A thorough analysis of the performance of delay distribution models for IEEE 802.11 DCF

Qi Wang[a,c,b], Katia Jaffrès-Runser[b,*], Jean-Luc Scharbarg[b], Christian Fraboul[b], Yi Sun[a], Jun Li[a], Zhongcheng Li[a]

[a] *Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China*
[b] *Université de Toulouse, INPT-ENSEEIHT / IRIT, Toulouse, France*
[c] *University of Chinese Academy of Sciences, China*

Abstract

Deriving the complete distribution of the end-to-end delay in a wireless multi-hop network is of paramount interest when delay-sensitive flows have to be conveyed over such networks. First works have proposed models to derive the total delay distribution of networks assuming the well-known IEEE802.11 DCF medium access (MAC) protocol. Current derivations can be decomposed into two main steps: *i*) the calculation of the total delay probability generating function (PGF) and *ii*) its numerical inversion. We show in this paper that there is a need for a thorough performance evaluation of these models since both steps introduce errors, naming modeling and inversion errors. We argue that both types of errors have to be analyzed separately to characterize the accuracy of the analytical derivations of the literature. Therefore, this paper defines two performance evaluation metrics that measure the magnitude of both types of errors. Both metrics are illustrated to select and optimize the most accurate model to calculate the single-hop end-to-end delay distribution of nodes using the IEEE802.11 DCF MAC protocol. The most accurate model is extended to calculate the end-to-end delay distribution for a 2-hop wireless communication.

*Corresponding author. Université de Toulouse, INPT-ENSEEIHT / IRIT, 2 rue Charles Camichel, BP 7122, 31061 Toulouse Cedex 7, France. Tel: +33 5 34 32 21 90

Email addresses: wangqi08@ict.ac.cn (Qi Wang), katia.jaffres-runser@irit.fr (Katia Jaffrès-Runser), jean-luc.scharbarg@irit.fr (Jean-Luc Scharbarg), christian.fraboul@enseeiht.fr (Christian Fraboul), sunyi@ict.ac.cn (Yi Sun), lijun@ict.ac.cn (Jun Li), zcli@ict.ac.cn (Zhongcheng Li)

Preprint submitted to Ad Hoc Networks *August 20, 2014*

Keywords: Performance measure, wireless networks, delay distribution, IEEE 802.11 DCF, saturated traffic

---

1. Introduction

Wireless networks based on the IEEE 802.11 technology [1] are now deployed widely for non-critical applications. The flexibility of wireless connectivity is gaining momentum in the context of real-time networks (wireless industrial fieldbuses, wireless embedded networks, etc..) [2, 3]. The main pitfall of wireless communications is of course the increased unreliability the medium suffers from due to interference and pathloss compared to shielded wires. Moreover, mainstream IEEE 802.11 technology is based on CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance), which is non-deterministic but highly flexible.

Carrying soft real-time data over wireless has been shown to be a feasible option in practice [3]. However, to be able to roll out such a technology, it is necessary to calculate the worst-case end-to-end delay the network offers to the real-time applications using it. If a CSMA/CA type of medium access control (MAC) protocol is considered, a probabilistic definition of the worst-case delay has to be taken into account, which relies on the full knowledge of the delay distribution [4]. The probabilistic worst-case delay can be defined as the delay d_{wc} for which the probability $P(d \geq d_{wc})$ to find a delay larger than d_{wc} is arbitrarily small (e.g. smaller than $\delta = 10^{-9}$ for instance).

The building block of this approach is the precise and accurate knowledge of the delay distribution. This delay distribution, defined more precisely by its probability mass function (PMF), can either be computed by simulations or using analytical models. This paper focuses on the analytical derivation since simulation-based ones do not scale and are too computationally intensive. Only a few works [5, 6, 7, 8] discuss the analytical derivation of the full delay distribution of CSMA/CA networks. Most references on the performance evaluation of IEEE 802.11 [5, 6, 8, 9, 10] mainly focus on the mean delay calculation since it is sufficient for the design of non-critical wireless networks.

The main works that discuss the delay distribution derivation propose different analytical models for the MAC and queuing delay. All these works calculate first the probability generating function (PGF) of the MAC and queuing delays. From these, they deduce the PGF of the total transmission delay which has to be inverted to obtain the total delay distribution. For instance, in [6] and [8], the MAC delay PGF is derived from the well-known Markov model of Bianchi [11]. An important step to get the delay distribution is to invert the PGF to obtain the corresponding probability mass function. This step can introduce errors. Similarly, different inversion methods have been proposed in these works.

Our aim in this paper is to propose a clear and precise performance evaluation method to *i*) assess the quality of the analytical model leading to the total delay PGF, *ii*) select the most accurate numerical inversion method. Therefore, we define two performance measures whose aims are to characterize the error originating from the analytical model on the one side and from the PGF inversion method on the other side. Computing the distribution is complex, and previous works have assessed the performance of their model only by comparing analytical distributions to their simulated counterpart. However, they have neglected the numerical inversion error. We argue in this paper that to trust the models, it is necessary to discriminate the impact of both errors on the final distribution. Unique to this work is to provide metrics to differentiate both types of error, which are then used to select the most accurate models and inversion errors of the litterature.

Our performance evaluation method is illustrated on the specific case of an IEEE 802.11 DCF medium access where two different types of queues are assumed, naming M/M/1 and M/G/1. Our main conclusions show that a pretty accurate model for the MAC delay is available while improvements are needed for the queuing delay distribution derivation.

The paper opens up on examining the feasibility of extending the single-hop theoretical delay distribution derivation to dual-hop communications assuming simple M/M/1 queues at the source and the relay nodes. We first show that the main assumptions of the single-hop case do not hold anymore, triggering an increased error for the total delay distribution. However, interestingly, the analytical model extension for the dual-hop case provides a good approximation of the distribution tail, which is the part required to calculate a probabilistic worst case delay.

This paper is organized as follows. In Section 2, the overall analytical derivation of the total delay distribution is presented. Detailed calculations for the individual MAC, queuing and total delays for IEEE 802.11 DCF protocol and specific queues are given in Section 3. In Section 4, we introduce the method proposed to assess the performance of a delay distribution model. This method is leveraged in Section 5 to select the most accurate analytical model that derives the total delay distribution of an IEEE 802.11 wireless network using DCF. We extend the total delay derivation to a 2-hop communication scenario in Section 6. Finally, Section 7 concludes the paper.

2. Derivation of the total delay distribution

This section starts by introducing the wireless system of interest and then presents the overall analytical derivation of the total delay distribution.

2.1. System model

In this paper, a source node is directly transmitting its packets to a destination node. These two nodes belong to a set of n stationary nodes sharing a common wireless medium. Each emitted packet experiences a *total transmission delay* d_t which is measured from its time of generation to the time its sender gets an acknowledgement (ACK) from the destination node or a maximum number of transmission trials has been reached.

At the time of generation, the emitted packet enters the transmission queue. Once it has reached the head of its queue, it will compete for channel access with the other stations. If the packet has been emitted, the sender waits for a positive ACK from the destination. Thus, the total delay is the sum of: *i)* a *queueing delay*, which is the time for the packet to reach the head of the transmission queue, and *ii)* a *medium access delay (MAC delay)*, which is the time needed by the medium access protocol to either successfully deliver the packet or drop it in case of repeated failures.

The rest of the paper recalls the analytical derivation of the distribution of the total delay experienced by packets for an IEEE 802.11 DCF medium access (with or without RTS/CTS mechanism). We consider a saturated traffic where all nodes of the network always have a packet ready for transmission. Two types of queues are investigated as well, naming M/M/1 and M/G/1. Ideal channel conditions are assumed as well (no channel errors, no hidden terminals).

2.2. Global analytical modeling

MAC and queuing delays can be assumed as independent discrete random variables as shown in [8]. Indeed, the MAC delay experienced by a head of line packet is completely independent from the time it has spent in the queue. It is just a function of the number of nodes contenting for medium access with him.

In the rest of the paper, the following notation is adopted: $d_t(k)$, $d_q(k)$ and $d_m(k)$ represent the probability mass functions (PMF) of the total transmission, queuing and MAC delays, respectively. $D_t(Z)$, $D_q(Z)$ and $D_m(Z)$ are the probability generating functions (PGF) of total, queuing and MAC delay, respectively. We recall that the probability-generating function of a discrete random variable X is the Z-transform of its PMF. It is calculated following $D(Z) = \sum_{k=0}^{\infty} d(k)Z^k$, with $Z \in \mathbb{C}$ and $d(k)$ the PMF of X.

Since MAC and queueing delay random variables are independent, the PGF of the total delay $D_t(Z)$ is equal to the product of $D_m(Z)$ with $D_q(Z)$:

$$D_t(Z) = D_m(Z)D_q(Z) \tag{1}$$

The mean of the total delay $E[D_t] = E[D_m] + E[D_q]$ is obtained by summing the mean MAC and queuing delays with $E[D] = D'(Z)|_{Z=1}$.

In this work we are interested in extracting the probability mass function $d_t(k)$ of the total delay. Therefore, we will need first to derive $D_m(Z)$ and $D_q(Z)$, the PGF of MAC and queuing delays respectively. Previous works have tackled these problems with different perspectives and models. Our aim in this paper is to present a performance evaluation analysis of these derivations to select the ones which provide the best trade-off between accuracy and complexity.

Having $D_t(Z) = \sum_{k=0}^{\infty} d_t(k)Z^k$, $d_t(k)$ is obtained by the Z-transform inversion of the PGF. This last inversion step is critical and can introduce errors. The error introduced by the inversion is added to the error an imperfect analytical model can create. We argue in this paper that to have a clear view of the performance of a given analytical derivation of a delay distribution, its validation has to be done in two steps. First, the model used to derive the individual PGFs has to be validated before numerical inversion. Second, the numerical inversion has to be tailored to reduce the inversion error.

3. PGF derivations of MAC and queueing delays

This section recalls briefly the derivation of the individual PGFs for the MAC and queueing delays we have selected from the literature.

3.1. PGF of MAC delay

Two different types of models have been proposed in the literature to characterize the MAC delay distribution. The models of Zhai et al. [8] and Vardakas et al. [6] rely on the well-known Markov chain originating from the work of Bianchi [11]. Vu and Sakurai have given in [7] the main lines of a different probabilistic derivation for the MAC PGF, together with a very limited performance evaluation. In the rest of this paper, we will focus mainly on the PGF derivation of [6] and [8] that builds on the Markov chain model of [11].

For conciseness purposes, we refer the reader to [1], [6] and [8] for a detailed description of the Markov chain representing the IEEE 802.11 DCF MAC protocol. From this Markov chain, at steady state, the probability τ that a node transmits in a randomly selected time slot is extracted. In the following, the derivation of the MAC PGF given by Vardakas et al. in [6] is summarized. The saturated state are considered, where each of n nodes always has a packet to transmit in its transmission queue.

The interruption of the backoff period is a result of two different events: the collision of two or more nodes with probability p and the transmission of only one node other than the tagged one, with probability

$$p' = \binom{n-1}{1}.\tau.(1-\tau)^{n-2} \tag{2}$$

Following [6], the binary exponential backoff algorithm can be envisioned as a function of two coordinates (x, y), where $(x \in [0, m])$ is the backoff stage and $(y \in [0, W_x - 1])$ is the value of the backoff counter at the backoff stage x. The authors of [6] deduce that the PGF of the duration a packet stays in stage x with backoff counter y is given by:

$$B_{x,y}(Z) = \frac{(1-p).Z^{\sigma}}{1-(p'S(Z)+(p-p')C(Z))} \tag{3}$$

where Z^{σ} is the PGF of the propagation time σ, $S(Z) = Z^{T_s}$ and $C(Z) = Z^{T_c}$ are the PGFs of the duration of a successful transmission period T_s and of a collision period T_c, respectively. They depend on the type of service (basic or RTS/CTS) and their derivation can be found in [6]. Main DCF timing values considered in this paper can be found in [12].

The PGF of the duration the packet stays in the backoff stage x follows:

$$B_x(Z) = \begin{cases} \sum_{y=0}^{W_x-1} \frac{B_{x,y}(Z)^y}{W_x}, 0 \leq x \leq m' \\ B_{m'}(Z), m' < x \leq m. \end{cases} \tag{4}$$

From this, the PGF of the MAC delay is derived as:

$$\begin{aligned} D_m(Z) \quad &= (1-p).S(Z).\sum_{x=0}^{m}[(p.C(Z))^x \prod_{i=0}^{x} B_i(Z)] + \\ &(p.C(Z))^{m+1}.\prod_{i=0}^{m} B_i(Z) \end{aligned} \tag{5}$$

It represents the duration for the packet to reach the end state (i.e. being transmitted successfully or discarded after maximum m retransmission failures) from the start state (i.e. beginning to be served). The first term relates to the delay of a successfully transmission including the delay spent in the previous x and y backoff stages, while the second term calculates the delay for dropping the packet after m trials.

Mean MAC delay $E[D_m]$ is given by the first derivative of $D_m(Z)$ at $Z = 1$: $E[D_m] = D'_m(Z)|_{Z=1}$

3.2. PGF of queueing delay

This section presents the derivation of the queuing delay PGF, $D_q(Z), Z \in \mathbb{C}$ for both M/M/1 and M/G/1 queues. Packets enter the queue according to a Poisson distribution of rate λ. The packet transmission process introduced by the DCF medium access can be modeled as a general single server whose service time distribution is known from Eq. (5).

3.2.1. Assuming an M/M/1 queue

For the M/M/1 queue, the service time is exponentially distributed with parameter μ. Thus, the cumulative distribution function (CDF) and probability density function (PDF) of the service delay are $F(t) = 1 - e^{\mu t}$ and $f(t) = \mu e^{-\mu t}$, respectively. The service times have an average value of μ equal to the mean MAC delay: $\mu^{-1} = E[D_m]$. The Laplace transform of F is the function $L_f(s) = \frac{\mu}{s+\mu}$ [13]. According to the Pollaczek-Khintchine (P-K) transform equation, the Laplace transform $L_{D_q}(s)$ of the queueing delay can be expressed as:

$$L_{D_q}(s) = \frac{s(1-\rho)}{s-\lambda+\lambda L_f(s)} \tag{6}$$

with $\rho = \lambda/\mu$ the server utilization. According to the relationship between Laplace and Z-transform [13] (*cf.* Appendix in [12]), it is possible to deduce the Z-transform $D_q(Z)$ from $L_{D_q}(s)$ by substituting $s = -\ln Z$ into (6):

$$D_q(Z) = \frac{-ln(Z)(1-\rho)}{-ln(Z)-\lambda+\lambda L_f(-ln(Z))} \tag{7}$$

We can derive the PMF $d_q(k)$ by inverting $D_q(Z)$.

3.2.2. Assuming an M/G/1 queue

The M/G/1 queue is a single-server system with Poisson arrivals and arbitrary service-time distribution.

Similarly, Laplace transform of the queueing delay gives $L_{D_q}(s) = \frac{s(1-\rho)}{s-\lambda(1-L_f(s))}$, where $L_f(s)$ is the Laplace transform of the service time distribution function. The Z-transform of the queueing delay $D_q(Z)$ is derived as in [5]:

$$D_q(Z) = \frac{(1-Z)(1-\rho)}{1-Z-\lambda(1-D_m(Z))} \tag{8}$$

Similarly to the M/M/1 case, the PMF $d_q(k)$ can be derived by inverting $D_q(Z)$.

3.3. PGF of total delay

The PGF of the total delay is computed by multiplying the PGF of queuing and MAC delay as given in (1). We investigate two different models, a very simple and a more accurate one. The first one assumes an M/M/1 queue and a Markovian MAC delay distribution as well. The second one assumes an M/G/1 queue and a MAC delay that follows the $D_m(Z)$ PGF of (5). This last model is very heavy to compute compared to the simpler M/M/1 one.

3.3.1. Assuming an M/M/1 queue

In this total delay derivation, we assume that the service times are exponentially distributed with an average μ^{-1} equal to the mean MAC delay deduced from (4). We do not use the MAC delay distribution of (5), but assume that the packets are served by the MAC with an exponential distribution of PDF $f(t) = \mu e^{-\mu t}$ of mean MAC delay $\mu^{-1} = E[D_m]$. The corresponding PGF of the exponential MAC delay is given by:

$$D_m(Z) = \frac{\mu}{\mu - \ln(Z)} \tag{9}$$

This assumption may of course introduce errors but its derivation is much simpler. The point of this paper is to state whether the loss due to this approximation is reasonable or not compared to a precise (and complex) M/G/1 formulation and complete MAC delay derivation.

Similarly to (1), the Laplace transform of the total delay $L_{D_t}(s)$ is computed as the produce of the Laplace transforms of the queuing and MAC delay PDFs.

$$L_{D_t}(s) = L_f(s) L_{D_q}(s) = L_f(s) \frac{s(1-\rho)}{s - \lambda + \lambda L_f(s)} \tag{10}$$

From $L_f(s) = \mu/(s + \mu)$ and (10), $L_{D_t}(s)$ is given by

$$L_{D_t}(s) = \frac{\mu(1-\rho)}{s + \mu - \lambda} = \frac{\mu - \lambda}{s + \mu - \lambda} \tag{11}$$

The Z-transform of the total transmission delay $D_t(Z)$ can be expressed as

$$D_t(Z) = L_{D_t}(-\ln Z) = \frac{\mu - \lambda}{-\ln Z + \mu - \lambda} \tag{12}$$

The mean queueing delay $E[D_q]$ for M/M/1 queue is computed using Little's law as $\rho/(\mu - \lambda)$ and the corresponding mean total delay $E[D_t]$ as $1/(\mu - \lambda)$.

3.3.2. Assuming an M/G/1 queue

In this case, the service times of the queue are distributed according to the MAC delay distribution given by the PGF $D_t(Z)$ in Eq. (5). Following (1), the Z-transform of the total delay $D_t(Z)$ follows:

$$D_t(Z) = D_m(Z)D_q(Z) = \tfrac{D_m(Z)(1-Z)(1-\rho)}{1-Z-\lambda(1-D_m(Z))} \tag{13}$$

The mean queueing delay $E[D_q]$ for M/G/1 is derived by the Pollaczek-Khinchin mean value formula [14] [Kleinrock 1975 (sec 5.6)], given through the second moment of D_m: $E[D_q] = \frac{\lambda E[D_m^2]}{2(1-\rho)}$. $E[D_m^2]$ is given by $E[D_m^2] = var(D_m)+(E[D_m])^2$ and $var(D_m)$ by $var(D_m) = D''_m(Z)|_{Z=1} + D'_m(Z)|_{Z=1} - (D'_m(Z)|_{Z=1})^2$.

4. Evaluating the accuracy of a delay distribution model

This section proposes a performance evaluation measure to characterize the accuracy of a given delay distribution model. As presented in Section 2, it is very convenient to express the delay distribution as a PGF. Thus, to obtain the PMF values $p(k)$, the corresponding PGF $D(Z)$ has to be inverted. This numerical inversion introduces errors. Thus, we argue that directly comparing the final $p(k)$ with the PMF obtained by simulations $p^s(k)$ is not appropriate to validate the quality of the analytical model. It can not discriminate the error originating from the model itself from the numerical inversion error. In other words, it is not possible with such a comparison, as done in [8], to know whether the errors between the simulated and analytical PMF come from an inaccurate model or originate from the inversion of $D(Z)$.

We show in this paper that to have a clear view of the performance of a given analytical derivation of a delay distribution, its validation has to be done in two steps. First, the model used to derive the individual PGFs has to be validated before numerical inversion. Second, the numerical inversion has to be tailored to reduce the inversion error.

Indeed, even though the models proposed in previous works [6][7][8] are very interesting, they suffer from a limited or not convincing performance evaluation of the delay distribution. More specifically, the MAC delay distribution models of [6] and [7] show little results in their papers. Vardakas et al.[6] mostly validate the *average* MAC delay against simulations but don't give results for the total distribution. Vu and Sakurai [7] present a single figure to validate their model against simulations and [8]'s results. Zhai et al. [8] provide PMF results for several cases, but they directly compare the PMF to the simulated distribution, completely ignoring the fact that errors can originate from the numerical inversion of the PGF. There is clearly a need for a clean performance evaluation measure capable of assessing the quality of a delay distribution model.

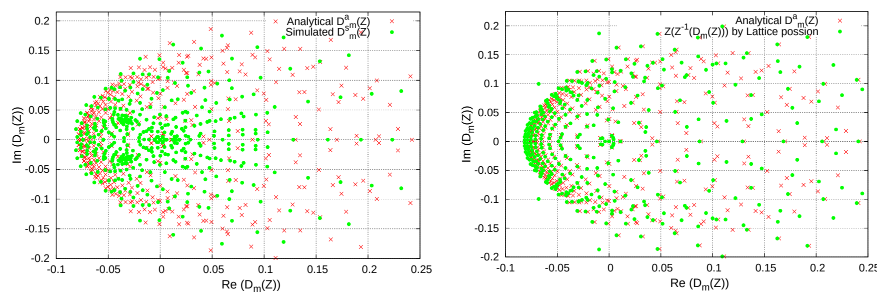
In the following, we describe first a performance measure to assess the analytical model's accuracy. Then, we give a performance measure to calculate the error introduced by the PGF numerical inversion.

4.1. Performance measure for the analytical model quality

From now on, we will denote the PMF (resp. PGF) values obtained by simulation using $d^s(k)$ (resp. $D^s(Z)$) and the ones obtained analytically using $d^a(k)$ (resp. $D^a(Z)$).

It is straightforward to calculate the PMF values $d^s(k)$ from the statistics of the delay obtained by simulation. Thus, to avoid the inversion of the analytical PGF for its performance evaluation, we propose to compare directly the analytical PGF values $D^a(Z), Z \in \mathbb{C}$ to the PGF $D^s(Z)$ derived from the simulated PMF. The value of the PGF $D^s(Z)$ for any complex $Z \in \mathbb{C}$ is given by the Z-transform of the $d(k)$: $D^s(Z) = \sum_{k=0}^{\infty} d^s(k) Z^k$. This calculation doesn't introduce any errors. Thus for a same set of complex values $\mathcal{C} \subseteq \mathbb{C}$, it is possible to calculate the analytical PGF $D^a(Z), Z \in \mathcal{C}$ and its simulated counterpart $D^s(Z), Z \in \mathcal{C}$. Figure 1(a) illustrates both analytical and simulated complex sets ($D^a_m(Z)$ and $D^s_m(Z)$) in a real and imaginary plot obtained for the MAC delay. Analytical PGF is derived following Eq. (5). The set of complex values used to calculate $D^a_m(Z)$ and $D^s_m(Z)$ is here defined as $\mathcal{C} = \{re^{-i\pi h/k}\}$ where $r = 10^{-4/k}$, k varies from 1 to 50 with step 5 and h varies from $-k$ to $+k$ with step 1. This set samples the complex unit circle with an accuracy of 10^{-8}.

For a perfect analytical model, the points calculated for $D^a(Z)$ would ex-



(a) MAC Modeling error: $D_m^a(Z)$ vs. $D_m^s(Z)$ for $n = 5$

(b) Inversion error: $D_m(Z)$ for $n = 5$ and $\gamma = 10^{-6}$

Figure 1: Model and inversion error in the complex space.

actly match the ones obtained by simulation (providing that the simulation is extensive enough). From Figure 1(a), it is clear that there is an error between analytical and simulated values. Therefore we propose to quantify this error by defining a normalized root mean squared error (NRMSE) criterion as:

$$f_{model} = \frac{1}{\text{Card}(\mathcal{C})} \sum_{Z \in \mathcal{C}} \sqrt{\frac{|D^s(Z) - D^a(Z)|^2}{|D^s(Z)|^2}} \tag{14}$$

4.2. Performance measure for PGF inversion quality

By definition, a perfect PGF inversion is characterized by:

$$\mathcal{Z}\{\mathcal{Z}^{-1}\{D(Z), Z \in \mathbb{C}\}\} \equiv \{D(Z), Z \in \mathbb{C}\}$$

where $D(Z)$ is the PGF of a delay distribution, $\mathcal{Z} : Z \in \mathbb{C} \to \sum_{k=0}^{\infty} d^a(k) Z^k$ is the Z-transform function and $\mathcal{Z}^{-1} : D(Z), Z \in \mathbb{C} \to d(k)$ is the inverse Z-transform function.

In the following, the PMF obtained after inversion is denoted $\{\hat{d}(k), k \in \mathbb{N}\}$. Thus, for a perfect inversion, there is a perfect match between the original PGF values $D(Z)$ and the Z-transform of the PMF $\hat{d}(k) = \mathcal{Z}^{-1}\{D(Z), Z \in \mathbb{C}\}, \forall k \in \mathbb{N}$. A non perfect inversion yields a difference between the two obtained complex sets. This is illustrated on Figure 1(b) for the MAC delay PGF calculated for $n = 5$. The same complex set $\mathcal{C}$ is used to plot the complex PGF values of Figure 1(a) and 1(b).

To assess the quality of a PGF inversion method, we propose to simply calculate, for each Z in a complex set $\mathcal{C} \subseteq \mathbb{C}$, the NRMSE between the original PGF and the Z-transform of the delay PMF obtained by inversion, naming $\hat{d}(k) = \mathcal{Z}^{-1}\{D(Z), Z \in \mathbb{C}\}, \forall k \in \mathbb{N}$. Formally, our performance measure is:

$$f_{inv} = \frac{1}{\mathrm{Card}(\mathcal{C})} \sum_{Z \in \mathcal{C}} \sqrt{\frac{|D(Z) - \mathcal{Z}\{\mathcal{Z}^{-1}\{D(Z)\}|^2}{|D(Z)|^2}} \tag{15}$$

5. Assessing the performance for IEEE 802.11 DCF

This section exploits first the analytical performance metric f_{model} to assess the quality of the models used to derive the individual MAC delay PGF and the queuing delay PGF introduced in Section 3. Second, PGF inversion metric f_{inv} is used to select and fine tune the numerical inversion method. Finally, the final total delay PMF is shown based on the choices made in terms of modeling and numerical inversion.

5.1. Simulation settings

A thorough simulation of the transmission delay is needed to compute the analytical performance metric. This section introduces the main simulation settings. The wireless network, composed of n nodes and one sink, is simulated using the discrete event-driven network simulator WSNet[1]. Presented results for the IEEE 802.11 DCF MAC delay are given with RTS/CTS mechanism. As no transmission errors are assumed, the distance between the nodes is short (maximum distance between a node and the destination is of about 7 meters), with a maximum range of 743 meters assuming a 2-way ground propagation model. In this case, RTS/CTS mechanism is compulsory to avoid hidden node problems. Indeed, in [15], Xu et al. show that if the distance between the source and the destination is above 0.56* R_{tx}, with R_{tx} the transmission range of the source, the RTS/CTS effectiveness drops rapidly. In all our simulations,

[1] http://wsnet.gforge.inria.fr/

Table 1: Simulated mean MAC delay with respect to the simulation duration ($n = 5$).

Simulation duration	Mean MAC delay (ms)	95% Confidence interval	Number of transmissions
12 hours	12.8189	[-35.9969, 61.6147]	$\sim$ 3 376 000
24 hours	12.8206	[-36.1339, 61.7750]	$\sim$ 6 742 000
30 hours	12.8202	[-36.1319, 61.7722]	$\sim$ 8 425 000
48 hours	12.8172	[-36.1009, 61.7355]	$\sim$ 13 473 604

the source to destination distance is much shorter than 0.56* R_{tx}, and thus it RTS/CTS mechanism has to be accounted for.

The DSSS-PHY layer is assumed with a data rate of 11Mbps, using packets of a constant size equal to 1400 bytes. Propagation delay δ is set to $1\mu s$. The queues implemented in the simulator have no limited size. Simulations have been conducted for 5 days. We have extracted the results of the first 24 hours, experimenting the transmission of $\sim$6 742 000 packets, since the 24-hour delay distribution was identical to the one calculated over longer periods. For conciseness purposes, we do not show the evolution of the distributions but just represent the average MAC delay for different durations in Table 1.

All nodes experience the same Poisson arrival rate λ. Since the MAC model assumes saturated conditions, λ should be chosen such as to satisfy the condition of a non-empty transmission queue. Therefore, utilization of the queue ρ should be more than 95%. And to satisfy the P-K transform equation condition, ρ has to be lower than 1 (i.e. $\lambda < \mu$). Since we have set $\mu^{-1} = E[D_m]$, values for λ are calculated for each network size using $\lambda = 0.95 * 1/E[D_m]$. Arrival rates for network sizes of $n \in \{5, 15, 30\}$ are given in Table 2.

Table 2: Queue parameters to reach saturation

Number of nodes	λ ($packet/ms$)	μ ($packet/ms$)
$n = 5$	0.07799	1/12.1808
$n = 15$	0.02665	1/36.4052
$n = 30$	0.01359	1/71.3596

5.2. Assessing the performance of the MAC models

This section illustrates our analytical performance measure f_{model} on the following two models of IEEE 802.11 DCF MAC:

1- the Markov chain based MAC PGF of [6], [8] of Eq (5).

2- the simple exponential MAC PGF of Eq. (9).

The PGF of the first model is complex to evaluate while the second one is very light, since it simply necessitates the derivation of the mean MAC delay.

Table 3: $E[D_m^a]$, $E[D_m^s]$, $\Delta = |E[D_m^a] - E[D_m^s]|$ (ms) and Confidence interval (95%)

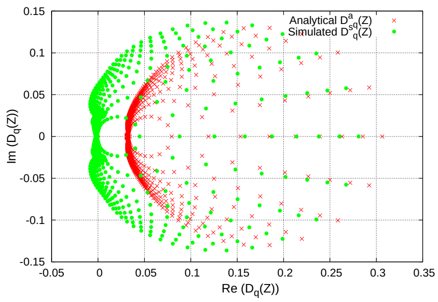
Number of nodes	$E[D_m^a]$	$E[D_m^s]$	Δ	Confidence interval
$n = 5$	12.1808	12.1123	0.0685	[-29.2188, 53.4434]
$n = 15$	36.4052	36.4096	0.0044	[-137.3915, 210.2106]
$n = 30$	71.3596	71.1140	0.2456	[-263.0347, 405.2628]

Most of the works on DCF modeling have been validated by comparing the mean MAC delay to the one obtained by simulations. Table 3 gives the analytical and simulated mean MAC delays obtained for different network sizes. By definition, the mean delay of the simple exponential MAC is set to the mean delay of the Markov MAC model (referred as $E[D_m^a]$). Simulated and analytical mean MAC delays are really close as shown in Table 3. However, looking at Figure 1(a), given in p.12, the analytical and simulated values of $D_m(Z)$ do not coincide for the Markov MAC model. Just comparing the mean MAC delay is not convincing, which calls for a more precise performance evaluation.

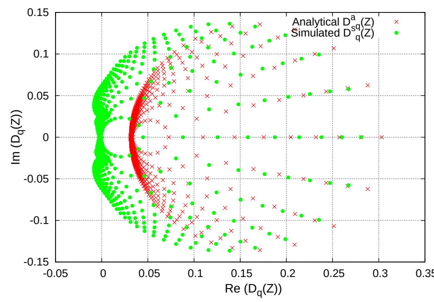
Table 4: Comparison of MAC distributions

Number of nodes	$n = 5$	$n = 15$	$n = 30$
f_{model} for Markov model	0.0547	0.0789	0.0729
f_{model} for Exponential MAC	0.1736	0.1258	0.1040

In Table 4, f_{model} quantifies the error induced by both MAC models (before inversion) compared to simulations. Not surprisingly, the Markov MAC model outperforms the simple exponential MAC. However, an interesting observation



(a) $M/G/1$



(b) $M/M/1$

Figure 2: Comparison of $D_q^a(Z)$ and $D_q^s(Z)$ for $M/G/1$ and $M/M/1$ queue with $n = 5$.

is that the simple exponential MAC becomes better as the number of nodes in the network increases.

5.3. Assessing the performance of the queueing models

This section assesses the queueing delay models (M/M/1 queue of Eq (7) and M/G/1 queue of Eq (8)) using f_{model}.

For the M/G/1 queue, $f_{model} = 0.03387$ and for the M/M/1 queue, $f_{model} = 0.10515$. Not surprisingly, M/G/1 outperforms the M/M/1 queue, but at the price of a much more intensive computation load. To illustrate the error, Figure 2 plots the analytical and simulated PGF values for M/G/1 (Figure 2(a)) and M/M/1 (Figure 2(b)) queues. There is still a noticeable error between these infinite queuing models and the simulated values. Thus, there is still room for improvement in the selection of the queuing model.

5.4. Assessing the performance of PGF inversion methods

In the first part of this section, the proposed PGF inversion performance metric is leveraged to select the most efficient numerical inversion method. In the second part, it is used to parameterize the numerical inversion method selected previously for both MAC and queuing models.

5.4.1. Considered PGF inversion methods

Inverting the probability generating function can be done by repeatedly differentiating and evaluating it at $Z = 0$: $d(k) = \left.\frac{D^{(k)}(Z)}{k!}\right|_{Z=0}$. This type of inversion has been done by Zhai et al. in [8] using numerical differentiation techniques and symbolic mathematical software. However, it is often difficult to achieve desired accuracy with numerical differentiation techniques, especially for large k_{max} (k_{max} being the number of PMF values obtained after inversion). It is also difficult to invoke symbolic mathematical software when the generating function is only expressed implicitly. Fortunately, in our setting, numerical inversion is a viable alternative that has been chosen by Vardakas et al. [6] and Vu and Sakurai [7]. The numerical inversion of a PGF is based on the Lattice-Poisson (LP) algorithm [16]. Two different derivations of the LP algorithm have been proposed to numerically invert a delay PGF in [6] and [7], respectively.

The LP inversion formula of Vardakas et al. [6] is:

$$d(k) \approx \tfrac{1}{2kr^k} \textstyle\sum_{j=1}^{2k} (-1)^j Re[D(re^{i\pi j/k})] \tag{16}$$

with $Re[D(Z)]$ the real part of the complex $D(Z)$. $d(k)$ is derived by summing $Re[D_m(Z)]$ over a circle of radius $r = 10^{-\gamma/(2k)}$ for an accuracy of $10^{-\gamma}$.

The LP inversion formula of Vu et Sakurai [7] is:

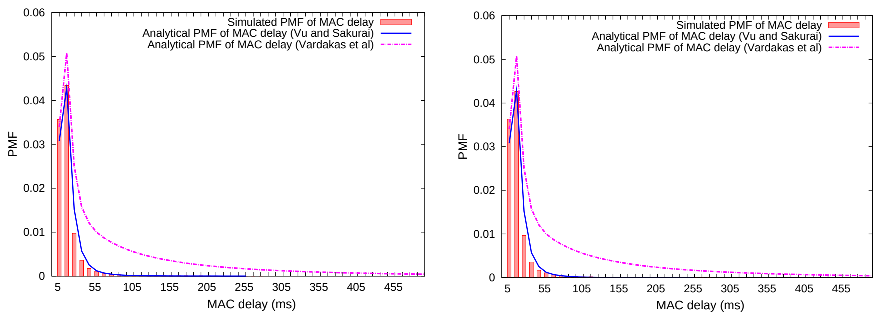
$$d(k) \approx \tfrac{1}{2klr^k} Re\left[\textstyle\sum_{j=-kl}^{kl-1} D(re^{-i\pi j/(kl)}) e^{i\pi j/l}\right] \tag{17}$$

where, $l = 1$ and $r = 10^{-\gamma/(2k)}$, which results in an accuracy of $10^{-\gamma}$ as well. Both formulas are almost equivalent. The only difference is how the real part is calculated. In Eq. (16), only the real part of $D(Z)$ is considered, while in Eq. (17), the real part of the whole sum is returned.

Table 5: Evaluation of LP algorithms using f_{inv} on the MAC delay PGF for $n = 5$.

Accuracy	Vu and Sakurai	Vardakas et al.
10^{-4}	0.0232	0.1814
10^{-6}	0.0195	0.0688

Using f_{inv}, it is possible to figure out which inversion formula is best. It is the one with the smallest value of f_{inv}. The values are given in Table 5 for an accuracy of 10^{-4} and 10^{-6}. It is the LP formula of Vu and Sakurai which introduces the less error for both accuracies. This fact can be observed



(a) Simulated PMF of MAC delay for 24 hours simulation duration

(b) Simulated PMF of MAC delay for 48 hours simulation duration

Figure 3: MAC delay PMF for different inversion methods (accuracy 10^{-6}, $n = 5$).

as well on Figure 3(a) and Figure 3(b), which presents the MAC delay PMF after numerical inversion.

It makes sense that Eq. (17) is correct since to inverse the PGF mathematically, a contour integral has to be computed for the complex values $D(Z)Z^{n-1}$. The real $d(k)$ values of this integral are obtained by taking the real part of the integration result and not by integrating $Re[D(Z)]Z^{n-1}$. As a consequence, we apply the LP formula of Vu and Sakurai for all results shown in the rest of the paper.

5.4.2. Tuning the LP algorithm for MAC and queuing models

It can be seen in the results of Table 5 that the accuracy with which the LP algorithm is computed directly influences the inversion error. If the accuracy is improved from 10^{-4} to 10^{-6}, the error measured with f_{inv} reduces from 0.0235 to 0.0195 for Vu and Sakurai's LP algorithm. These values have been calculated for the MAC delay PGF. Figure 4 plots the MAC delay PMF obtained after inversion with the two accuracies of 10^{-4} and 10^{-6}, for three different network sizes. The impact of the improved accuracy is clearly visible on these figures. The impact is the highest for $n = 5$.

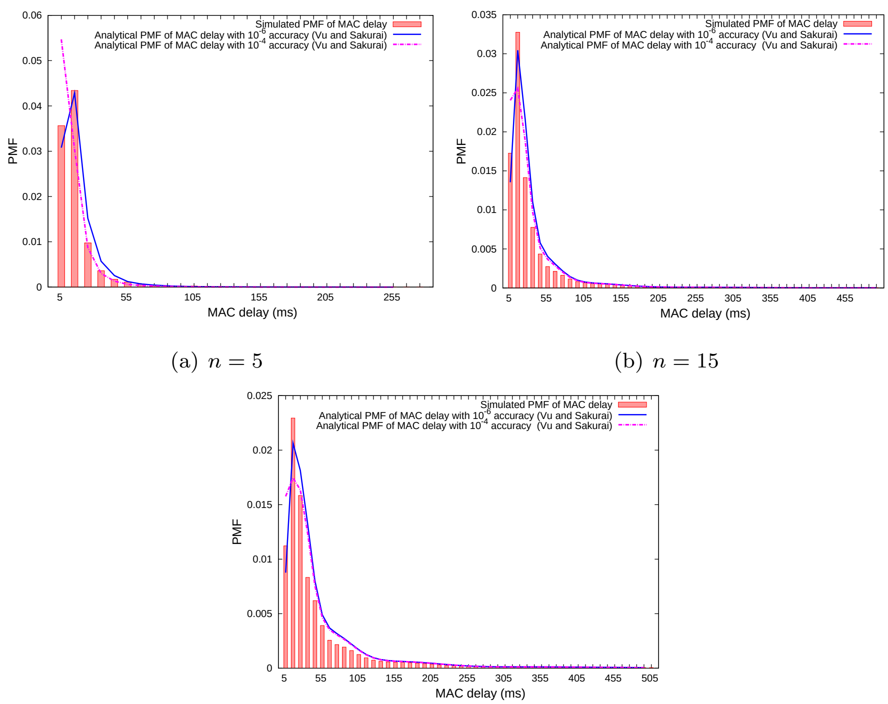
Using f_{inv} and f_{model}, the MAC delay distributions shown on Figure 4 present the best possible fits we have obtained with the models investigated.

It has been obtained using the Markov MAC model of Eq (5) and the Lattice Poisson algorithm of Vu and Sakurai for an accuracy of 10^{-6}.

Table 6: Impact of the accuracy of LP algorithm for the queuing delay using f_{inv} metric

Accuracy	M/G/1	M/M/1
10^{-6}	0.01477	0.01482
10^{-8}	0.007582	0.009189

A similar analysis has been performed using f_{inv} to determine the best accuracy for the inversion of the queuing delay PGF. Results are shown in Table 6 for both M/M/1 and M/G/1 queues. A good improvement is obtained by adopting an accuracy of 10^{-8} for the LP algorithm.

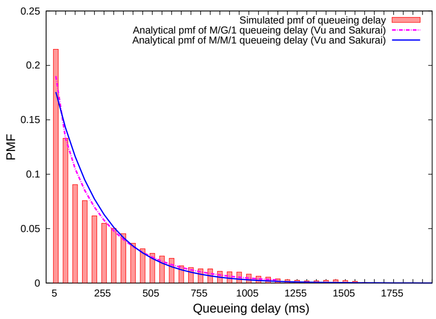


(a) $n = 5$

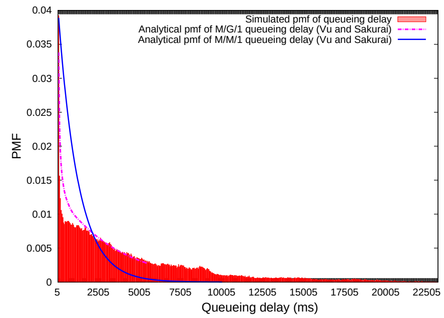
(b) $n = 15$

(c) $n = 30$

Figure 4: MAC delay PMF for different accuracies using Vu and Sakurai's LP formula.

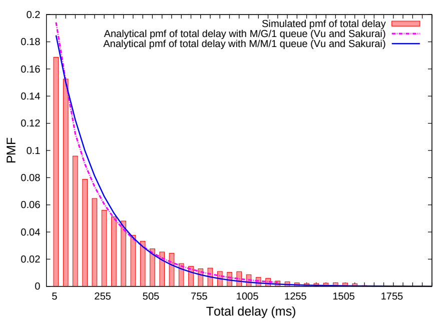


(a) $n = 5$

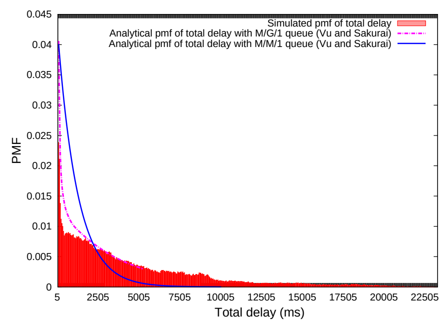


(b) $n = 15$

Figure 5: Analytical queueing delay PMFs for M/M/1 and M/G/1 queues vs. simulations.



(a) $n = 5$



(b) $n = 15$

Figure 6: Analytical total delay PMFs for M/M/1 and M/G/1 queues vs. simulations.

5.5. Final queuing and total delay distributions

Queueing and total delay distributions for M/M/1 and M/G/1 are presented in Figure 5 and Figure 6, respectively. They are given for $n = 5$ and 15.

The total delay is clearly dominated by the queueing delay. This is not surprising since we are working at a very high utilization ($\rho > 95\%$) to reach saturated conditions. For the queuing delay, the NRMSE calculated using the k_{max} PMF values is of 0.08519 ($n = 5$) and 0.10366 ($n = 15$) for the M/M/1 model and of 0.05057 ($n = 5$) and 0.01879 ($n = 15$) for the more precise M/G/1 model. It can be concluded that the M/G/1 model clearly better matches the simulated queuing delay, but it can be noticed that the error with M/M/1 stays limited for $n = 5$.

For the total delay, the NRMSE calculated using the k_{max} PMF values is of 0.08318 ($n = 5$) and 0.10455 ($n = 15$) for the M/M/1 model and of 0.05067 ($n = 5$) and 0.02061 ($n = 15$) for the more precise M/G/1 model. However, looking at the total delay, M/M/1 seems to be a good compromise between accuracy and complexity for small networks. In Table 4, it is clear using f_{model} that the MAC model of M/M/1 is less efficient than the Markov based model. We can conclude that the Markov MAC model with the M/G/1 queue is the most efficient one, but as well the most computationally demanding.

6. First extension to a 2-hop communication

The purpose of this section is to illustrate our performance evaluation method on a simple extension of the previous computation to handle a 2-hop communication. Several important works have discussed the delay performance of wireless multi-hop networks [17][10][18][19][20]. All these works investigate the average end-to-end delay using various models and assumptions. To the best of our knowledge, there are no previous works discussing the calculation of the complete end-to-end delay distribution for a multi-hop wireless network.

In this section, we still assume ideal channel conditions (no channel errors, no hidden terminals). Nodes work in saturated mode. The considered 2-hop communication is depicted on Figure 7. Packets emitted by the source node S can't reach directly the destination D. A node R relays all packets received from S and re-emits them to the destination on the fly. R only relays packets. Similarly to the single-hop network, there is a total of n nodes transmitting

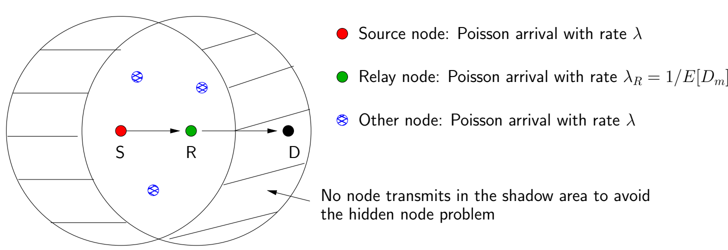


Figure 7: 2-hop communication configuration

packets concurrently, S and R being included in this set. All nodes but R have the same arrival rate of λ. The arrival rate of R is given by the rate at which it receives packets successfully from the source, which is given by $\lambda_R = 1/E[D_m]$, with $E[D_m]$ the mean MAC delay of the nodes.

6.1. Analytical PGF derivation

The aim of this section is to present a simple analytical model to retrieve the 2-hop total delay distribution. The 2-hop total delay is the time between the date the packet enters the queue of the source S and the date it is received at the destination. The packets that are lost are not accounted for in this first model. The derivation we propose builds on the 1-hop delay analysis detailed earlier. We calculate the analytical PGF of the 2-hop total delay distribution and invert it using the Lattice-Poisson algorithm selected in Section 5.4. To calculate the analytical PGF, we make the following two assumptions:

(1) The total delay (i.e. queuing plus MAC delay) a packet experiences on the first hop from S to R is independent from the total delay it experiences on the second hop from R to D.

(2) Packet arrival process at the relay follows a Poisson distribution with an arrival rate close to $1/E[D_m]$.

Knowing that the PGF of the sum of independent random variables is equal to the product of the PGF of each variable, the analytical PGF of the 2-hop total delay $D_t(Z)$ can be easily derived as the product of the PGFs of total delay calculated for each hop,

$$D_{t_{2hop}}(Z) = D_{t_{1st}}(Z) D_{t_{2nd}}(Z) \tag{18}$$

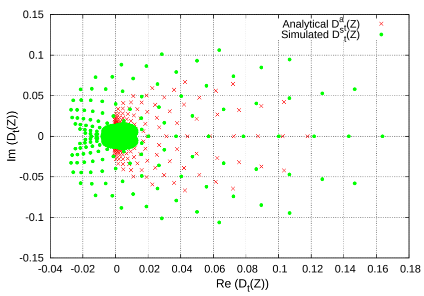
with $D_{t_{1st}}(Z)$ and $D_{t_{2nd}}(Z)$ the PGFs of the 1st hop and the 2nd hop total delays, respectively. For each hop, the PGF is calculated assuming:

- The MAC model is the simple exponential MAC with a mean equal to the mean MAC delay $E[D_m]$ extracted from Eq. (5),
- The queuing model is a simple M/M/1 queue. Thus, the 1-hop total delay PGF follows Eq. (12)

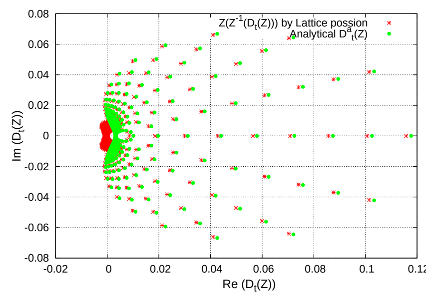
The numerical inversion of the 2-hop total delay PGF uses the LP algorithm of [7] with accuracy 10^{-8}.

The 2-hop total delay distribution derivation has been calculated for a small network of $n = 5$. Figure 8 plots the error in Z space related to the model in Figure 8(a) and to the inversion in Figure 8(b). It is clear on these plots that the error mostly originates from the model. The two metrics proposed in this paper bring us to the same conclusion: $f_{model} = 0.21664$ and $f_{inv} = 0.007917$.

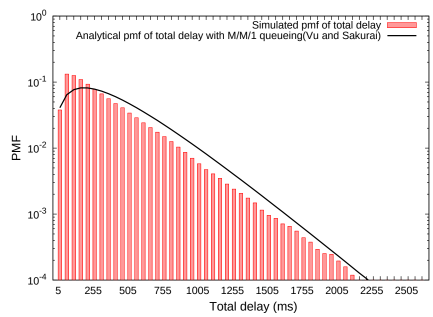
Figure 8(c) presents the 2-hop total delay distribution we are looking for. The y-axis is plot using a logarithmic scale. Most of the errors are concentrated for low 2-hop total delay values. The proposed distribution underestimates the occurrence of these low delay values. Interestingly, the high delay values are pretty well estimated with this model. This fact is explained in the next section which discusses the main assumptions done in the 2-hop study.



(a) Modeling error: $D_t^a(Z)$ and $D_t^s(Z)$



(b) Inversion error for $D_t^a(Z)$



(c) Total delay PMF for 2-hop

Figure 8: Results for a 2-hop communication of $n = 5$ nodes

6.2. Discussion on the main assumptions

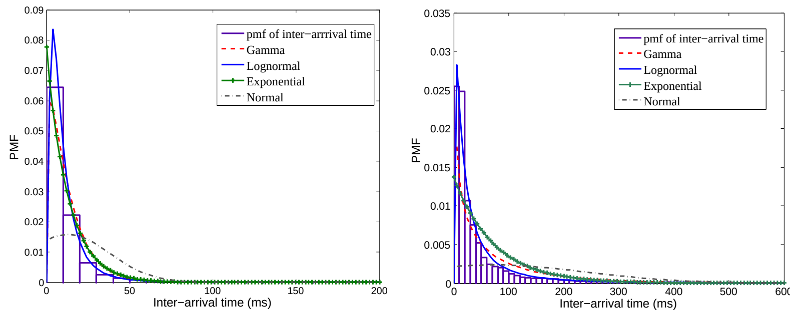
It is shown in this sections that the assumptions made are not realistic. In this section, we will mainly discuss the second assumption. For the first independence assumption, in most of the cases, the time spent by a packet to travel on the 2nd hop is independent of the time it has spent traveling on the 1st hop since R does not relay packets of other nodes. The dependence would exist for the packets being lost on the first hop. But here the delay of lost packets is not accounted for in the delay distribution calculation.

A more precise verification of the second assumption has been done to understand the results of Figure 8. To verify if the arrival date distribution is Poisson at the relay node, the inter-arrival time distribution of the packets arriving at R are plot on Figure 9 for $n = 5$ and $n = 30$. They are compared to gamma, lognormal, exponential and normal distributions. From Figure 9, the exponential distribution seems to provide a reasonably good approximation for the smaller network ($n = 5$). The log-normal distribution provides a good approximation for both cases. This conclusion is in line with the one of Zhai et al. [8].

For $n = 5$, the small delay values are the ones that are less well captured by the exponential distribution compared to the lognormal one. On the opposite, for high delay values, the exponential distribution is as precise as the lognormal distribution. Thus, it makes sense that the same kind of observation exists for the 2-hop total delay distribution of Figure 8(c), where the 2-hop delay model better captures higher delays than smaller ones.

7. Conclusion

This paper proposes a performance evaluation method to characterize the accuracy of a delay distribution derivation. This method is capable of decoupling the error originating from the analytical model from the error induced by the probability generating function inversion. The method has been illustrated on MAC, queuing and total delay distribution models for an IEEE DCF



(a) $n = 5$ (b) $n = 30$

Figure 9: Inter-arrival time distribution

medium access protocol under saturated conditions for a 1-hop and a 2-hop communication.

Future work will leverage the proposed performance evaluation method to provide an analytical model to capture the multi-hop end-to-end delay distribution and extend the current proposition to non-saturated networks. This work has highlighted that in the 2-hop scenario, the arrival process at the relay is not Poisson distributed anymore. The main challenge will be then to incorporate the log-normal inter-arrival time at the relay, using a G/G/1 queue. G/G/1 queueing has already been leveraged by Tickoo and Sikdar [21] to extract the average single hop end-to-end delay. The main difference is that only the relay uses a G/G/1 queue in the 2-hop scenario. Extending the model to more than two hops will be even more challenging since additional relays may modify the arrival distributions of the last nodes of the linear network.

References

1. IEEE Standard. 802.11-2012 IEEE Standard. Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. 2012.

2. Petersen S, Carlsen S. WirelessHART versus ISA100.11a: The format war hits the factory floor. *IEEE Industrial Electronics Magazine* 2011;5:23–34.

3. Pulgar T, Scharbarg JL, Jaffrès-Runser K, Fraboul C. Extending CAN over the air: an interconnection study through IEEE802.11. In: *IEEE ETFA*. Cagliari, Italy; 2013:10–5.

4. Jaffrès-Runser K. Worst case delay analysis for a wireless point-to-point transmission. In: *11th International Workshop on Real-Time Networks*. Pisa, Italy; 2012:.

5. Engelstad P, Osterbo O. The delay distribution of IEEE 802.11e EDCA and 802.11 DCF. In: *Proceeding of IPCCC 2006*. 2006:10–96.

6. Vardakas J, Sidiropoulos MK, Logothetis M. Performance behaviour of IEEE 802.11 distributed coordination function. *Circuits, Devices Systems, IET* 2008;2(1):50–9.

7. Vu H, Sakurai T. Accurate delay distribution for IEEE 802.11 DCF. *IEEE Communications Letters* 2006;10(4):317–9.

8. Zhai H, Kwon Y, Fang Y. Performance analysis of IEEE 802.11 MAC protocols in wireless LANs. *Wireless Communications and Mobile Computing* 2004;4(8):917–31.

9. Dong L, Shu Y, Chen H, MA M. Packet delay analysis on IEEE 802.11 DCF under finite load traffic in multi-hop ad hoc networks. *Science in China Series F: Information Sciences* 2008;51:408–16.

10. Xie L, Wang H, Wei G, Xie Z. Performance Analysis of IEEE 802.11 DCF in Multi-hop Ad Hoc Networks. In: *Proceedings of NSWCTC '09*; vol. 1. 2009:227–30.

11. Bianchi G. Performance analysis of the IEEE 802.11 distributed coordination function. *Selected Areas in Communications, IEEE Journal on* 2000;18(3):535–47.

12. Wang Q, Jaffrès-Runser K, Scharbarg J, Fraboul C, Sun Y, Li J, Li Z. Delay distribution of IEEE802.11 DCF: a comparative study under saturated conditions. In: *Proceedings of PE-WASUN 2013*. 2013:41–8.

13. Willig A. A short introduction to queueing theory. Tech. Rep.; TU Berlin, Telecommunication Networks Group; 1999.

14. Kleinrock L. Theory, Vol. 1, Queueing systems. Wiley-interscience; 1975.

15. Xu K, Gerla M, Bae S. Effectiveness of RTS/CTS handshake in IEEE 802.11 based ad hoc networks. *Ad Hoc Networks* 2003;1(1):107–23.

16. Abate J, Choudhury G, Whitt W. Computational Probability; chap. An Introduction to Numerical Transform Inversion and its Application to Probability Models. Boston: Kluwer; 1999:257–323.

17. Alizadeh-Shabdiz F, Subranmaniam S. Analytical models for single-hop and multi-hop ad hoc networks. *Mobile networks and Applications* 2005;11:75–90.

18. Abreu T, Baynat B, Begin T, Guérin-Lassous I. Hierarchical modeling of IEEE 802.11 Multi-hop Wireless Networks. In: *Proceedings of ACM MSWiM*. Barcelona, Spain; 2013:.

19. Tickoo O, Sikdar O. Modeling queueing and channel access delay in unsaturated IEEE 802.11 random access MAC based wireless networks. *IEEE/ACM Transactions on Networking* 2008;16(4):878–91.

20. Nguyen NV, Guérin-Lassous I, Moraru V, Razafindralambo T. Characterisation and application of idle period durations in IEEE 802.11 DCF-based multihop wireless networks. In: *MSWiM*. 2012:277–84.

21. Tickoo O, Sikdar B. Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks. In: *Proceedings of INFOCOM 2004*; vol. 2. 2004:1404–13.