A finite-source feedback queueing network as a model for the IEEE 802.11 Distributed Coordination Function

E.M.M. Winands[1], T.J.J. Denteneer[2], J.A.C. Resing[1] & R. Rietman[2]

[1]**Eindhoven University of Technology**
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[2]**Philips Research**
Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

Abstract: **The most mature medium access control protocol for wireless local area networks is the IEEE 802.11 standard. The primary access mode of this protocol is based on the mechanism of carrier sense multiple access with binary exponential backoff. We develop a finite-source feedback queueing model for this access mode. In this model we derive expressions for the throughput and the first moment of the delay and a set of equations to compute the Laplace Stieltjes Transform and higher moments of the delay. The accuracy of the model is verified by simulation.**

1. Introduction

The most important *medium access control* (MAC) protocol for a radio frequency wireless *local area network* (LAN) is the IEEE 802.11 MAC protocol [1]. The primary access mode of this protocol is the *distributed coordination function* (DCF), that creates an *asynchronous* type of service. The following subsection provides an introduction to the DCF. For more detailed information see [1].

1.1. Distributed coordination function

The basic scheme used for data transmission in the DCF is a *carrier sense multiple access* (CSMA) protocol, wherein stations listen to the channel before transmission. Stations access the channel using a basic access method or an optional RTS/CTS mechanism. Under the basic access method a station is allowed to transmit its packet, if the medium is idle for a period of time greater than the *distributed interframe space* (DIFS). Otherwise, it persists to monitor the channel until it is measured idle for a DIFS. At this time, a backoff interval is generated uniformly distributed between zero slots and the contention window size W. This backoff interval initializes the backoff timer of the transmitter.

Only when the medium is sensed as being idle, the backoff timer of a station is decremented. When the backoff timer reaches zero, the station is allowed to access the channel. A collision occurs when several stations start transmission simultaneously. Therefore, an *acknowledgement* (ACK) is sent by the receiving station after a period of time called the *short interframe space* (SIFS), which is smaller than a DIFS. If the ACK is not received, the transmitting station reschedules the corresponding packet according to the described backoff procedure.

After each unsuccessful transmission, the size of the contention window is changed according to the *binary exponential backoff* procedure. The contention window W at the first transmission attempt equals W_{min} and is doubled at each subsequent retransmission of the packet

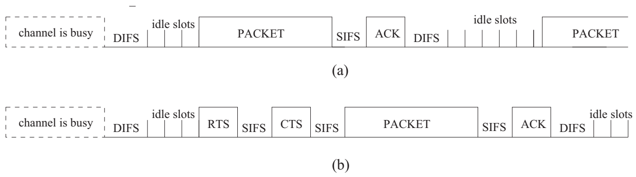

Figure 1: The IEEE 802.11 access mechanisms: (a) Basic access mechanism; (b) RTS/CTS mechanism.

until a predefined maximum W_{max} is reached. Retries shall continue until the transmission is successful or until the retry limit M' has been attained and the current packet is rejected. After a successful packet transmission or after reaching the retry limit, the window size is reinitialized to W_{min}.

To shorten the duration of collisions the DCF includes a four-way handshaking mechanism based on the exchange of the *request to send* (RTS) and the *clear to send* (CTS) frame. After a successful exchange of these control frames, the data packet can be sent by the transmitter. The basic and the RTS/CTS mechanism are depicted in Figure 1a and Figure 1b, respectively. In this paper we will focus on the RTS/CTS mechanism.

1.2. Quality-of-service

The goal of our research is to get more insight in the performance of an IEEE 802.11 wireless LAN and to measure the impact of aspects like the size of the network and the load. We use the following measures to characterize the quality-of-service level in the wireless network

1. *throughput* (γ): the number of messages transmitted per time slot;
2. *delay* (D) : the number of time slots needed to complete the transmission of a message.

1.3. Review of related literature

The performance of the DCF has been studied either by simulation or by approximate mathematical models. Most of the simulation models (e.g. see Anastasi et al. [2], Bianchi et al. [3] and Weinmiller et al. [17]) have been developed for the purpose of analyzing alternative design parameters or enhancements of the original MAC protocol. In most of the analytical models, simplifying assumptions have been made on the retransmission algorithm. In particular, a constant backoff window has been analyzed in Chhaya and Gupta [8] and Huang and Chen [11], while Cali et al. [6] assume a backoff interval sampled from a geometric distribution. Ho and Chen

[10] make the assumption that a station only doubles its contention window after the first collision and does not change it anymore even if the station suffers more collisions. Next to the protocol simplifications, many models assume simple traffic conditions to facilitate their analysis. Bianchi [4] has analyzed the DCF under a saturation condition, i.e. all stations are active. Wu et al. [19] and Chatzimisios et al. [7] extended the model proposed by Bianchi and took the maximum retry limit into account. Vishnevskii and Lyakov [15] and [16] advanced this model even further by including noise and the seizing effect.

We develop a queueing model in which the number of active stations is variable. With the help of this model, we quantify the performance of a network that deploys the IEEE 802.11 DCF. The main contribution of the present paper is that it not only gives approximations for the throughput and the *mean* delay, but also for *higher moments* of the delay.

1.4. Outline

The rest of the paper is organized as follows. In Section 2 we develop a product-form queueing model for an IEEE 802.11 wireless LAN. Furthermore, it describes the methods that are employed to inject our model with the characteristics of the wireless multiaccess communication. The analytical performance study is presented in Section 3. The numerical evaluation of the queueing model in relation with the wireless LAN is discussed in Section 4. Finally, Section 5 outlines the most important conclusions of our research. Besides, recommendations for future research are made.

2. Model

2.1. Assumptions

We first list the traffic assumptions for the wireless IEEE 802.11 network.

1. Identical, independent and stationary stations. We assume that the wireless LAN contains an arbitrary but finite number N of identical, independent and stationary stations.

2. ON/OFF arrival process with geometrically distributed message sizes. We characterize the bursty packet arrivals process of a station by a stream of packets during talkspurts (ON-periods) and no arrivals during silences (OFF-periods). The assumption is made that a station is in saturation during its ON-period, i.e. it immediately has a packet available for transmission after the completion of each successful transmission. The length of an ON-period is determined by the number of packets a station wants to transmit; the *message* size L. We assume that the message size is a shifted geometrically distributed random variable with parameter q

$$P(L = k) = (1 - q)q^{k-1}, \quad k = 1, 2, \ldots . \tag{1}$$

Thus an ON-process can be thought of as one in which each station generates a new packet with probability q after a previous packet has been transferred. Yet another view is provided by the notion that each station instantaneously generates a geometrically distributed number of packets with parameter q at the start of a new ON-period.

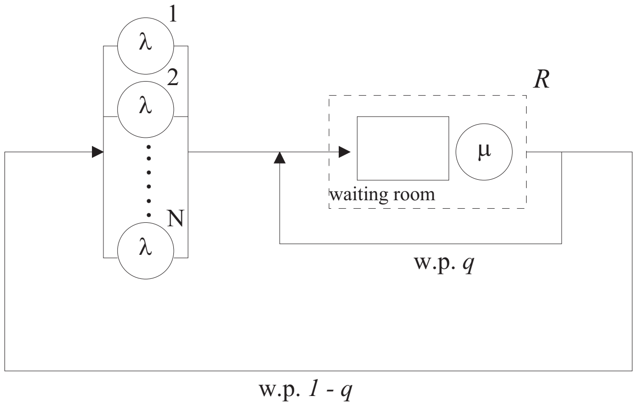

Figure 2: The model.

The length of the OFF-period is assumed to be an exponentially distributed random variable with parameter λ.

3. Exponentially distributed packet payload sizes. A transmission length is composed of the packet payload P and all kinds of fixed-length overhead such as DIFS, SIFS, ACK etcetera. In order to facilitate the mathematical analysis of this paper, we assume that the packet payloads are exponentially distributed.

4. Ideal channel conditions. We ignore the possibility of errors due to noise and the possibility of *capture*, the ability of a receiver to successfully receive a transmission from a given station when multiple stations are transmitting simultaneously (see Chhaya and Gupta [8]).

2.2. Model

In order to study the system under the assumptions made, the shared wireless communication medium is represented by a waiting room and a server, which serves the packets of the stations. Requests for service are generated by the finite number N of identical, independent and stationary stations. The stations become active, independently, after an exponentially distributed OFF-period with parameter λ. Active stations queue up in the waiting room, where their packets are served according to the implicit *service discipline* of the DCF (see Subsection 2.3). The time needed to serve one single packet of an active station is called the *service time* (see Subsection 2.3). After having been served, a station generates a new packet with probability q. On the other hand, with probability $1 - q$ each station becomes inactive again. It then becomes active again on the generation of a new message (see Figure 2).

Remark 1. Obviously, the proposed model is an extension of the well-known *machine repair model*. In the machine repair model requests for service are generated by a finite number of identical machines that operate independently of each other. The machines break down after an exponentially distributed working time. Broken machines queue up in the single server repair queue, where they are repaired. In the proposed extension, with probability q machines are not repaired as it should be. For consistency, we adopt the terminology corresponding to the wireless LAN in this paper. This means that, for example, we do not speak of machines that break down but of stations that generate packets.

□

2.3. Service characteristics

Before the proposed model can be analyzed, we must be explicit about the following service characteristics.

1. Service discipline. The service discipline of the DCF determines which packet will be transmitted next after a successful transmission. This discipline is determined by two contrasting effects. On the one hand, stations with old packets are only delayed a residual backoff interval and, therefore, have a high probability to be the first to transmit. On the other hand, a station that has just completed transmission randomizes its new attempt over the smallest possible contention window. In Winands [18] it is shown by means of a simulation study that the resulting service discipline of the DCF on a packet level is slightly more unfair than *random order of service* (ROS). However, it turns out that this difference vanishes on a message level (see Section 4.3). Hence, we assume that the server handles only one packet (of a randomly chosen station) at a time. After a service completion, the next station to be served is chosen randomly out of the active stations. Finally, a station stays active until all its phases of service are completed, i.e. until all its packets have been successfully transmitted.

2. Service time distribution. In order to obtain the service time distribution we use results of Bianchi [4] for the wireless LAN in *overload*, also called *saturation*, conditions. More specifically, the service time distribution can be obtained by analyzing the distribution of the time between successive successful transmissions given i active stations; the *virtual transmission time* T_i. The virtual transmission time consists of three components: length of a successful transmission, collision intervals and idle periods. Bianchi [4] gives an approximation of the mean virtual transmission time $E[T_i]$. From this mean virtual transmission time, the packet service rate μ_i given i active stations can be obtained by

$$\mu_i = \frac{1}{E[T_i]}. \tag{2}$$

In Bianchi [4] it is also shown that T_i is virtually independent of the number of active stations, when the RTS/CTS access mechanism is deployed. Therefore, the packet service rate μ in the proposed model with N stations is assumed to be state-independent and is approximated by the following arithmetic mean

$$\mu = \frac{\sum_{i=1}^{N} \mu_i}{N}, \tag{3}$$

where μ_i is given by Equation (2). In Winands [18], Bianchi's work is extended with an approximation of the *variance* of T_i. It is shown that the coefficient of variation of T_i is close to one and that the individual virtual transmission time realizations are mutually independent. Henceforth, we assume that the packet service requirements of a station are exponentially distributed random variables with parameter μ.

Remark 2. The idea to obtain service rates from Bianchi's saturation analysis was proposed by Foh and Zukerman [9]. Using the results of Bianchi, the authors construct several queueing models with state-dependent service rates. Litjens et al. [13] develop a processor sharing queueing model with state-dependent service rates, obtained with the help of Bianchi's analysis, and a finite number of service positions. Approximations for the throughput and the *mean* delay are derived in these papers. It is emphasized here that the present paper gives approximations for *higher moments* of the delay as well.

□

3. Performance analysis

3.1. Equilibrium distribution

At an arbitrary instant the state of the system is fully described by the number of stations k that is inactive; then, the number of active stations is $N-k$. We denote the random variables X and Y as the number of stations inactive and active, respectively, at an arbitrary epoch. The equilibrium distributions of X and Y can be obtained via a straightforward balance argument or using known queue length results for this network, which is of product-form, that is

$$P[X = k] = P[Y = N-k] = \frac{\frac{\rho^k}{k!}}{\sum_{i=0}^{N} \frac{\rho^i}{i!}}, \; k = 0, \ldots, N, \tag{4}$$

where

$$\rho = \frac{\mu(1-q)}{\lambda}.$$

Further, for the mean and variance of X and Y we can easily obtain the following expressions

$$\begin{aligned} E[X] &= N - E[Y] \\ &= \rho(1 - B_N(\rho)), &(5)\\ \text{var}[X] &= \text{var}[Y] \\ &= E[X] - \rho B_N(\rho)(N - E[X]), &(6) \end{aligned}$$

where $B_N(\rho)$ denotes Erlang's loss probability, which is given by

$$B_N(\rho) = \frac{\frac{\rho^N}{N!}}{\sum_{i=0}^{N} \frac{\rho^i}{i!}}. \tag{7}$$

The determination of the delay depends on the distribution of the system state as seen by a tagged station at instants both at which the tagged station becomes active and at which the service of the tagged station ends, but the ON-period continues. Stochastic quantities related to these moments will be denoted by a subscript 1. From the arrival theorem for a closed product-form network (see Sevcik and Mitrani [14]), it follows that these distributions are related simply to the equilibrium distribution of the network with the population decreased by one station, i.e.

$$\begin{aligned} P[X_1 = k] &= P[Y_1 = N-1-k] \\ &= \frac{\frac{\rho^k}{k!}}{\sum_{i=0}^{N-1} \frac{\rho^i}{i!}}, \quad k = 0, \ldots, N-1. &(8) \end{aligned}$$

3.2. Performance measures

We can easily derive the following relationship

$$\gamma = \lambda E[X] = \mu(1-q)(1-B_N(\rho)), \tag{9}$$

that equates the throughput to the mean number of stations becoming active per time slot. Moreover, the following relationship between the mean delay $E[D]$ and the mean number $E[Y]$ of active stations can be obtained with the help of Little's law

$$\begin{aligned} E[D] &= \frac{E[Y]}{\gamma} = \frac{1}{\mu(1-q)}\frac{N-\rho(1-B_N(\rho))}{1-B_N(\rho)} \\ &= \frac{1}{\mu(1-q)}(N-\rho(1-B_{N-1}(\rho))). \end{aligned} \tag{10}$$

For N large and $N >> \frac{\mu(1-q)}{\lambda}$, Erlang's loss probability $B_N(\rho)$ goes to zero like $\frac{\rho^N}{N!}$. The following are sharp approximations for the throughput and mean delay, respectively, in this regime (cf. Boxma et al. [5])

$$\gamma \approx \mu(1-q), \tag{11}$$

$$E[D] \approx \frac{N}{\mu(1-q)} - \frac{1}{\lambda}. \tag{12}$$

Remark 3. Since the message size L is a shifted geometrically distributed random variable, the *Laplace Stieltjes Transform* (LST) of the total service requirement T becomes

$$\begin{aligned} \tilde{T}(\omega) &= E[e^{-\omega T}] = \sum_{k=1}^{\infty} E[e^{-\omega T}|L=k]P[L=k] \\ &= \frac{\mu(1-q)}{\mu(1-q)+\omega}, \end{aligned} \tag{13}$$

which is recognized as the LST of the exponential distribution with parameter $\mu(1-q)$. The equilibrium distribution and the above mean performance measures can also be obtained by assuming that the packets of a station are served in one stretch with service time T. Now, the model reduces to the standard machine repair model, for which the equilibrium distribution and average performance measures can be found in e.g. Kleinrock [12].

□

We now turn to the distribution of the delay. A tagged station is studied at the moment that it becomes active. Let $D_{k,i}$ denote the (remaining) steady-state delay of the tagged station when there are k other stations in the waiting room and it is either in the service position ($i=1$) or waiting ($i=0$). Introduce

$$\phi_{k,i}(\omega) := E[e^{-\omega D_{k,i}}],\ \mathrm{Re}(\omega) \geq 0,\ k=0,\ldots,N-1, \tag{14}$$

as the corresponding LST. If there are stations active upon the moment at which the tagged station becomes active, the tagged station is placed in the waiting room before it is taken into service. Only when the waiting room is empty upon arrival, the tagged station is taken into service immediately. Therefore, the following expression

$$\begin{aligned} \tilde{D}(\omega) = E[e^{-\omega D}] &= \sum_{k=1}^{N-1} P[Y_1=k]\phi_{k,0}(\omega) + \\ & \quad P[Y_1=0]\phi_{0,1}(\omega), \end{aligned} \tag{15}$$

corresponds to the LST of the (remaining) delay D of the tagged station at the service facility. Moments of the delay can be obtained by the LST as follows

$$E[D^m] = \sum_{k=1}^{N-1} P[Y_1=k]E[D_{k,0}^m] + P[Y_1=0]E[D_{0,1}^m], \quad m \in \mathbb{Z}^+, \tag{16}$$

with

$$E[D_{k,i}^m] = (-1)^m [\frac{d^m}{d^m \omega}\phi_{k,i}(\omega)]_{\omega=0}, \quad m \in \mathbb{Z}^+. \tag{17}$$

The following set of $2N-1$ equations for equally many unknowns $\phi_{1,0}(\omega),\ldots,\phi_{N-1,0}(\omega)$, $\phi_{0,1}(\omega),\ldots,\phi_{N-1,1}(\omega)$ holds

$$\begin{aligned} \phi_{k,0}(\omega) &= \frac{\mu+(N-k-1)\lambda}{\mu+(N-k-1)\lambda+\omega}\Big[\frac{(N-k-1)\lambda}{\mu+(N-k-1)\lambda}\phi_{k+1,0}(\omega) + \\ & \frac{\mu q}{\mu+(N-k-1)\lambda}(\frac{1}{k+1}\phi_{k,1}(\omega) + \frac{k}{k+1}\phi_{k,0}(\omega)) + \\ & \frac{\mu(1-q)}{\mu+(N-k-1)\lambda}(\frac{1}{k}\phi_{k-1,1}(\omega) + \frac{k-1}{k}\phi_{k-1,0}(\omega))\Big], \\ & \qquad k=1,\ldots,N-1, \end{aligned} \tag{18}$$

$$\begin{aligned} \phi_{k,1}(\omega) &= \frac{\mu+(N-k-1)\lambda}{\mu+(N-k-1)\lambda+\omega}\Big[\frac{(N-k-1)\lambda}{\mu+(N-k-1)\lambda}\phi_{k+1,1}(\omega) + \\ & \frac{\mu q}{\mu+(N-k-1)\lambda}(\frac{1}{k+1}\phi_{k,1}(\omega) + \frac{k}{k+1}\phi_{k,0}(\omega)) + \\ & \frac{\mu(1-q)}{\mu+(N-k-1)\lambda}\Big], \qquad k=0,\ldots,N-1. \end{aligned} \tag{19}$$

The distinction is made whether the tagged station is in service or not. We first explain Equation (18). After the tagged station has become active, the next event occurs after a time period with LST $\frac{\mu+(N-k-1)\lambda}{\mu+(N-k-1)\lambda+\omega}$. This event is either a packet transmission or a station becoming active. The latter occurs first with probability $\frac{(N-k-1)\lambda}{\mu+(N-k-1)\lambda}$ and then the memoryless property of the exponential distribution implies that the tagged station sees the system as if it only now becomes active, meeting $k+1$ stations. With probability $\frac{\mu}{\mu+(N-k-1)\lambda}$ a packet transmission occurs first and the served station either stays active with probability q or becomes inactive with probability $1-q$. Now, the tagged station is taken into service with either probability $\frac{1}{k+1}$ in the case that the served station stays active or with probability $\frac{1}{k}$ in the case that the served station becomes inactive. If the tagged station does not enter the service position, it sees the waiting room as if it only now becomes active meeting either $k-1$ stations or k stations dependent on whether the served station stays active or not. The explanation of Equation (19) is quite similar. After the service of the tagged station has been started, a station becomes active with probability $\frac{(N-k-1)\lambda}{\mu+(N-k-1)\lambda}$, the tagged station requires another phase of service with probability $\frac{\mu q}{\mu+(N-k-1)\lambda}$ and its service

Physical configuration parameters		
Parameter	**Abbreviation / notation**	**Duration** (time slots)
Acknowledgement	ACK	4.80
Clear to send frame	CTS	4.80
Distributed interframe space	DIFS	2.56
Headers	H	8.00
Request to send frame	RTS	5.76
Short interframe space	SIFS	0.56
Slot time	σ	1.00
Mean packet payload size	$E[P]$	163.68
Protocol parameters		
Access mode	**Parameter**	**Value**
Binary exponential backoff	W_{min}	32
	W_{max}	1024
	M'	∞
Connection parameters		
Parameter	**Abbreviation / notation**	**Value**
Number of stations	N	$(10, 25)$
Mean message size (packets)	$E[L]$	(20)
Load	$\frac{\Lambda}{\mu(1-q)}$	$(0.25, 0.50, 1, 2, 4, 8)$

Table 1: System parameter set-up.

ends with probability $\frac{\mu(1-q)}{\mu+(N-k-1)\lambda}$. If the tagged station stays active, it is immediately taken into service again with probability $\frac{1}{k+1}$ and it is placed in the waiting room with probability $\frac{k}{k+1}$. It is noticed that the pre-factors of the terms $\phi_{0,0}(\omega)$, $\phi_{N,0}(\omega)$ and $\phi_{N,1}(\omega)$ in Equations (18) and (19) equal zero. Furthermore, the set of equations defined by (18) and (19) can be used to obtain numerical values of the higher moments of the delay D (see Section 4.3).

Remark 4. It is noteworthy that substituting $q = 0$, that means single-packet messages, leads to the set of equations for the LST of the delay in the machine repair model with ROS service discipline derived by Boxma et al. [5]. Their model has been developed for contention resolution in cable networks using contention trees.

□

4. Numerical results

In this section, we present numerical results that show that the proposed model can be used as approximation for the performance measures in the wireless LAN. Since we have no explicit expressions for the performance measures of the wireless network, these are obtained via a *discrete-event* simulation. We simulate a wireless network that deploys the IEEE 802.11 DCF (cf. [1]) under the traffic assumptions of Subsection 2.1. All simulation results in this section are shown with an approximate 95% confidence interval. The corresponding performance measures in the queueing model have been derived in Section 3.

4.1. Network settings

The parameters characterizing the wireless LAN can be classified into three groups.

1. Physical configuration parameters. The physical configuration parameters characterize the configuration of the wireless LAN. The detailed specification of the physical configuration parameters is given in the first part of Table 1. These parameter values are those specified for the *frequency hopping spread spectrum* (FHSS) physical layer. The mean packet payload size $E[P]$ has been chosen about one fourth of the maximum data frame size for the FHSS physical layer (cf. Bianchi [4]). Furthermore, these values correspond to a channel bit rate of 1 Mbit/s.

2. Protocol parameters. We make the assumption that the stations in the wireless network deploy the RTS/CTS access mechanism without retry limit. The second part of Table 1 gives a concise overview of the protocol parameter setting. This parameter setting is typically studied in the literature. Both the physical configuration and protocol parameters remain unchanged in the course of our performance evaluation.

3. Connection parameters. We consider two network sizes, $N = 10$ and $N = 25$ and assume that a message consists on average of twenty packets, $E[L] = 20$. We investigate and compare the performance for various settings of the load $\frac{\Lambda}{\mu(1-q)}$, where $\Lambda = N\lambda$ equals the total traffic intensity.

Table 2 lists the consequences of the network settings for the mean service time. It is seen that the mean service time is almost insensitive to the network size. Obviously, this result is closely related to the insensitivity result for the virtual transmission time in small wireless networks (see Bianchi [4]).

N	**Mean service time ($\frac{1}{\mu}$)**
10	197.6
25	196.4

Table 2: Mean service time.

Mean delay ($E[D]$)				
	$N = 10$		$N = 25$	
$\frac{\Lambda}{\mu(1-q)}$	**Simulation**	**Analysis**	**Simulation**	**Analysis**
0.25	$0.511\ 10^4 \pm 0.007\ 10^4$	$0.501\ 10^4$	$0.526\ 10^4 \pm 0.008\ 10^4$	$0.513\ 10^4$
0.5	$0.672\ 10^4 \pm 0.009\ 10^4$	$0.649\ 10^4$	$0.733\ 10^4 \pm 0.012\ 10^4$	$0.713\ 10^4$
1	$1.076\ 10^4 \pm 0.017\ 10^4$	$1.079\ 10^4$	$1.623\ 10^4 \pm 0.035\ 10^4$	$1.649\ 10^4$
2	$2.027\ 10^4 \pm 0.022\ 10^4$	$2.050\ 10^4$	$4.781\ 10^4 \pm 0.059\ 10^4$	$4.942\ 10^4$
4	$2.882\ 10^4 \pm 0.022\ 10^4$	$2.964\ 10^4$	$7.263\ 10^4 \pm 0.057\ 10^4$	$7.364\ 10^4$
8	$3.378\ 10^4 \pm 0.020\ 10^4$	$3.457\ 10^4$	$8.307\ 10^4 \pm 0.052\ 10^4$	$8.592\ 10^4$

Table 4: Mean delay ($E[L] = 20$).

Standard deviation of the delay (σ_D)				
	$N = 10$		$N = 25$	
$\frac{\Lambda}{\mu(1-q)}$	**Simulation**	**Analysis**	**Simulation**	**Analysis**
0.25	$0.541\ 10^4 \pm 0.011\ 10^4$	$0.550\ 10^4$	$0.565\ 10^4 \pm 0.012\ 10^4$	$0.572\ 10^4$
0.50	$0.758\ 10^4 \pm 0.014\ 10^4$	$0.755\ 10^4$	$0.871\ 10^4 \pm 0.019\ 10^4$	$0.868\ 10^4$
1	$1.251\ 10^4 \pm 0.023\ 10^4$	$1.272\ 10^4$	$1.984\ 10^4 \pm 0.051\ 10^4$	$2.067\ 10^4$
2	$2.147\ 10^4 \pm 0.026\ 10^4$	$2.185\ 10^4$	$4.977\ 10^4 \pm 0.068\ 10^4$	$5.100\ 10^4$
4	$2.881\ 10^4 \pm 0.026\ 10^4$	$2.983\ 10^4$	$7.317\ 10^4 \pm 0.070\ 10^4$	$7.380\ 10^4$
8	$3.335\ 10^4 \pm 0.025\ 10^4$	$3.446\ 10^4$	$9.242\ 10^4 \pm 0.067\ 10^4$	$8.580\ 10^4$

Table 5: Standard deviation of the delay ($E[L] = 20$).

4.2. Throughput

The throughput γ as derived in Equation (9) is used as an approximation for the throughput in the wireless LAN. For presentation reasons, we have not listed the throughput itself in Table 3, but the fraction of bandwidth used for transmission of payload bits instead. This quantity can be derived by multiplying the throughput γ by the mean number $E[L]$ of packets in a message and by the mean packet payload size $E[P]$.

Fraction of bandwidth used for transmission of payload				
	$N = 10$		$N = 25$	
$\frac{\Lambda}{\mu(1-q)}$	**Simulation**	**Anal.**	**Simulation**	**Anal.**
0.25	0.203 ± 0.002	0.201	0.205 ± 0.002	0.206
0.5	0.382 ± 0.003	0.383	0.400 ± 0.003	0.404
1	0.648 ± 0.003	0.651	0.711 ± 0.003	0.714
2	0.814 ± 0.004	0.813	0.836 ± 0.005	0.833
4	0.841 ± 0.005	0.828	0.836 ± 0.005	0.834
8	0.840 ± 0.005	0.829	0.843 ± 0.005	0.834

Table 3: Fraction of bandwidth used for transmission of payload bits ($E[L] = 20$).

From this table we can draw two conclusions. First of all, it immediately strikes the eye that the throughput in the queueing model provides an accurate approximation to the throughput in the wireless LAN. Secondly, it can be verified that as the arrival intensity Λ increases the throughput approaches the service parameter μ multiplied by the factor $(1-q)$, which follows from Equation (11). A related performance measure is the *saturation throughput*, which is defined as the limit of the throughput as the offered load increases (see Bianchi [4]). This performance measure indicates the overhead required by the DCF to perform its coordination task among stations (see Cali et al. [6]). The entries in Table 3 with $\Lambda >> \mu(1-q)$ are relevant for this case. It is seen that the limit of the fraction of bandwidth used for transmission of payload bits is approximately equal to 84%.

4.3. Delay

We now turn to a comparison of the delay in the wireless LAN and the delay in the queueing model. We first consider first moments and afterwards standard deviations. The mean delay is approximated by the mean delay $E[D]$ as derived in Equation (10). Table 4 shows the mean delay as a function of the total traffic intensity for different values of N. As seen from this table, the figures manifest good agreement between our analysis and the conducted simulations for all traffic intensities. Secondly, the mean delay increases with the traffic intensity Λ. Lastly, we observe that the mean delay displays a linear dependence on the number of stations for the case $\Lambda >> \mu(1-q)$, which follows from Equation (12).

Table 5 shows the standard deviation σ_D of the delay as a function of the total traffic intensity for different values of N. Several conclusions can be drawn from this table. Firstly, the standard deviation of the delay becomes larger as the traffic intensity increases. Secondly, we observe that the standard deviations grow approximately linear with the network size in the case $\Lambda >> \mu(1-q)$. Finally, the table shows that the analytical figures are almost identical to the results obtained by simulation.

Remark 5. The analytical figures in this section have been obtained under the assumption that the service discipline at the service facility is ROS. Obviously, the throughput and the mean delay are the same under any work-conserving non-preemptive service discipline that does not pay attention to the actual service requests of the stations. However, higher moments of the delay strongly depend on the adopted service policy.

$\square$

5. Conclusions and model extensions

In this paper we have developed a finite-source product-form queueing model that incorporates a bursty arrival process as well as the IEEE 802.11 protocol operations. The throughput, the first moment of the delay and a set of equations to compute the LST and higher moments of the delay are derived in this model. Simulation shows that the model is realistic for all network sizes and traffic values. The model can quite naturally be extended with generally distributed service times and message sizes. These extensions will be part of a future paper. An interesting open research topic is the study of the impact of quality-of-service differentiation mechanisms on the performance of a wireless LAN.

Acknowledgement

The authors are grateful to S.P.P. Pronk (Philips Research) and O.J. Boxma (Eindhoven University of Technology) for some valuable comments and stimulating discussions.

REFERENCES

[1] LAN MAN Standards Committee of the IEEE Computer Society (1999). *"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"*. ISO/IEC 8802-11:1999(E); ANSI/IEEE Std 802.11, 1999 Edition, USA.

[2] Anastasi, G., De Stefano, E., Lenzini, L., (2000). *"QoS Provided by the IEEE 802.11 Wireless LAN to Advanced Data Applications: a Simulation Analysis"*. Wireless Networks, vol. 6, no. 2, pp. 99-108, Baltzer.

[3] Bianchi, G., Fratta, L., Oliveri, M., (1996). *"Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs"*. Proceedings PIMRC 1996, Taipei, vol. 2, pp. 392-396.

[4] Bianchi, G., (2000). *"Performance Analysis of the IEEE 802.11 Distributed Coordination Function"*. IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, pp. 535-547.

[5] Boxma, O.J., Denteneer, D., Resing, J.A.C., (2003). *"Delay Models for Contention Trees in Closed Populations"*. Performance Evaluation, Vol. 53, pp. 169-185.

[6] Cali, F., Conti, M., Gregori, E., (1998). *"IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement"*. Proceedings IEEE INFOCOM 1998, San Francisco, vol. 1, pp. 142-149.

[7] Chatzimisios, P., Vitsas, V., Boucouvalas, A.C., (2002). *"Throughput and Delay Analysis of IEEE 802.11 Protocol"*. 5th IEEE International Workshop on Networked Appliances (IWNA5), Liverpool.

[8] Chhaya, H.S., Gupta, S., (1995). *"Throughput and Fairness Properties of Asynchronous Data Transfer Methods in the IEEE 802.11 MAC Protocol"*. Proceedings PIMRC 1995, Toronto, vol. 2, pp. 613-617.

[9] Foh, C.H., Zukerman, M., (2002). *"Performance Analysis of the IEEE 802.11 MAC Protocol"*. Proceedings EW, Florence, pp. 184-190.

[10] Ho, T.S., Chen, K.C., (1996). *"Performance Analysis of IEEE 802.11 CSMA/CA Medium Access Control Protocol"*. Proceedings PIMRC 1996, Taipei, vol. 2, pp. 407-411.

[11] Huang, K.C., Chen, K.C., (1995). *"Interference Analysis of Nonpersistent CSMA with Hidden Terminals in Multicell Wireless Data Networks"*. Proceedings PIMRC 1995, Toronto, vol. 2, pp. 907-911.

[12] Kleinrock, L., (1976). *"Queueing Systems, Volume 2: Computer Applications"*. John Wiley & Sons, New York.

[13] Litjens, R., Roijers, F., Berg, J.L. van den, Boucherie, R.J., Fleuren, M., (2003). *"Performance Analysis of Wireless LANs: an Integrated Packet/Flow Level Approach"*. ITC 18, Berlin.

[14] Sevcik, K.C., Mitrani, I., (1979). *"The Distribution of Queueing Network States at Input and Output Instants"*. In: M. Arato et al. (eds), Performance '79, North-Holland Publishing Company, Amsterdam, pp. 319-335.

[15] Vishnevskii, V.M., Lyakhov, A.I., (2001). *"A Wireless LAN under High Load and in Noise: Throughput Evaluation"*. Automation and Remote Control, vol. 62, no. 8, pp. 1298-1311.

[16] Vishnevskii, V.M., Lyakhov, A.I., (2002). *"IEEE 802.11 Wireless LAN: Saturation Throughput Analysis with Seizing Effect Consideration"*. Cluster Computing 5, pp. 133-144, Kluwer Academic Publishers.

[17] Weinmiller, J., Woesner, H., Ebert, J.P., Wolisz, A., (1996). *"Analyzing and Tuning the Distributed Coordination Function in the IEEE 802.11 DFWMAC Draft Standard"*. MASCOTS'96, San Jose.

[18] Winands, E.M.M., (2003). *"A Finite-Source Feedback Queueing Network as a Model for the IEEE 802.11 Distributed Coordination Function"*. Master's Thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven.

[19] Wu, H., Peng, Y., Long, K., Cheng, S., (2001). *"A Simple Model of IEEE 802.11 Wireless LAN"*. Proceedings ICII 2001, Beijing, vol. 2, pp. 514-519.