to a reduced system throughput due to excessive collisions. The IEEE 802.11 standard also defines an optional, centralised MAC protocol called Point Coordination Function (PCF). In the PCF-scheme a central node polls the stations to access the shared medium, thus eliminating the need for contention and enabling the support of delay-sensitive services. In the rest of this paper we focus on the DCF MAC protocol.

The IEEE 802.11 WLAN can operate in infrastructure mode or in ad-hoc mode, see Figure 1. In ad-hoc mode all stations can transmit packets directly to other stations that are within the sending range (Basic Service Set (BSS)). In the infrastructure mode an Access Point (AP) is present to link the stations in a BSS. An AP may be connected to a distribution system (e.g. a wired LAN) via which stations linked to other AP's or e.g. a remote server can be reached.

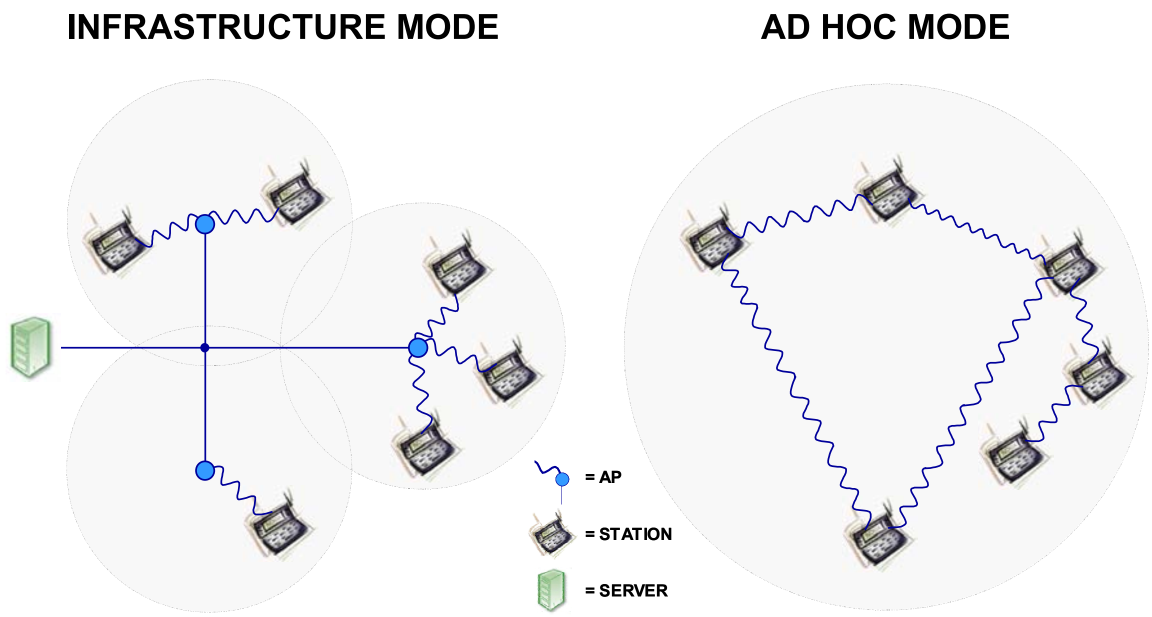

Figure 1: Infrastructure and ad hoc mode WLANs.

A number of papers have studied the throughput performance of IEEE 802.11 DCF MAC for both the BASIC and RTC/CTS access modes. Several of them are based on simulation, see e.g. [21]. Other papers use analytical models, but with simplifying assumptions about the DCF MAC layer operations and/or the traffic conditions in order to enable mathematical analysis. In particular, a strongly simplified backoff mechanism has been used in [4, 5], while [9, 23] assume Poisson sources generating fixed size data packets. A more detailed mathematical performance model of DCF has been developed and analysed by Bianchi [2] that was slightly extended by Wu et al. [22]. The key approximation made in these papers enabling a relatively simple Markov chain analysis is the assumption of independent transmissions by different flows, as well as constant and independent collision probabilities, regardless of the number of erroneous transmissions already experienced. Comparison with simulation shows that the analytical results are mostly very accurate. [2, 22] both assume a constant number of persistently active sources and a simplified physical layer model. [11, 12] consider a more realistic model of the physical layer. In particular, they study the impact of packet capture on the aggregate system throughput by two different capture models, which specify the likelihood that transferred data packets survive a collision with concurrently transferred packets.

CONTRIBUTION In the present paper we extend the work mentioned above in two directions. First, we further elaborate on Bianchi's packet level model [2] by integrating the various modelling enhancements on the physical and MAC layer proposed by other authors (see above) into one single DCF performance model, still allowing analytical treatment. Our second extension covers the practical situation that the number of active stations is not constant (as in e.g. [2, 22]) but varies in time due to the random user behaviour, i.e. the initiation and completion of flow transfers. In order to enable mathematical analysis of flow throughputs and transfer delays in the system with the extensions mentioned above we propose an integrated packet/flow level modelling approach. In particular, from the flow level point of view, the WLAN is considered as a queueing system with Poisson flow arrivals and a Processor Sharing (PS) type of service discipline, which reflects the IEEE 802.11 DCF MAC design principle of distributing the transmission capacity fairly among the active flows. The rate at which the flows are served depends on the number of flows simultaneously present in the system (i.e. the number of active stations). These service rates are obtained from the analysis of our extended packet level model describing the behaviour of DCF in detail for the situation with a constant number of active stations. The resulting PS queueing model with state-dependent service rates is analytically tractable (see e.g. [6]) and yields closed-form expressions for e.g. the (conditional) expected flow transfer time in the WLAN system. Our modelling approach provides also some important, more general insights in the essential WLAN performance characteristics. In particular, the well-known insensitivity property of the Processor Sharing model implies that the expected flow transfer times are independent of the flow size distribution, apart from its mean. In addition, the conditional expected flow transfer time is linear in the flow size. These attractive properties and the accuracy of our analytical performance results are excellently validated by simulation. To our knowledge there is only one other paper, [8], which does consider flow transfer times in a WLAN with non-persistent bursty sources. Using the DCF performance model and results of [2], the authors construct a continuous time Markov chain describing the system dynamics when the number of active stations varies in time. The steady state distribution of this Markov chain is numerically solved from the balance equations and yields approximations for the mean throughput and flow transfer delay. Summarising, the main contributions of the present paper are the inclusion of an enhanced DCF and physical layer model that remains analytically solvable, and the recognition that the resulting flow level model is an analytically tractable PS queue, which opens the possibility for additional performance analysis of e.g. Call Admission Control.

OUTLINE This paper is organised as follows. In Section 2, the IEEE 802.11 DCF MAC protocol is described in more detail. Section 3 describes the system, traffic and capture models underlying the analytical performance study, which is presented in Section 4. In Section 5 we present an extensive numerical study in order to validate the accuracy of our analytical model (by comparison with simulation) and to illustrate the impact of various model parameters on the system performance. Finally, the principal conclusions of our investigation as well as some topics for further research are outlined in Section 6.

2 DISTRIBUTED COORDINATION FUNCTION

The Distributed Coordination Function (DCF) [13] is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme. Whenever a station wants to transmit a packet, it first senses the channel to determine whether or not it is already in use by another station. If the channel is idle, and remains idle for a contiguous period of time called DIFS (Distributed InterFrame Space), the station can transmit the packet. Otherwise the station waits until the channel becomes idle for a period DIFS, after which it has to wait a random number of time slots before it is permitted to send a packet. This random back-off procedure is intended to reduce the probability of multiple stations sending at the same time resulting in a *collision*.

The back-off procedure draws a random value for a back-off counter from a discrete uniform distribution between 0 and $cw_r - 1$, where cw_r is the so-called contention window at the r-th re-attempt to send the packet. As long as the channel remains idle after a DIFS period, a station will decrement its back-off counter by 1 for each time slot. When the back-off counter of a particular station reaches 0, the station transmits a packet. If the packet is received correctly, the destination responds by sending an acknowledgment (ACK) to the source.

In case multiple packets are transmitted concurrently, the packet with the strongest received signal may be *captured* by the intended destination, as long as the carrier-to-interference ratio exceeds a minimum threshold. If a station does not receive an ACK, it assumes that the packet was lost and it will retransmit the packet. The contention window cw_r is doubled and a new random value is chosen from the interval $[0, cw_r - 1]$. cw_r is given by expression (1), where $r_{\max}$ is the maximum number of retransmissions for one packet and r^* is the maximum number of times that cw_r may be doubled after a failed transmission attempt.

$$cw_r = \begin{cases} 2^r\left(cw_{\min}+1\right), & 0 \le r \le r^\star, \\ 2^{r^\star}\left(cw_{\min}+1\right), & r^\star \le r \le r_{\max}. \end{cases} \quad (1)$$

The DCF operates in two different access modes, BASIC access and RTS/CTS-access. Figure 2 illustrates the principle of the BASIC access scheme. The source station sends a data packet to the destination station, which responds by sending an ACK after a time period of length SIFS (Short InterFrame Space). This period is needed by a station to switch from receiving mode to sending mode. As a SIFS is shorter than a DIFS, the ACK will be transmitted before other stations are allowed to send their packets. If the source does not receive the ACK within a pre-defined time out period, it will resend the packet. After the packet is successfully transmitted, the cw_r is reset to cw_0 and the whole procedure repeats for subsequent packets.

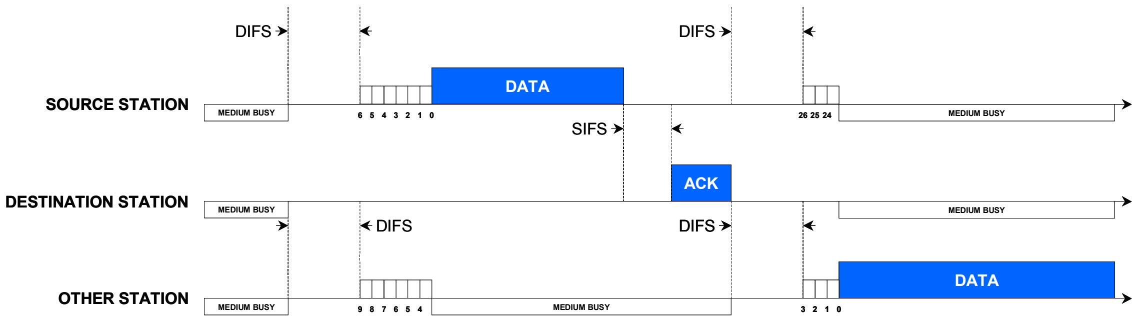

Figure 2: BASIC access mode in the distributed coordination function.

The operation of the RTS/CTS mode is illustrated in Figure 3. In this mode a source station first sends a small RTS frame (Request To Send) when it is ready to send a data packet. If the destination station is able to receive a packet from the source, it responds with a CTS frame (Clear To Send). After receipt of the CTS the source transmits the data packet which is subsequently acknowledged by the destination. All of these frames are separated by a time period of length SIFS so that the other stations cannot intervene the transmissions.

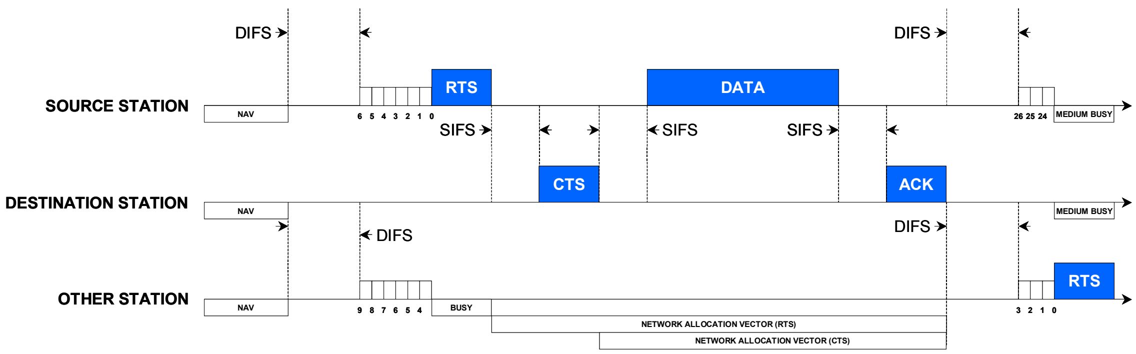

Figure 3: RTS/CTS access mode in the distributed coordination function.

The advantages of the RTS/CTS access scheme are twofold. First it is an efficient way to reduce the impact of a collision that is detected when the sender does not receive a CTS frame. The RTS frame is typically smaller than a data packet and therefore the time wasted by a collision will be shorter. A second benefit of the RTS/CTS scheme is that it reduces the so-called *problem of hidden stations*, i.e. stations which cannot hear each others packet transfers. A hidden station may be able to hear the CTS frame, which contains a *duration field* that indicates the total transmission time up to the ACK. This information is used to set the station's NAV (Network Allocation Vector) so that it is aware of the medium being used, even if it cannot sense the transmitting station directly *(virtual carrier sensing)*.

Although RTS/CTS access mode reduces the effect of collisions, it involves more overhead than the BASIC access mode. In particular, for small packets and a small number of users (when the probability of a collision is small), BASIC access mode is usually more efficient.

3 MODEL

This section sets the framework for the presented performance analysis by describing the system, capture and traffic models in generic terms. Concrete parameter settings are specified in Table 1 in Section 5.

3.1 SYSTEM MODEL

We consider a single basic service set with stations contending for the shared WLAN radio access medium. The fixed channel rate of the medium is denoted $r_{\text{WLAN}} \in \{1, 2, 5.5, 11\} \cdot 10^3$ kbits/s, while the physical layer preamble (required for synchronisation purposes) and header are always transmitted at a fixed rate of 1 Mbits/s to ensure compatibility between the IEEE 802.11 and IEEE 802.11B standards. The applied DCF is considered in both BASIC and RTS/CTS access mode. The DCF operations at the MAC layer are modelled in significant detail in the Markov chain that is taken from [2, 22] and specified in Section 4. The DCF model includes DIFS and SIFS timers, MAC layer acknowledgements, an exponentially increasing contention window, and a randomly sampled backoff counter that is decremented towards a packet transfer attempt and potentially 'frozen' if the shared medium is sensed busy. Furthermore, we integrate a more realistic physical layer model into the setting of [2, 22] by taking into account the possibility of capture, in case of concurrent packet transfers. The considered capture models are discussed below after a specification of the traffic model.

3.2 TRAFFIC MODEL

The considered WLAN network serves stations which generate data flows according to a Poisson process with rate λ, and are assumed to be located at similar distances from their intended receiver(s). Data flows are assumed to be transfers of files with generally distributed sizes. The mean file size is denoted $1/\mu$ (in kbits). Each file is segmented into packets of a given size (with a final packet containing the flow's remainder) which are processed at the WLAN's MAC layer. The data traffic load is denoted $\rho \equiv \lambda / (\mu r_{\text{WLAN}})$. A CAC scheme is deployed to limit the number of contending data flows to $n_{\max}$ and thus ensure system stability and provide some minimum Quality of Service (QOS).

3.3 CAPTURE MODELS

We apply two distinct capture models, denoted CM-1 and CM-2, which specify the likelihood that transferred data packets survive a collision with concurrently transferred packets [11, 12, 14, 20]. The common assumption underlying both capture models is that in a collision of multiple packets, only the one with the strongest signal has a chance of successful capture. In our analytical performance evaluation model, the effects of capture appear in the form of the capture functions $\mathbf{P}_s(k)$ and $\mathbf{P}_s^\star(k)$ for $k \geq 1$. The former function is defined as the probability that the *strongest* data packet among k concurrently transferred packets is successfully captured, while the latter function denotes the probability that a *tagged* data packet is successfully captured in a simultaneous transfer with $k-1$ *other* data packets. Both capture models specified below have the desired property that both $\mathbf{P}_s(k)$ and $\mathbf{P}_s^\star(k)$ are nonincreasing in k.

Capture model CM-1 assumes that a packet transfer is successful if and only if there are no concurrent transfers from other flows. Expressed in the above-defined capture functions:

$$\mathbf{P}_s(k) = \mathbf{P}_s^\star(k) = \begin{cases} 0 & \text{if } k > 1, \\ 1 & \text{if } k = 1. \end{cases}$$

Capture model CM-1 is the most basic option imaginable and is implicitly applied in e.g. [2, 22].

Capture model CM-2 is taken from [11, 12] and assumes that all signals in a collision have some uniform local mean received power $\overline{p}$, determined by attenuation and shadowing effects, while the instantaneous received signal powers are independent and exponentially distributed around this mean, which is a direct consequence of an assumption of Rayleigh fading (see e.g. [15]). Under these assumptions the capture function $\mathbf{P}_s(k)$ is equal to the probability that the carrier-to-interference ratio of the strongest signal in a collision ensemble of k signals exceeds the threshold $\Gamma^\star$ required for successful capture, i.e.

$$\mathbf{P}_s(k) = \int_{p_1=0}^{\infty}\int_{p_2=0}^{\infty}\cdots\int_{p_k=0}^{\infty}\left(\prod_{i=1}^{k}\varphi_{\overline{p}}(p_i)\right) 1\left\{\Gamma_{\max}(p_1,\cdots,p_k) \geq \Gamma^\star\right\} dp_1 dp_2\cdots dp_k, \tag{2}$$

where the p_i's, $i = 1,\cdots,k$, denote the instantaneous received signal powers of all signals involved in the collision, $\varphi_{\overline{p}}$ is the exponential PDF with mean $\overline{p}$, $1\{\cdot\}$ is the indicator function, and $\Gamma_{\max}(p_1,\cdots,p_k)$ is defined as the carrier-to-interference ratio of the strongest signal in the collision ensemble. The independence of $\mathbf{P}_s(k)$ with respect to the local mean received power $\overline{p}$ follows from a substitution of $q_i \equiv p_i/\overline{p}$ in the above integral. The capture function is readily evaluated analytically, observing that

$$\{(p_1,\cdots,p_k) : \Gamma_{\max}(p_1,\cdots,p_k) \geq \Gamma^\star\} = \bigcup_{i=1}^{k}\left\{(p_1,\cdots,p_k) : (1+\Gamma^\star)\,p_i \geq \Gamma^\star\sum_{i=1}^{k}p_i\right\}.$$

For $\Gamma^\star \geq 1$ the sets $\left\{(p_1,\cdots,p_k) : (1+\Gamma^\star)\,p_i \geq \Gamma^\star\sum_{i=1}^{k}p_i\right\}$ are disjoint, so that

$$\begin{aligned}
\mathbf{P}_s(k) &= k\int_{p_1=0}^{\infty}\int_{p_2=0}^{\infty}\cdots\int_{p_k=\Gamma^\star(p_1+\cdots+p_{k-1})}^{\infty}\left(\frac{1}{\overline{p}}\right)^{k}\exp\left(-\sum_{i=1}^{k}\frac{p_i}{\overline{p}}\right)dp_1dp_2\cdots dp_k \\
&= k\int_{p_1=0}^{\infty}\int_{p_2=0}^{\infty}\cdots\int_{p_{k-1}=0}^{\infty}\left(\frac{1}{\overline{p}}\right)^{k-1}\exp\left(-\frac{(1+\Gamma^\star)}{\overline{p}}\sum_{i=1}^{k-1}p_i\right)dp_1dp_2\cdots dp_{k-1} \\
&= \frac{k}{(1+\Gamma^\star)^{k-1}}\left[\int_{p_1=0}^{\infty}\int_{p_2=0}^{\infty}\cdots\int_{p_{k-1}=0}^{\infty}\left(\frac{1+\Gamma^\star}{\overline{p}}\right)^{k-1}\exp\left(-\frac{(1+\Gamma^\star)}{\overline{p}}\sum_{i=1}^{k-1}p_i\right)dp_1dp_2\cdots dp_{k-1}\right] \\
&= \frac{k}{(1+\Gamma^\star)^{k-1}},
\end{aligned}$$

where the expression between the brackets is simply the integrated probability mass of a $(k-1)$-dimensional joint exponential distribution with uniform parameter $(1+\Gamma^\star)/\overline{p}$. Observe that the resulting expression does indeed not depend on $\bar{p}$. In the alternate case that $\Gamma^\star \leq 1$ the sets $\left\{(p_1,\cdots,p_k) : (1+\Gamma^\star)\,p_i \geq \Gamma^\star\sum_{i=1}^{k}p_i\right\}$ are not disjoint, so that, although it is still straightforward, evaluation of $\mathbf{P}_s(k)$ requires substantial bookkeeping. Still, a relatively 'nice' expression can be derived for the case of $\Gamma^\star \in [0.5, 1]$:

$$\mathbf{P}_s(k) = \frac{k}{(1+\Gamma^\star)^{k-1}} - \frac{1}{2}k(k-1)\left(\frac{1-\Gamma^\star}{1+\Gamma^\star}\right)^{k-1}.$$

Since all signal powers are assumed to be independent and identically distributed, the probability that a tagged signal is the strongest one in a collision ensemble of k signals is equal to $1/k$, so that capture function $\mathbf{P}_s^\star(k)$ for CM-2 is given by

$$\mathbf{P}_s^\star(k) = \frac{1}{k}\mathbf{P}_s(k).$$

In case $\Gamma^\star \geq 1$, the resulting expression for $\mathbf{P}_s^\star(k)$ reflects a form of independence in the sense that $\mathbf{P}_s^\star(k) = (\mathbf{P}_s^\star(2))^{k-1}$, i.e. the probability that a tagged signal is sufficiently stronger than the sum of $k-1$ interfering signals is equal to the probability that the tagged signal is sufficiently stronger in each of $k-1$ pairwise comparisons with the individual interfering signals.

The carrier-to-interference ratio requirement is given by $\Gamma^\star \equiv z_0 g(S_f)$. Here z_0 denotes the required energy-per-bit to interference density ratio, which typically lies somewhere in the range $6-24$ dB. Assuming rectangular chip pulses at the receiver, the inverse processing gain $g(S_f) \equiv 2/(3S_f)$ is a function of the spreading factor S_f. S_f is equal to 11 for $r_{\text{WLAN}} \in \{1, 2\} \cdot 10^3$ kbits/s (IEEE 802.11: Binary/Quadrature Phase Shift Keying), resulting in $\Gamma^\star \in [0.2413, 15.2236]$. A spreading factor of $S_f = 8$ is used for $r_{\text{WLAN}} \in \{5.5, 11\} \cdot 10^3$ kbits/s (IEEE 802.11B: Complementary Code Keying), which leads to $\Gamma^\star \in [0.3318, 20.9324]$.

4 PERFORMANCE ANALYSIS

The analytical evaluation of the WLAN performance is split into two stages. STAGE I concentrates on the packet level dynamics at the MAC layer, generalising the analysis first presented in [2] (and subsequently improved by [22]), by incorporating the possibility of capture at the physical layer. The outcome of STAGE I is the aggregate system throughput as a function of the number of persistently active data flows. STAGE II then focuses on the flow level performance using a generalised processor sharing queueing model, and includes the impact of the dynamics of flow arrivals and departures. At this level, for the analysis of e.g. the flow transfer delay we utilise the throughput function as provided by STAGE I.

4.1 STAGE I: THROUGHPUT ANALYSIS FOR PERSISTENT FLOWS

The STAGE I analysis builds upon the approach presented in [2, 22] and generalises the considered model with the incorporation of packet capture in case of concurrent transfers. In a scenario with n persistent data flows, the MAC layer operations of a single tagged data flow are modelled by a Markov chain, while the impact of the other $n-1$ flows is incorporated by means of the packet error probability $\mathbf{P}_e$. In turn, from the equilibrium distribution, which is expressed in terms of the Markov chain's characterising parameter $\mathbf{P}_e$, an expression for the packet transfer probability $\mathbf{P}_t^\star$ of an individual flow can be derived, requiring the numerical determination of a unique fixed point. Subsequently, the equilibrium distribution is utilised to derive a closed-form expression for the expected data throughput.

A key approximation that is made in the analysis is the independence of the different flows' transfer events, which implies that the packet error probability is independent of the number of transfer reattempts the tagged data flow required thus far (cf. [2]). In practice, when a tagged flow's data packet collides irrepairably, not only the tagged flow's contention window size is doubled, but typically also that of the interfering data flow, which in turn decreases the probability that the next packet transfer attempt fails as well.

The evaluation approach is worked out in more detail below.

MARKOV CHAIN ANALYSIS

Consider a single tagged data flow contending for the WLAN's shared medium with $n-1$ other flows. Denote with $b(t)$ the stochastic process representing the tagged flow's backoff counter, and with $r(t)$ the stochastic process counting the number of transfer reattempts for the packet at the head of the tagged flow's queue that currently awaits a successful transfer. The embedded jump chain following the state transition of the two-dimensional semi-Markov process $(r(t), b(t))_{t\geq 0}$ is modelled by an irreducible discrete-time Markov chain $(r(k), b(k))_{k\in\mathbb{N}_0}$, with states denoted (r, b), see Figure 4. Observe that in each state at the left of the diagram, i.e. with $b = 0$, the tagged flow (re)attempts a packet transfer, while in all states to the right of such a 'transfer state', the tagged flow is decrementing its backoff counter. The contention window size cw_r, as specified in (1), sets the upper bound for the randomly sampled initial backoff value. The state space $\mathbb{S}$ of the Markov chain is

$$\mathbb{S} \equiv \{(r, b) \in \mathbb{N}_0 \times \mathbb{N}_0 : 0 \leq b \leq cw_r - 1 \text{ and } 0 \leq r \leq r_{\max}\}.$$

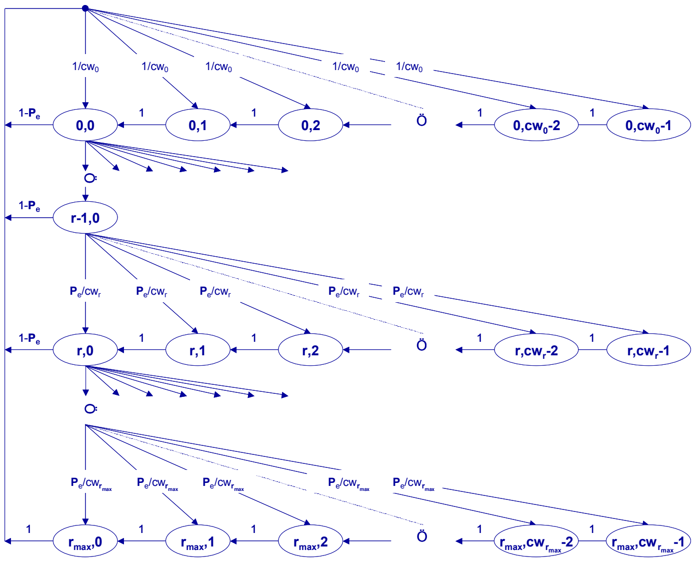

Figure 4: The embedded Markov chain model.

The influence of the other $n-1$ data flows sharing the wireless medium is incorporated in the Markov chain by means of the packet error probability $\mathbf{P}_e$, i.e. the probability that a packet transfer collides irrecoverably with one or more other simultaneous packet transfers. It is stressed that a temporary freeze of the backoff counter due to the sensed activity of another data flow, affects only the *time* between subsequent decrements of $b(t)$, *not* the evolution of the embedded jump chain considered here. This effect is included in the throughput analysis below. The Markov chain's one-step transition probabilities corresponding to a succesful $((1-\mathbf{P}_e)/cw_0)$ or erroneous $(\mathbf{P}_e/cw_r)$ packet transfer, a backoff counter decrement (with probability 1), and a reset of the retransmission counter $(1/cw_0)$ after the $r_{\max}^{\text{th}}$ MAC layer packet transfer reattempt (regardless of whether the transfer is succesful or not), are indicated in Figure 4. The last-mentioned event type is where [2] and [22] differ: while in [2] the considered station continues to attempt the packet transfer until it is successful, in [22] the station gives up after $r_{\max}$ reattempts, as is the case in our model. For the MAC layer throughput analysis of persistent data flows presented in this section, it is irrelevant whether a packet that suffers from $r_{\max}+1$ unsuccessful transfer attempts is discarded or scheduled for retransmission by higher-layer protocols (e.g. UDP or TCP).

Since the discrete-time Markov chain is irreducible and has a finite state space, a unique equilibrium distribution $(\pi(r, b), (r, b) \in \mathbb{S})$ exists, and is given by (cf. [22])

$$\pi(r, b) = \pi(0, 0) \cdot \begin{cases} \mathbf{P}_e^r & \text{for } 0 \leq r \leq r_{\max} \quad \text{and } b = 0, \\ \\ \frac{cw_r - b}{cw_r}\mathbf{P}_e^r & \text{for } 0 \leq r \leq r_{\max} \quad \text{and } 1 \leq b \leq cw_r - 1, \end{cases} \tag{3}$$

while the normalisation condition for the equilibrium distribution is imposed to determine $\pi(0, 0)$:

$$1 = \sum_{r=0}^{r_{\max}} \sum_{b=0}^{cw_r - 1} \pi(r, b) \Longleftrightarrow \pi(0, 0) = \frac{2(1-\mathbf{P}_e)}{\left(1-\mathbf{P}_e^{r_{\max}+1}\right) + (1-\mathbf{P}_e)\sum_{r=0}^{r_{\max}} cw_r \mathbf{P}_e^r}. \tag{4}$$

Given the cw_r as specified in (1), expression (4) can be written more explicitly using

$$\sum_{r=0}^{r_{\max}} cw_r \mathbf{P}_e^r = \begin{cases} (cw_{\min}+1)\left(\frac{1-(2\mathbf{P}_e)^{r_{\max}+1}}{1-2\mathbf{P}_e}\right) & \text{if } r_{\max} \le r^\star, \\ (cw_{\min}+1)\left\{\frac{1-(2\mathbf{P}_e)^{r^\star+1}}{1-2\mathbf{P}_e} + 2^{r^\star}\frac{\mathbf{P}_e^{r^\star+1}-\mathbf{P}_e^{r_{\max}+1}}{1-\mathbf{P}_e}\right\} & \text{if } r_{\max} > r^\star, \end{cases}$$

The equilibrium distribution is then completely specified by (3) and (4) as a function of the (still unknown) packet error probability $\mathbf{P}_e$.

The next step is to express $\mathbf{P}_e$ in terms of the equilibrium distribution derived for a tagged data flow. Firstly, we derive the equilibrium probability $\mathbf{P}_t^\star$ that a specific flow transmits a data packet (successfully or unsuccessfully) at a randomly selected event, given by

$$\mathbf{P}_t^\star = \sum_{r=0}^{r_{\max}} \pi(r,0) = \pi(0,0)\sum_{r=0}^{r_{\max}} \mathbf{P}_e^r = \frac{1-\mathbf{P}_e^{r_{\max}+1}}{1-\mathbf{P}_e}\pi(0,0). \tag{5}$$

In a system with n data flows, the probability that a tagged data packet is erroneous can be determined by conditioning on the number of simultaneous packet transfers:

$$\mathbf{P}_e = \sum_{k=1}^{n} \mathcal{B}(n-1, \mathbf{P}_t^\star, k-1)\,(1-\mathbf{P}_s^\star(k)), \tag{6}$$

where $\mathcal{B}(n,p,k)$ denotes the binomial probability of k successes out of n attempts given per-attempt success probability p. Expression (6) utilises the assumed independence of the different flows' packet transfer attempts. The $\mathbf{P}_s^\star(k)$ are specified in Section 3 and depend on the applied capture model. Note that unlike in [11, 12], the effects of capture are incorporated in the dynamics of the Markov chain, and hence influences the equilibrium distribution and, in particular, the packet transfer probability $\mathbf{P}_t^\star$. Observe that if we substitute (4) in (5), $\mathbf{P}_t^\star$ is expressed in terms of $\mathbf{P}_e$, while $\mathbf{P}_e$ in turn is expressed as a function of $\mathbf{P}_t^\star$ in (6).

Theorem 1 *A unique tuple $(P_t^\star, P_e)$ exists which satisfies expressions (4), (5) and (6).*

Proof. See the appendix. ■

THROUGHPUT ANALYSIS

In order to determine the expected aggregate data throughput, we first need to introduce some additional notation and parameters. Let τ denote the IEEE 802.11 time slot duration and $\mathbf{T}_s$ ($\mathbf{T}_c$) denote the expected inter-event time in case of a successful (erroneous) packet transfer which may or may not include the tagged data flow. In case the BASIC access mode is used, these inter-event times are

$$\begin{cases} \mathbf{T}_s^{\text{BASIC}} = \text{PHY} + \text{MAC} + r_{\text{WLAN}}^{-1}\mathbf{E}\{P\} + \delta + \text{SIFS} + \text{ACK} + \delta + \text{DIFS}, \\ \mathbf{T}_c^{\text{BASIC}} = \text{PHY} + \text{MAC} + r_{\text{WLAN}}^{-1}\mathbf{E}\{\overline{P}\} + \delta + \text{DIFS}, \end{cases}$$

where PHY and MAC denote the physical header (plus preamble) and MAC header sizes (converted to seconds), $\mathbf{E}\{P\}$ is the expected net payload size (in kbits), δ is the propagation delay between sender and receiver (in seconds), ACK is the acknowledgement message size (converted to seconds), and $\mathbf{E}\{\overline{P}\}$ is the expected net payload size of the largest packet involved in a collision (in kbits). In the RTS/CTS access mode we have

$$\begin{cases} \mathbf{T}_s^{\text{RTS/CTS}} = \text{RTS} + \delta + \text{SIFS} + \text{CTS} + \delta + \text{SIFS} + \text{PHY} + \text{MAC} + r_{\text{WLAN}}^{-1}\mathbf{E}\{P\} + \delta + \text{SIFS} + \text{ACK} + \delta + \text{DIFS}, \\ \mathbf{T}_c^{\text{RTS/CTS}} = \text{RTS} + \delta + \text{DIFS}, \end{cases}$$

The values of DIFS, SIFS, PHY, MAC, δ, RTS, CTS, ACK, τ, $\mathbf{E}\{P\}$ and r_{WLAN} are specified in Section 5.

From a *single flow*'s perspective in a system with n persistent data flows, the expected call-average data throughput is equal to the *event rate* $\times$ *the fraction of events that correspond with successful packet transfers* *(for the considered flow)* $\times$ *the expected transfer volume in case of a successful packet transfer:*

$$\mathbf{R}_{\text{flow}}(n) \equiv (\mathbf{E}\{\text{inter-event time}\})^{-1}\mathbf{P}_t^\star(1-\mathbf{P}_e)\,\mathbf{E}\{P\}$$

$$= \frac{\frac{1}{n}\sum_{k=1}^{n}\mathcal{B}(n,\mathbf{P}_t^\star,k)\,\mathbf{P}_s(k)\,\mathbf{E}\{P\}}{\mathcal{B}(n,\mathbf{P}_t^\star,0)\;\tau+\sum_{k=1}^{n}\mathcal{B}(n,\mathbf{P}_t^\star,k)\,\{\mathbf{P}_s(k)\,\mathbf{T}_s+(1-\mathbf{P}_s(k))\,\mathbf{T}_c\}}, \tag{7}$$

(in kbits/s) where the expected inter-event time (the inverse of the event rate) is determined by conditioning on the occurrence of three different event types: *(i)* none of the data flows attempts a packet transfer; *(ii)* some of the data flows attempt a packet transfer and the data packet with the strongest signal is succesfully captured by the intended receiver; *(iii)* some of the data flows attempt a packet transfer which all collide irrepairably. Note that the duration of a temporary freeze of the considered flow's backoff counter is incorporated in the expected inter-event time in the denominator, i.e. those times when the considered data flow does not attempt a packet transfer but one or more other data flows do.

To conclude this section, the expected *aggregate* data throughput is equal to

$$\mathbf{R}(n) \equiv n\mathbf{R}_{\text{flow}}(n)\,. \tag{8}$$

Observe that in the simple case of capture model CM-1 (see Section 3) expression (8) simplifies to

$$\mathbf{R}(n) = \frac{\mathcal{B}(n,\mathbf{P}_t^\star,1)\,\mathbf{E}\{P\}}{\mathcal{B}(n,\mathbf{P}_t^\star,0)\;\tau+\mathcal{B}(n,\mathbf{P}_t^\star,1)\,\mathbf{T}_s+\sum_{k=2}^{n}\mathcal{B}(n,\mathbf{P}_t^\star,k)\,\mathbf{T}_c}$$

$$= \frac{n\mathbf{P}_t^\star(1-\mathbf{P}_t^\star)^{n-1}\,\mathbf{E}\{P\}}{(1-\mathbf{P}_t^\star)^{n}\,\tau+n\mathbf{P}_t^\star(1-\mathbf{P}_t^\star)^{n-1}\,\mathbf{T}_s+\left(1-(1-\mathbf{P}_t^\star)^{n}-n\mathbf{P}_t^\star(1-\mathbf{P}_t^\star)^{n-1}\right)\mathbf{T}_c},$$

which is identical to the aggregate throughput expression given in [2, 22].

4.2 STAGE II: TRANSFER TIME ANALYSIS FOR NON-PERSISTENT FLOWS

From the flow level point of view we consider the WLAN as a service center serving flows at varying rates depending on the number of stations simultaneously active. In particular, when n stations are active the service rate per flow (station) is $\mathbf{R}(n)/n$, where $\mathbf{R}(n)$ is the aggregate data throughput derived in the previous section for the situation with n persistently active flows, $n = 1, ..., n_{\max}$. The resulting model is known as a Processor Sharing queueing model with state-dependent service rates and a finite number of service positions. Assuming, as in our case, that the time instants at which new flow transmissions start constitute a Poisson process, this PS model is analytically tractable. In particular, the equilibrium distribution of the number of flows simultaneously in progress is given by

$$\widetilde{\pi}(n) = \frac{\rho^n\varphi_n}{\sum_{j=0}^{n_{\max}}\rho^j\varphi_j} \quad \text{with} \quad \varphi_n \equiv \left(\prod_{j=1}^{n}\frac{\mathbf{R}(j)}{r_{\text{WLAN}}}\right)^{-1},$$

(see [6]) where $\rho \equiv \lambda/(\mu r_{\text{WLAN}})$ and $\varphi_0 \equiv 1$ by convention.

From the equilibrium distribution we can compute the expected number of flows present in the system and, using Little's formula [19], the expected flow transfer time $\mathbf{T}$:

$$\mathbf{T} \equiv \sum_{n=0}^{n_{\max}}\frac{n\,\widetilde{\pi}(n)}{\lambda(1-\widetilde{\pi}(n_{\max}))},$$

Some additional interesting results for the PS model have been derived (see [6]). In particular, the conditional expected transfer time $\mathbf{T}(x)$ of a flow of given size $x \geq 0$ can be computed explicitly and grows linearly in x:

$$\mathbf{T}(x) \equiv \frac{x}{r_{\text{WLAN}}}\sum_{n=0}^{n_{\max}}\frac{n\,\widetilde{\pi}(n)}{\rho(1-\widetilde{\pi}(n_{\max}))}, \tag{9}$$

a result which expresses the fair allocation of capacity to the served flows.

An important feature of the PS model is that these performance measures are *insensitive* with respect to the specific form of the flow size distribution, depending on its mean $1/\mu$ only. These attractive properties suggested by our modelling approach will be validated by simulation results of the WLAN system to be presented and discussed in the next section.

5 NUMERICAL RESULTS

In this section we present numerical results obtained from our analysis and compare them with simulation results, which have been produced by implementing the traffic and capture models, as well as a detailed representation of the IEEE 802.11 MAC layer in a C program. We present numerical results both for the throughput analysis for persistent flows (STAGE I) as well as for the transfer time analysis for non-persistent flows (STAGE II). For STAGE I our model is also compared with the models of [11, 12] in the considered scenario with capture. Note that for the situation without capture, the analytical model simplifies to that of [22]. The parameter settings for the MAC layer and the DSSS physical layer given in Table 1 are used[2], with the following default choices: a channel rate of $r_{\text{WLAN}} = 1$ Mbits/s and (therefore) a spreading factor of $S_f = 11$ and a required energy-per-bit to interference density ratio of $z_0 = 15$ dB in case of capture. Sufficient independent replications were run to obtain 95% confidence intervals with a relative precision no worse than 5%, except when otherwise noted.

parameter	value		parameter	value
r_{WLAN}	$\in \{1, 2, 5.5, 11\} \cdot 10^3$ kbits/s		δ	1 μs
$n_{\max}$	100		τ	20 μs
S_f	$\in \{11, 11, 8, 8\}$		SIFS	10 μs
z_0	$\in [6, 24]$ dB		DIFS	SIFS $+ 2 \times \tau = 50$ μs
PHY	192 bits	~ 192 μs	$cw_{\min}$	31
MAC	272 bits	$\sim 272/r_{\text{WLAN}}$ μs	$cw_{\max}$	1023
RTS	PHY + 160 bits	$\sim 192 + 160/r_{\text{WLAN}}$ μs	$r^\star$	5
CTS	PHY + 112 bits	$\sim 192 + 112/r_{\text{WLAN}}$ μs	$r_{\max}$	3 (BASIC access mode)
packet size	12 kbits		$r_{\max}$	6 (RTS/CTS access mode)
ACK	PHY + 112 bits	$\sim 192 + 112/r_{\text{WLAN}}$ μs		

Table 1: Parameter settings for the MAC layer and the DSSS physical layer.

5.1 STAGE I: THROUGHPUT RESULTS FOR PERSISTENT FLOWS

Consider the aggregated system throughput as a function of the number of persistent flows. Figure 5 (left) shows for the BASIC access mode both the scenario with (CM-2) and without (CM-1) capture, and Figure 5 (right) shows the corresponding results for the RTS/CTS access mode.

The results of Figure 5 show that both in the BASIC and RTS/CTS access mode our analytical model captures the behaviour of the WLAN extremely well. For the situation without capture, this was also observed in [22]. In the scenario with capture, we have also compared our analytical model with the model of [11, 12] (HVS model). As can be seen from the graphs, both in the BASIC and RTS/CTS access mode our model outperforms that of HVS that predicts the aggregate system throughput fairly well but too conservatively. In conclusion, our numerical results indicate that our analytical model accurately represents WLAN behaviour.

In the RTS/CTS access mode the aggregate system throughput increases for small numbers of persistent data flows but decreases for larger numbers. With more persistent flows, the average idle times between transmission attempts decreases as with more flows the probability that a flow has finished its contention window waiting

[2] Although in practice $\mathbf{E}\{P\}$ and $\mathbf{E}\{\overline{P}\}$ tend to be slightly smaller than the given packet size, due to the contribution of the typically smaller packet containing a non-persistent flow's remainder, we assume that $\mathbf{E}\{P\} = \mathbf{E}\{\overline{P}\} = 12$ kbits, in line with the assumption made in [22].

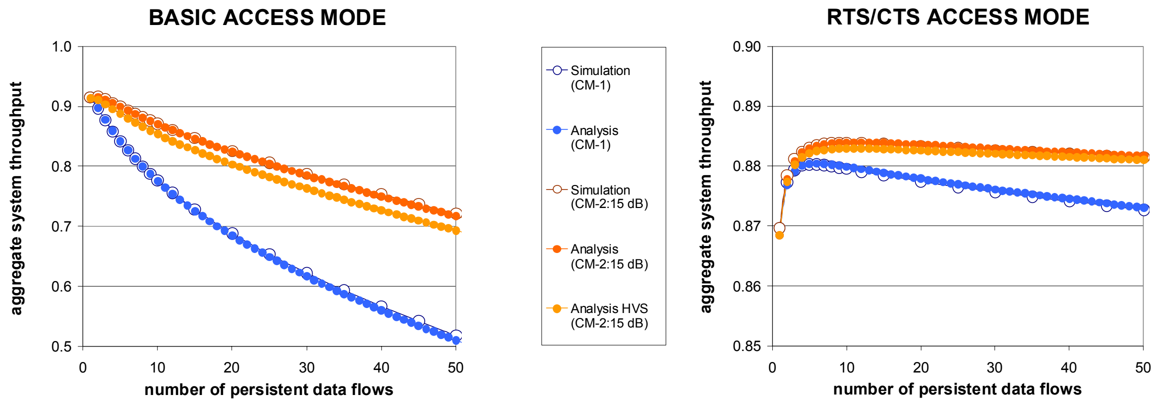

Figure 5: Aggregate system throughput as a function of the number of persistent data flows.

time is larger. For a larger number of flows this effect is dominated by a loss of efficiency due to destructive packet collisions. For the BASIC access mode we observe that the latter effect always dominates, as the system throughput appears to be strictly decreasing when the number of flows is larger than one. Note that the aggregate system throughput is substantially smaller than the channel rate of 1 Mbits/s due to the inefficiency on the MAC layer caused by waiting time, non-data packets and destructive collisions. This effect is even stronger for high channel rates as the duration of the DIFS, SIFS and PHY is independent of the channel speed.

As can be seen from both graphs, the effect of capture on the aggregate system throughput is significant. With $z_0 = 15$ dB in the case of the BASIC access mode the aggregate system throughput improves up to 40%. In the RTS/CTS access mode the effect of capture is much smaller, as the effect of collisions even without capture on the aggregate system throughput is less significant. Hence with an increasing number of persistent flows, the aggregate system throughput decreases only slightly when the number of persistent flows and thus the probability of a collision increases.

Comparing BASIC and RTS/CTS mode observe that for a small number of persistent data flows the BASIC access mode leads to a slightly higher aggregate system throughput as a consequence of the inherently higher resource efficiency if few collisions occur. In contrast as the number of persistent data flows increases, the additional overhead of the RTS/CTS access mode pays off in the sense that the aggregate system throughput decreases much less dramatically. Our results suggest the implementation in IEEE 802.11 of a combination of BASIC and RTS/CTS mode, where the system switches from BASIC to RTS/CTS mode at roughly 10 flows, yielding an interesting improvement of WLAN performance.

5.2 STAGE II: TRANSFER TIME RESULTS FOR NON-PERSISTENT FLOWS

In STAGE II we study the transfer time of files on flow level corresponding to the realistic scenario in which the number of active flows varies due to users intiating and terminating their data flows.

Let us first consider the mean flow transfer time $\mathbf{T}$ as a function of the offered traffic load ρ. Figure 6 shows results for the BASIC (left) and RTS/CTS (right) access mode for different flow size distributions in the scenario with capture for the following flow size distributions: deterministic, exponential and two hyperexponential distributions with coefficient of variation 2 and 4. The parameters of the hyperexponential distributions have been determined using the method of 'balanced means' (see [19]). The mean flow size for all considered distributions is equal to 120 kbits.

From these graphs, observe that the expected flow transfer time appears to be very insensitive to the specific form of the flow size distribution, depending only on its mean. Hence the insensitivity property as observed in

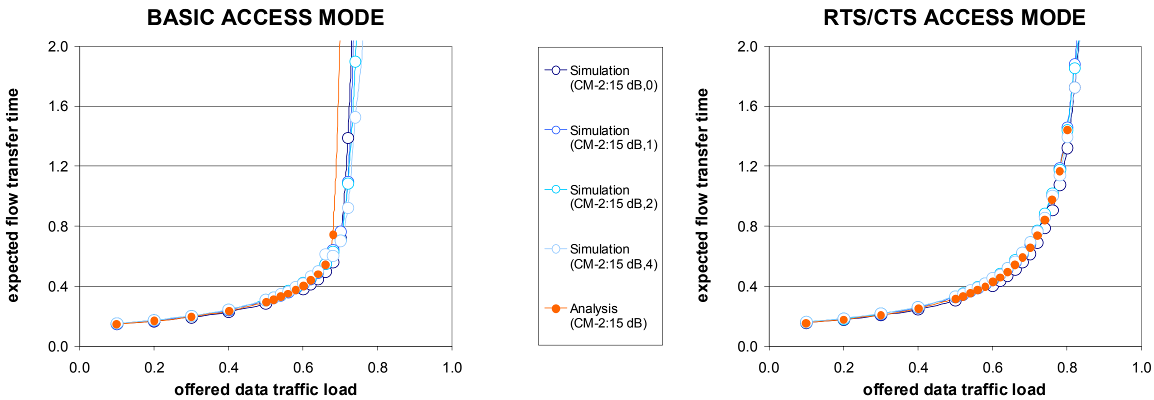

Figure 6: Expected flow transfer time as a function of the offered data traffic load.

Section 4.2 seems to be supported by the simulation results. Moreover, our analytical model approximates the WLAN behaviour extremely well both in BASIC and RTS/CTS access mode. The small deviation for high traffic loads in the BASIC model is due to numerical inaccuracies in the simulations as in this range the system becomes unstable.

Both for the BASIC and RTS/CTS access mode the expected flow transfer time increases gradually up to an offered data traffic load of 0.6 for BASIC access and 0.7 for RTS/CTS. At these values the expected transfer time increases rapidly as the expected number of present data flows approaches the CAC threshold. With more flows the available throughput per user does not only decrease because the aggregate throughput has to be shared with more users, but also because the aggregate throughput itself decreases when more flows share the channel (see Figure 5). In the RTS/CTS access mode this transfer time degradation is less dramatic (more gradual and at higher load values) than in the BASIC access mode which is directly reflected in the observation that the stability regime in the RTS/CTS access mode is larger than in the BASIC access mode.

Comparison of the expected flow transfer times for the RTS/CTS and BASIC access modes reveals that they hardly differ for an offered data traffic load up to about 0.5, and that in this regime the BASIC access mode appears to perform slightly better than the RTS/CTS access mode. In support of this observation, note that for lower loads the number of users in the system will be mostly in the range where the BASIC access mode outperforms the RTS/CTS access mode in STAGE I with respect to aggregate system throughput. The large difference in aggregate system throughput in STAGE I between the RTS/CTS and BASIC access mode is on the flow level reflected only by the larger stability regime of the RTS/CTS access mode, not by a lower expected transfer time for low to moderate traffic loads, i.e., due to the higher throughput the WLAN can handle more flows. Note that the traffic load the channel can handle is significantly smaller than 1 Mbits/s due to the inefficiencies and overhead incurred.

As a second experiment, consider the conditional expected transfer time $\mathbf{T}(x)$ as a function of the flow size x. In Figure 7 (left) results are shown for the BASIC access mode for different flow size distributions for the scenario with capture. In particular, the expected flow transfer time observed during the simulation are scattered against the corresponding flow sizes. The figure also contains the averaged flow transfer times. The results support the analytical claim of the insensitivity property as observed in Section 4.2 as the conditional expected transfer time is found to be independent of the chosen flow size distribution. Moreover, the simulation results confirm the linearity of $\mathbf{T}(x)$ in x (recall Equation (9)). These results indicate that our analytical model provides an excellent approximation for the behaviour of the WLAN.

Figure 8 depicts the expected flow transfer time as a function of the traffic load for the scenario with capture and for different channel rates. Observe that the expected flow transfer time rapidly increases for an offered

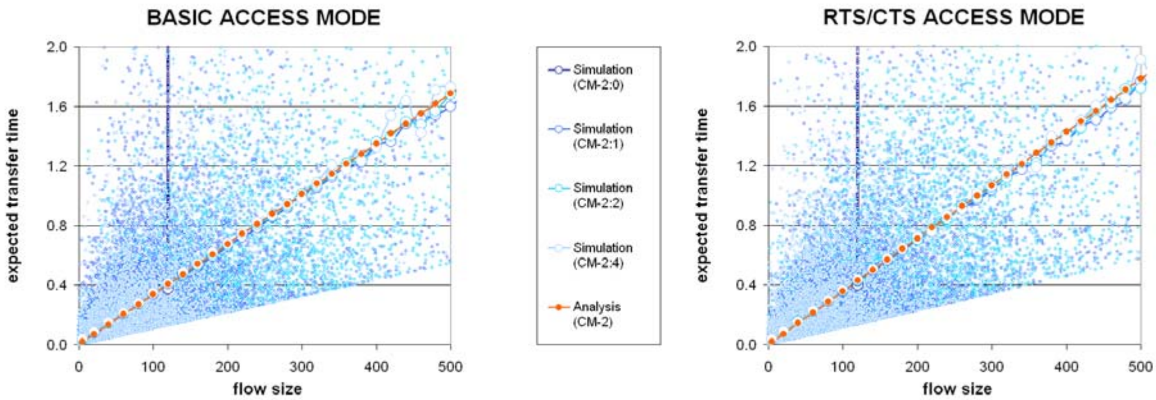

Figure 7: Scatter plot of flow transfer times as a function of the flow size. The plotted curve shows the averaged values.

traffic load between 0.4 and 0.7, depending on the channel rate. Note that the relative efficiency of the channel decreases for higher bit rates, as the time duration of the DIFS, SIFS and PHY is independent of the channel rate (cf. [13]). Hence, the maximal traffic load the channel can handle is smaller for high bit rate channels. Comparing the BASIC and RTS/CTS access mode we observe (as expected) that the RTS/CTS access mode can handle a higher data traffic load.

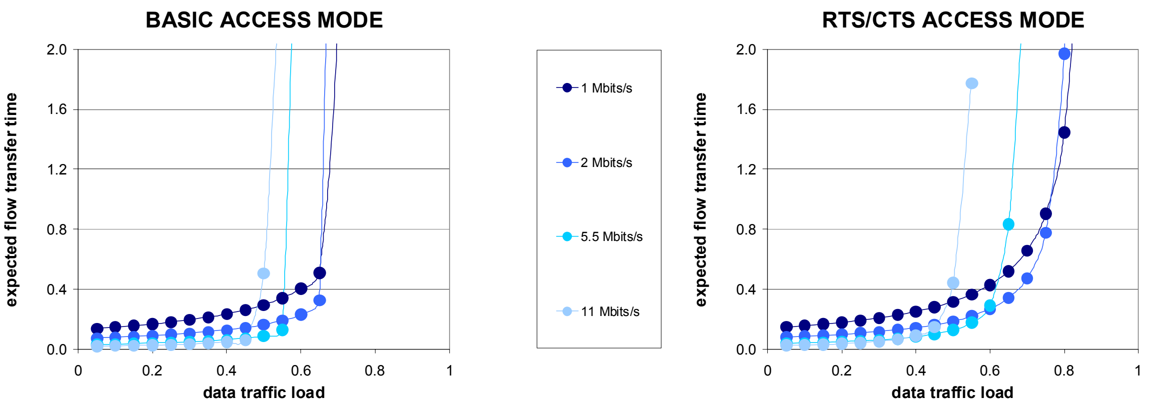

Figure 8: The expected flow transfer time as a function of the offered traffic load for various channel rates.

Finally, let us investigate the influence of both the channel rate and the required energy-per-bit to interference density ratio on the expected flow transfer time for a fixed traffic load of 0.6. In Figure 9 the results for the BASIC access mode (left) and the RTS/CTS access mode are presented. For the BASIC access mode we observe that for higher required energy-per-bit to interference density ratio the expected flow transfer time rapidly increases, as with an increasing capture threshold, the effective channel throughput decreases. For higher channel rates this will occur at lower capture thresholds due to the lower relative efficiency (see the discussion of Figure 8). For the RTS/CTS access mode the impact of the capture threshold is very small as was already observed in Section 5.1.

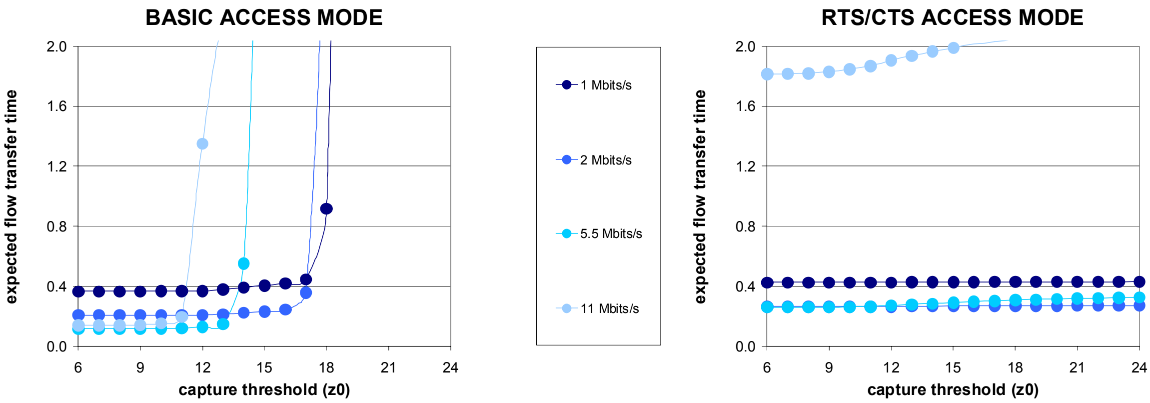

Figure 9: The expected flow transfer time as a function of the capture threshold.

6 CONCLUSIONS AND FURTHER RESEARCH

We have presented an integrated packet/flow level modelling approach for performance evaluation of a IEEE 802.11 WLAN with non-persistent bursty traffic sources. At the packet level, different physical and MAC layer system aspects are integrated into a single model, while still allowing explicit analytical evaluation.. The flow level model is based on the observation that a WLAN system behaves approximately as a queueing system with a Processor Sharing service discipline. Exploiting known performance results for Processor Sharing queues we have derived an analytical approximation for (conditional) expected flow transfer times. The accuracy of the approximation has been investigated by comparison of the analytical results with results obtained by simulation. The main conclusion from the numerical experiments is that the approximation yields very accurate results for all considered scenarios. In particular, the approximation very well reflects the sudden and steep increase of the mean flow transfer times when the offered traffic load approaches the maximum system throughput. Further, the numerical results show that the positive effect of packet capture on WLAN system throughput and flow transfer delays is considerable yet often ignored in WLAN performance studies. Our modelling approach also provides interesting general insights in the performance characteristics of WLANs. In particular, known results for the Processor Sharing model imply that expected flow transfer times are insensitive to the flow size distribution, which is confirmed by the simulation results.

One interesting application of our approximation, which deserves further study, is Call Admission Control. In particular, the analytical model enables swift determination of suitable flow admission thresholds for given requirements on expected flow transfer delays, expected throughputs and flow blocking probabilities. Other topics for further research are the inclusion of TCP flow control and MAC layer QOS differentiation mechanisms in our modelling approach (cf. [1] and [16]).

ACKNOWLEDGMENTS

The authors are indebted to Robert E. Kooij of TNO Telecom (The Netherlands) for filling in an exasperating gap in the proof of Theorem 1. This work has been partially funded by the Dutch Ministry of Economics Affairs under the program "Technologische Samenwerking ICT-doorbraakprojecten", project TSIT1025 BEYOND 3G. The TNO contribution to the reported research has been commissioned by KPN, The Netherlands.

REFERENCES

[1] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11", *Proceedings of IEEE INFOCOM '01*, Anchorage, USA, 2001.

[2] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function", *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535-547, 2000.

[3] K.G. Binmore, *"Mathematical analysis: a straightforward approach"*, Cambridge University Press, Cambridge, UK, 1982.

[4] F. Cali, M. Conti and E. Gregori, "IEEE 802.11 wireless LAN: capacity analysis and protocol enhancement", *Proceedings IEEE INFOCOM '98*, San Francisco, USA, 1998.

[5] H.S. Chhaya and S. Gupta, "Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol", *Wireless networks*, vol. 3, no. 3, pp. 217-234, 1997.

[6] J.W. Cohen, "The multiple phase service network with generalized processor sharing", *Acta informatica*, vol.12, pp. 245-284, 1979.

[7] D. Dardari, V. Tralli and R. Verdone, "On the capacity of slotted Aloha with Rayleigh fading: the role played by the number of interferers", *IEEE Communication letters*, vol. 2, no. 5, pp. 155-157, 2000.

[8] C.H. Foh and M. Zukerman, "Performance analysis of the IEEE 802.11 MAC protocol", *Proceedings of European Wireless '02*, Florence, Italy, 2002.

[9] T.S. Ho and K.C. Chen, "Performance Analysis of IEEE 802.11 CSMA/CA Medium Access Control protocol", *Proceedings of IEEE PIMRC '96*, Taipei, Taiwan, pp. 407-411, 1996.

[10] H. Honkasalo, K. Pehkonen, M.T. Niemi and A.T. Leino, "WCDMA and WLAN for 3G and beyond", *IEEE Wireless communications magazine*, vol. 9, no. 12, pp. 14-18, 2002.

[11] Z. Hadzi-Velkov and B. Spasenovski, "Capture effect in IEEE 802.11 wireless LANs", *Proceedings of IEEE ICWLHN '01*, Singapore, 2001.

[12] Z. Hadzi-Velkov and B. Spasenovski, "IEEE 802.11 DCF with capture over Ricean-fading channel", *presentation given at the third IEEE Workshop on WLANs '01*, Boston, USA, 2001, available at `http://www.wlan01.wpi.edu/proceedings/wlan26d.pdf`.

[13] IEEE, "IEEE standard for wireless LAN Medium Access Control (MAC) and PHYsical layer (PHY) specifications, P802.11, 1997.

[14] C.T. Lau and C. Leung, "Capture models for mobile packet radio networks", *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 917-925, 1992.

[15] W.C.Y. Lee, *"Mobile communications design fundamentals"*, Howard W. Sams & Co., Indianapolis, USA, 1986.

[16] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz and L. Stibor, "IEEE 802.11E wireless LAN for quality of service", *Proceedings of European Wireless '02*, Florence, Italy, 2002.

[17] N. Prasad and A. Prasad (editors), *"WLAN systems and wireless IP for next generation communications"*, Artech House, Norwood, USA, 2002.

[18] D.J. Struik (editor), *"A source book in mathematics 1200-1800"*, Princeton University Press, Princeton, USA, pp. 89-93, 1986.

[19] H.C. Tijms, *"Stochastic modelling and analysis: a computational approach"*, John Wiley & Sons, Chichester, England, 1986.

[20] C.Ware, J.F.Chicharo and T.Wysocki, "Modelling capture behaviour in IEEE 802.11 radio modems", *Proceedings of IEEE VTC '01*, Atlantic City, NJ, 2001.

[21] J. Weinmiller, M. Schlager, A. Festag and A. Wolisz, "Performance study of access control in wireless LANs IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN", *Mobile networks and applications*, vol. 2, no. 1, pp. 55-67, 1997.

[22] H. Wu, Y. Peng, K. Long, S. Cheng and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement", *Proceedings of IEEE INFOCOM '02*, New York, USA, 2002.

[23] W. Yue and Y. Matsumoto, "An exact analysis for CSMA/CA protocol integrated voice/data wireless LANs", *Proceedings of IEEE GLOBECOM '00*, San Francisco, USA, 2000.

APPENDIX: PROOF OF THEOREM 1

Proof. Expressions (5) and (6), along with $\pi(0,0)$ as specified in (4), explicitly define functions f and g given by

$$\begin{cases} \mathbf{P}_t^\star = f(\mathbf{P}_e) = 2\left(1 + \frac{\sum_{r=0}^{r_{\max}} cw_r \mathbf{P}_e^r}{\sum_{r=0}^{r_{\max}} \mathbf{P}_e^r}\right)^{-1}, \\ \mathbf{P}_e = g(\mathbf{P}_t^\star) = \sum_{k=1}^{n} \binom{n}{k} (\mathbf{P}_t^\star)^k (1-\mathbf{P}_t^\star)^{n-k} (1 - \mathbf{P}_s^\star(k)), \end{cases}$$

In overview, we will prove Theorem 1 by showing that f is a nonincreasing function from $[0,1]$ to $[f(1), f(0)]$ with $f(0) = 2/(1+cw_0) \geq f(1) = 2(r_{\max}+1)/\left(\sum_{r=0}^{r_{\max}}(1+cw_r)\right) > 0$, while g is a nondecreasing function from $[0,1]$ to $[0, 1-\mathbf{P}_s^\star(n)]$ with $0 \leq 1-\mathbf{P}_s^\star(n)$, which by Brouwer's fixed point theorem (e.g. [3]) implies the existence of a unique fixed point.

First we prove that f is nonincreasing. Defining

$$\psi(\mathbf{P}_e) \equiv \frac{\sum_{r=0}^{r_{\max}} cw_r \mathbf{P}_e^r}{\sum_{r=0}^{r_{\max}} \mathbf{P}_e^r},$$

$\mathbf{P}_t^\star$ can be written as

$$\mathbf{P}_t^\star = f(\mathbf{P}_e) = 2(1+\psi(\mathbf{P}_e))^{-1}.$$

We will prove that f is nonincreasing by showing that ψ is nondecreasing. Observe first that

$$\psi(0) = cw_0 \leq \psi(1) = \frac{1}{r_{\max}+1}\sum_{r=0}^{r_{\max}} cw_r,$$

as cw_r is nondecreasing in r. For $\psi(0) = \psi(1) = \psi^\star$, since cw_r is nondecreasing in r it must be that $cw_r = cw_0$, $r = 1, \cdots, r_{\max}$, and the function ψ must be constant (and hence indeed nondecreasing). Alternatively, for the case of $\psi(0) < \psi(1)$, we need to derive that for an arbitrary $\psi^\star \in (\psi(0), \psi(1))$, the number of times the function ψ crosses $\psi^\star$ is equal to 1. Notice that

$$\frac{\sum_{r=0}^{r_{\max}} cw_r \mathbf{P}_e^r}{\sum_{r=0}^{r_{\max}} \mathbf{P}_e^r} = \psi^\star \Longleftrightarrow \sum_{r=0}^{r_{\max}} (cw_r - \psi^\star)\mathbf{P}_e^r = 0. \tag{10}$$

Since cw_r is nondecreasing in r and $cw_0 < \psi^\star < cw_{r_{\max}}$, $(cw_r - \psi^\star)$ changes sign precisely once as r runs from 0 to $r_{\max}$. Invoking Descartes' sign rule (e.g. [18]), the polynomial in (10) has no more than a single positive root, and hence ψ crosses $\psi^\star$ no more than once. The continuity of ψ then implies that ψ must cross $\psi^\star$ precisely once. As a consequence, ψ is indeed a nondecreasing function, and hence f is nonincreasing in $\mathbf{P}_e$.

Secondly, we prove that g is nondecreasing, Using $0^0 = 1$, observe from (6) that $g(0) = \mathbf{P}_e(0) \leq \mathbf{P}_e(n-1) = g(1)$, and

$$\begin{aligned}
\frac{d}{d\mathbf{P}_t^\star} g(\mathbf{P}_t^\star) &= \frac{d}{d\mathbf{P}_t^\star} \left\{ \sum_{k=1}^{n} \binom{n}{k} (\mathbf{P}_t^\star)^k (1-\mathbf{P}_t^\star)^{n-k} (1-\mathbf{P}_s^\star(k)) \right\} \\
&= \sum_{k=1}^{n} (1-\mathbf{P}_s^\star(k)) \binom{n}{k} \left\{ k(\mathbf{P}_t^\star)^{k-1}(1-\mathbf{P}_t^\star)^{n-k} - (n-k)(\mathbf{P}_t^\star)^k (1-\mathbf{P}_t^\star)^{n-k-1} \right\} \\
&= \sum_{k=1}^{n-1} (\mathbf{P}_t^\star)^k (1-\mathbf{P}_t^\star)^{n-k-1} \frac{n!}{(n-k-1)!k!} \{\mathbf{P}_s^\star(k) - \mathbf{P}_s^\star(k+1)\} \geq 0,
\end{aligned}$$

for $\mathbf{P}_t^\star \in [0,1]$, since $\mathbf{P}_s^\star(k)$ was assumed to be nonincreasing in k. Hence g is nondecreasing in $\mathbf{P}_t^\star$. ■