Analytical Approaches for Short-range Wireless Technologies Evaluation

Vitaly Petrov
Tampere University of Technology
Tampere, Finland
vitaly.pertov@tut.fi

Abstract—Analytical evaluation approaches play a very important role, when the applicability of short-range wireless technology is assessed for emerging applications and scenarios, like machine-to-machine, device-to-device, heterogeneous networking, etc. Due to a set of limitations of measuring and simulation strategies, a closed-form equation describing the network characteristics as a function of parameters could be very suitable in the cases, that were not studied before: new application, traffic pattern or scenario. This paper presents an overview of existing approaches aiming to evaluate different types of short-range networking. Starting from simple cases: channel modeling and saturated system analysis, we finally come to the stability criteria and different traffic patterns affect on the network performance.

Keywords—Short-range networks, Wireless networking, Analytical approach, Saturated network analysis, Channel modeling, Markov models, Performance evaluation, Overload control, IEEE 802.11, Wi-Fi, CSMA/CA, Machine-to-Machine.

I. INTRODUCTION

Significant changes happened to the paradigm of wireless technologies evaluation over the last decade. Only ten-fifteen years ago majority of performance estimation results were obtained by analytical approaches, that incorporate different mathematical apparatus from probability theory, stochastic processes, Markov models, etc. However, with respect to lots of improvements being applied to existing wireless protocols, it became more and more difficult to build an accurate network model, any of the known analytical technique is applicable to. As such, following the trend of CPU performance growth, the main role in networks evaluation passed to simulation techniques.

Currently, plenty of different simulation methodologies [1] are used for a variety of applications, starting from overload control [2] and coming to sparse data traffic handling [3]. World leading standardisation committees (3GPP [4] and IEEE [5]) now accepting results from link-level and system-level simulators and even publish some calibration methodologies, regarding software implementations of particular signals scheme, for example Uplink Channel in IEEE 802.16 [6] and Random Access Channel (RACH) in LTE [7].

In addition to simulation, a set of performance evaluation results could be obtained by another strategy, very appreciated by the industry: test bench design. A number of test benches, starting from simple link-level [8] up to huge system-level, that can include tens of nodes [9], are now widely used by many research groups.

Low complexity of network simulation, using special tools such as OPNET [10] or NS-3 [11], as far as the need of a real prototype design, when proposing amendments to existing standards, resulted in belief that analytical approaches are not quite necessary to be applied for wireless network evaluation.

However, even though simulation and test bench approaches are very beneficial to measure the particular technology performance under a fixed set of parameters, most of them are not applicable to solving optimizations tasks. The full search through all the possible combinations of parameters values is time-consuming.

At the same time, using a set of equations describing system dynamics, some can easily find an optimal set of parameters for a given metric [12]. Moreover, analytical approaches can provide accurate estimations for emerging technologies performance, e.g. Machine-to-Machine [13], [14], when simulation results can only be obtained over long runs (several hours of system uptime) and test bench could not be easily assembled due to financial limitations: thousands of devices to be deployed.

Therefore, in this paper we present a survey of existing analytical approaches for wireless technologies evaluation. We start with the channel modeling, as one of the ways to estimate the channel conditions and, consequently, predict values of Bit Error Rate (BER), Block Error Rate (BLER) and Packet Error Rate (PER) (Section II). Then, in Section III we focus on performance evaluation of widely used short-range wireless access technology - CSMA/CA. Finally, we note the effect of particular types of traffic on the overall network performance (see Section III-C). The paper ends with some conclusion remarks.

II. WIRELESS CHANNEL MODELING

Accurate modeling of wireless channel characteristics provides a starting point in any performance modeling framework. They are often represented in terms of the stochastic process as a function of propagation environment and transmission characteristics such as emitted power, modulation, noise, length of the frame, etc. Depending on how many of these properties are taken into account we distinguish between (i) models representing the received signal strength or related process at
the physical layer (ii) PDU error models at layers higher than physical.

To represent characteristics of a wireless channel in terms of the received signal strength, propagation models need to be used. We distinguish between theoretical, and empirical propagation models. Empirical models are based on measurements. In such models all propagation phenomena are taken into account regardless of whether they can be separately recognized. Their accuracy depends on similarities between the environment under consideration and the environment where measurements have been carried out. Theoretical models are based on principals of radio propagation such as diffraction, reflection, and scattering. Unfortunately, most of these models are relatively simple and cannot provide the required accuracy. As a result, empirical models need to be used.

As an example we consider the measurement-based model proposed in [15], where the authors modeled the signal-to-noise ration (SNR) process of IEEE 802.11b wireless channel. They found SNR process to follow normal distribution with geometric autocorrelation function and proposed to capture these parameters using autoregressive process of order 1, AR(1). AR(1) process is given by $X(n+1) = \phi_0 + \phi_1 X(n) + \epsilon(n+1)$, $n = 0, 1, \ldots$, where $\phi_0$ and $\phi_1$ are some constants, $\epsilon(n), n = 0, 1, \ldots$ are independently and identically distributed random variables having the same normal distribution with zero mean and variance $\sigma^2[\epsilon]$. This process is fully characterized by three parameters $(\phi_0, \phi_1, \sigma^2[\epsilon])$. The authors showed that they can be found as

\[
\begin{align*}
\phi_1 &= \frac{K_Y(1)}{1 - \phi_1} , \\
\phi_0 &= \mu_Y(1 - \phi_1) , \\
\sigma^2[\epsilon] &= \sigma^2[Y](1 - \phi_1^2) ,
\end{align*}
\]

where $K_Y(1)$, $\mu_Y$, and $\sigma^2[Y]$ are the lag-1 autocorrelation coefficient, mean, and variance of SNR process, respectively. This model is valid when SNR is averaged over relatively long time intervals, i.e. more than 0.1s.

A. PDU error models

Another way to describe wireless channel characteristics is to use PDU error models including symbol, bit, frame, and packet error models. These models are represented by binary stochastic processes where 0 stands for correct PDU reception, 1 refers to incorrectly received PDU. Both bit and frame error processes are conventionally modeled using Markov chain starting from simple two-state Gilbert model [16] to quite complicated large-states high-order Markov models [17], [18]. There are a number of PDU error models for layers higher than data-link. The models of bit error observations can be classified into two categories. There are models based on partitioning of SNR observations and models and models based on direct fitting of parameters.

A Markov modulated model based on partitioning of SNR was originally proposed in [19]. Assuming the Rayleigh fading the authors associated each state of the finite-state Markov model with a bit error probability. To parameterize this model one has to determine the transition probability matrix and the bit error probability vector $\vec{p} = (p_1, p_2, \ldots, p_M)$. This model is often referred to as finite state Markov chain (FSMC).

The major advantage of PDU error models is that they can be derived directly from the PDU error traces. The first work in this area dates back to 1960 when Gilbert proposed a simple two-states Markov chain [16]. One of the states is error free while another one is associated with a certain bit error probability. Elliott extended this model allowing errors to occur in both states [20]. One more extension is due to Fritchman [21] who used three-states Markov chain. A model with finite Markov chain based on direct fitting of parameters has been formulated in [22].

An example of the Markov model based on direct fitting of parameters is discussed below. Let the bit error process be covariance stationary with a certain BER and lag-1 autocorrelation. In order to capture these two parameters it is sufficient to use two-states Markov chain. The model is defined by $2 \times 2$ matrices $D_E(k)$, $k = 0, 1$, containing transition probabilities from state $i$ to state $j$ with correct $(k = 0)$ and incorrect $(k = 1)$ reception of a bit. The authors in [23], [24] showed that the model is parameterized as

\[
\begin{align*}
\alpha_E &= [1 - K_E(1)]E[W_E] , \\
\beta_E &= [1 - K_E(1)](1 - E[W_E]) , \\
f_0,E(1) &= 0 , \\
f_1,E(1) &= 1
\end{align*}
\]

where $\alpha_E$ and $\beta_E$ are transition probabilities from state 0 to state 1 and from state 1 to state 0, respectively, $K_E(1)$ is the lag-1 autocorrelation of bit error observations, $E[W_E]$ is the mean of bit error observations, $f_0,E(1)$ and $f_1,E(1)$ are probabilities of incorrect reception of a bit in states 1 and 2, respectively.

B. Stationarity of wireless channel

In [25], the authors carried out statistical analysis of IEEE 802.11b frame error traces and applied a number of well-known channel models. Their results shows inconsistency in using Markov models to represent PDU error processes, that is, no models were found to provide good results for all available measurements. Statistical analysis of IEEE 802.11b frame error traces was also done in [26], where it was shown that Markov models with large state spaces are capable to capture autocorrelation of frame error traces. The authors also noticed some uncertainty as the number of states required for accurate modeling varied from trace to trace. These observations serve as an indication that some important statistic is undetected.

One possible reason is non-stationarity of channel process. When a users is allowed to freely move and the received signal strength may substantially vary. In [25], the authors studied statistical characteristics of IEEE 802.11b frame error traces and revealed that the frame error process has exceptionally strong memory lasting longer than geometrically decaying function. A number of other studies reported similar properties of PDU error processes at higher layers (see e.g. [18]). This implies that either the underlying processes have exceptional memory or they are not stationary.

Wireless channel modeling studies often assume covariance stationarity providing no reason for that. The authors in [27] found their GSM bit error traces to be non-stationary and proposed an algorithm to extract covariance stationary parts. The isolated parts were shown to closely follow simple Markov models. The authors in [28] observed similar behavior of
IEEE 802.11b SNR observations. They used exponentially-weighted moving average (EWMA) change-point test to isolate covariance stationary parts. This procedure is implemented using upper and lower control limits, $E[LY] \pm C(n)$, where

$$C(n) = k \sigma[Y] \left( \frac{\gamma}{2} \right) \left( 1 + \phi_1(1-\gamma) \right),$$

where the actual EWMA statistics $\{LY(n), n = 0, 1, \ldots \}$ is

$$LY(n) = \gamma Y(n) + (1-\gamma)LY(n-1),$$

$\gamma$ is the smoothing parameter, $k$ is a parameter depending on the properties of the underlying process, $\sigma[Y]$ is the variance of original observations, $\phi_1$ is the parameter of AR(1) model used to model covariance stationary parts. More details are provided in [28]. This approach was extended to frame error observations in [29].

### III. Modeling of CSMA/CA WLANs

In distributed (random) wireless access environment there is contention for resources between mobile stations. IEEE 802.11 wireless local area network (WLAN) is a perfect example of such environment. Modeling of random access systems is more complicated task compared to those with centralized access as in addition to performance degradation caused by wireless channel characteristics we have to take into account collisions caused by simultaneous transmissions from multiple stations.

#### A. Random channel access protocols

ALOHA was the first protocol proposed for distributed access environment. According to it a station having a frame for transmission begins its transmission immediately. Obviously, when the number of mobile stations grows performance of ALOHA quickly degrades. Slotted modification of ALOHA reduces collision probability between stations dividing time into fixed length units. In slotted ALOHA frames start to be transmitted only at slot boundaries. Although its performance in terms of throughput was found to be twice of the original protocol it is still unacceptably low, see [30] for analysis.

Carrier sense multiple access (CSMA) protocols is the most popular approach for distributed channel access. CSMA with binary exponential back-off (BEB), also known as CSMA with collision avoidance (CSMA/CA), is a modified version of non-persistent CSMA algorithm specifically designed for IEEE WLANs. According to it, if the channel is busy a station delays its transmission attempt for a certain time. This time is chosen randomly from $(0, CW_i)$, where $CW_i$ is the length of the contention window (CW) at the transmission attempt $i$. With each unsuccessful transmission attempt the length of CW doubles. The number of transmission attempts is limited to $N$. If a frame is still incorrectly received after $N$ transmission attempts an error is returned to higher layers.

#### B. CSMA/CA analysis

The fundamental model proposed for CSMA/CA is due to Bianchi [31]. He evaluated throughput obtained by a single station in saturation conditions. The author studied CW evolution using two-dimensional Markov chain, $\{S(n), B(n), n = 0, 1, \ldots \}$, where $S(n) \in \{0, 1, \ldots , m\}$ is the back-off stage, $B(n)$ is the back-off counter, $B(n) \in \{CW_{i,min}, CW_{i,max}\}$, $CW_{i,min} = W_i$, $CW_{i,max} = 2^n W_i$, $W_i = 2^T$, $T = CW_{0,min}$, $i \in \{0, 1, \ldots , m\}$. Thanks to the structural similarities existing in the Markov process the proposed model allows for closed form solution for probability that a station transmits in an arbitrary slot in the following form

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1)+pW(1-(2p)^m)},$$

where $p$ is the frame collision probability, $W$ is the minimum value of CW.

In order to obtain the final solution we need additional equation binding $p$ and $\tau$ together. There is a collision when two or more stations starts transmitting at the same slot, i.e.

$$p = 1 - (1-\tau)^{-1}. $$

The system of two equations now reads as [31]

$$\left\{\begin{array}{l}
\tau = \frac{2(1-2p)}{(1-2p)(W+1)+pW(1-(2p)^m)} \\
\tau = 1 - (1-\tau)^{-1} \end{array} \right. $$

The system (7) is solved numerically to obtain the stationary regime of CSMA/CA mechanism $(\tau, p)$. Using the probability of transmission in an arbitrary slot one can obtain throughput of the system and mean delay experienced by an arbitrary frame as discussed in [31]. Although the model was shown to provide fairly accurate results in saturation scenario, it takes a number of unrealistic assumptions that was addressed in later studies discussed below.

The first attempt to relax saturation assumption was taken in [32]. The authors assumed the finite arrival rate at each node, $\lambda$, and get the mean CW size in unsaturation condition as $E[W] = (1 - \tau_0)E[W_S]$, where $\tau_0$ is the probability that the node’s buffer is empty, $E[W_S]$ is the mean CW size in saturated condition. The parameter $\tau_0$ is approximated by $(1-\rho)$ holding for M/M/1 queueing system, with $\rho = \lambda/\mu$, where $\mu$ is the mean service rate of frames. Observing that $k$th retransmission attempts increases the mean waiting time by $2^k/W_{min}$ the mean CW size is given by

$$E[W] = \frac{(1-p)W_{min}}{2} + \frac{p(1-p)2W_{min}}{2} + \ldots + \frac{p^{m+1}2^mW_{min}}{2} = 1 - p - (2p)^m W_{min}.$$
all the stations are independent of each other, probability that $k$ out of $N$ stations are active at arbitrary instant of time is given by binomial distribution with parameters $(b_0, k, N)$. To get $b_0$, the authors used M/G/1/K queue, where the service time distribution is the distribution of the back-off window size found at previous step. Note that the solution of the model is iterative in nature. To derive transition probabilities of the Markov model describing the back-off mechanism, the probability that the node is busy is needed. This probability is obtained by solving M/G/1/K queue which needs the service time distribution.

There are a number of other models proposed recently for unsaturated conditions, see e.g. [33], [34], [35], [36]. The only difference between them is how the service time distribution for queueing systems is computed. In [35] the authors used the probability generating functions while Ozdemir and McDonald derive it using probabilistic arguments. The authors in [35] also considered the error of replacing M/G/1/K queuing system by M/M/1/K one and it happened to be negligible. In [36], the authors extended the work performed in [35] developing more accurate approach to estimate the service time distribution. Additional step forward has been performed in [37], where the authors used a queuing model with correlated arrivals.

One notable shortcoming of all those abovementioned studies is that wireless channel conditions were assumed to be perfect presuming that collision-related losses is the main source for performance degradation. The authors in [38] relaxed this assumption showing that the second equation in Bianchi model should look like

$$p = 1 - (1 - p_L)(1 - \tau)^{n-1}$$

(9)

where $p_L$ is the probability of loss due to fading.

Packet losses due to wireless channel characteristics were also taken into account in [36]. The common shortcoming of both [36] and [38] is that the loss probability was assumed to be constant for all frames. The the best of our knowledge [39] is the most general study of CSMA/CA mechanism as it includes most factors mentioned above, i.e., unsaturated channel conditions, channel and collision related losses, CW evolution according to BEB, and limited buffer space of end systems. The only simplifying assumption is constant collision probability at different stages of the back-off procedure. The solution proposed in [39] is iterative in nature.

A very special approach has been taken in [40]. The authors proposed a state-dependent M(i)/PH/1/K queueing system as a model of the frame service process in CSMA/CA environment. The arrival process in this system represents the number of stations that become active, that is, receives a frame to be transmitted. Once a station becomes active it start contention for resources on a shared medium. The station arrival process is assumed to be Poisson. When there are $i$ stations competing for resources the service time distribution of a single packet can be obtained applying the original Bianchi model with $i$ saturated stations. The authors in [40] investigated the distribution of the service time of a packet and found that it closely follows PH distribution. The proposed model is of special interest as it combines two separate models to represent behavior of a complex system.

One more interesting approach was introduced in [41], where the authors replaced the Markov model used in [31] by the closed queuing network model. A back-off stage of the BEB algorithm is modeled by a separate queue while the number of queues corresponds to the number of back-off stages. The service time at the queue $i$, $i = 0, 1, \ldots, N$ corresponds to the back-off interval and follows discrete uniform distribution between $0$ and $CW_i$. The probability of leaving a network after getting service at the queue $i$ $(1-p)$ describes the situation of successful frame transmission at the $i$th transmission attempt. With complementary probability $p$ the frame enters the next queue and is delayed for some additional time uniformly distributed between $0$ and $CW_{i+1}$. The probability $p$ is, thus, the probability of collision and assumed to be the same for all back-off stages. From the closed nature of a network it is easy to observe that served customers immediately return back to the beginning of the system implying saturated conditions. Since the proposed structure was intended to represent the whole CSMA/CA contention environment with constant number of stations, the authors had to ensure that no frames are waiting in queues before they start getting service. To fulfill this requirement they used G/G/$\infty$ queues as elements of their network. Due to the chain-like structure and observing that the probability of a transmission attempt in an arbitrary slot is $l/N$, where $l$ is the total rate of attempts and $N$ is the number of stations, the analysis is straightforward. Furthermore, the solution for $\tau$ can be obtained in closed form. Similarly to [31], the authors obtained an independent expression for $\tau$ inverting $p = 1 - (1 - \tau)^{n-1}$ and then solved the system of equations numerically to get $(\tau, p)$. Using $\tau$ and $p$, the system throughput and mean delay metrics have been obtained.

C. TCP modeling in WLAN environment

Modeling performance of TCP running over wireless channels with distributed access scheme is more complicated compared to stations having real-time traffic only that does not use loss-controlled protocols such as TCP. As we observed above the distributed access is complicated to model analytically even when the arriving traffic is independent of the network conditions. Adding TCP congestion and flow control procedures on top of this would bring a lot of additional challenges to a performance analyst. For example, the congestion window size of a TCP source, and therefore, the amount of frames that need to be transmitted depends on both RTT and the frame loss probability. In addition to those losses caused by buffer overflow and error-prone nature of a wireless channel, the latter parameter should now take into account the probability of frame collisions which, in turn, depends on the congestion window size. The situation becomes even more complex when TCP timeouts need to be taken into account. Indeed, as it was demonstrated in [42] the frame delay distribution is multimodal in CSMA/CA networks. This property may lead to frequent TCP timeouts. As a result, little to none progress has been achieved so far in analytical modeling of TCP performance in distributed access environment.

One of the promising ways to approach the problem of TCP modeling in distributed access environment is to divide the whole system into a number of parts, analyze these parts separately and then use a certain technique to solve the resulting equations simultaneously. This modeling approach resembles the fixed point approximation that is widely used to analyze performance of TCP sources in wired networks,
see e.g. [43], [44]. In order to accomplish this a framework similar to that one proposed by Foh and Zukerman for real-time applications can be used [40] (discussed in the previous section). Firstly, one can represent the number of active stations using a queuing system of state-dependent arrivals, e.g. Geo(1)/G/1/K queue. In this queue the arrival process represents those stations that become busy, i.e. receive a new frame for transmission. The service process reflects stations that become empty, i.e. transmitted all their frames and have no new frames for transmission. The overall state of the system is represented by the number of active stations. The service process in each state of the system can be obtained using the set of queuing systems of G/G/1/K type. In those queues the arrival process and service times remain unchanged and correspond to $i$ active stations in the system. The service time distribution can be obtained using the model originally presented by Bianchi in [31] and then extended to the case of non-ideal channel conditions in [38]. The solution of G/G/1/K queuing systems gives the steady-state probabilities $p_i$ that there are $i = 0, 1, \ldots, N$ active stations in the system. Weighting the service time distribution in each state of the system with respective steady-state probabilities of these states we get the overall service time distribution for Geo(1)/G/1/K queuing system. Finally, we can use the fixed-point approximation technique to determine the throughput of TCP sources in CSMA/CA environment. Note that the model in this form can be computationally intensive as it requires to solve $i + 1$ queuing systems to get a single performance measure. Such a model has been recently proposed in [45].

IV. Conclusions

Wireless networks play an important role in human daily life. However, there are certain conditions, when currently deployed solutions can not efficiently handle traffic from some applications, such as VoIP or M2M. Withing the last decades a number of techniques was proposed in order to predict network performance in particular scenarios. In this paper a survey on analytical approaches to perform short-range wireless networks evaluation is given. Listed methods can be applied for different characteristics estimation, starting from PER values on a single Wi-Fi link up to a system-level performance of the whole network with dynamic data flows.

Presented techniques and their modifications could be used both for conventional network analysis as far as prediction of future networks behavior, handling emerging applications, for example machine-to-machine [46] or device-to-device [47].

REFERENCES


