Adaptive Communications and Signal Processing laboratory
Cornell University, Ithaca, NY 14853

Measurement-Based Models for Cognitive Medium Access in WLAN Bands

S. Geirhofer, L. Tong, and B. M. Sadler

Technical Report ACSP-TR-02-07-02

March 2007

Contents

Abstract

This technical report corroborates the design of Cognitive Medium Access (CMA), a protocol enabling coexistence with multiple parallel WLAN channels by keeping interference below a given constraint. The derivation of CMA in [1] is solidified by measurement-based models, including an experimental study of the physical layer interaction between both systems, as well as a stochastic model allowing to predict the WLAN's medium access. Due to space limitations our paper [1] focused on the derivation of CMA and provided only the main results of our experimental studies. In this technical report we cover the measurement aspect in detail and provide additional information to substantiate our results.

1 Introduction

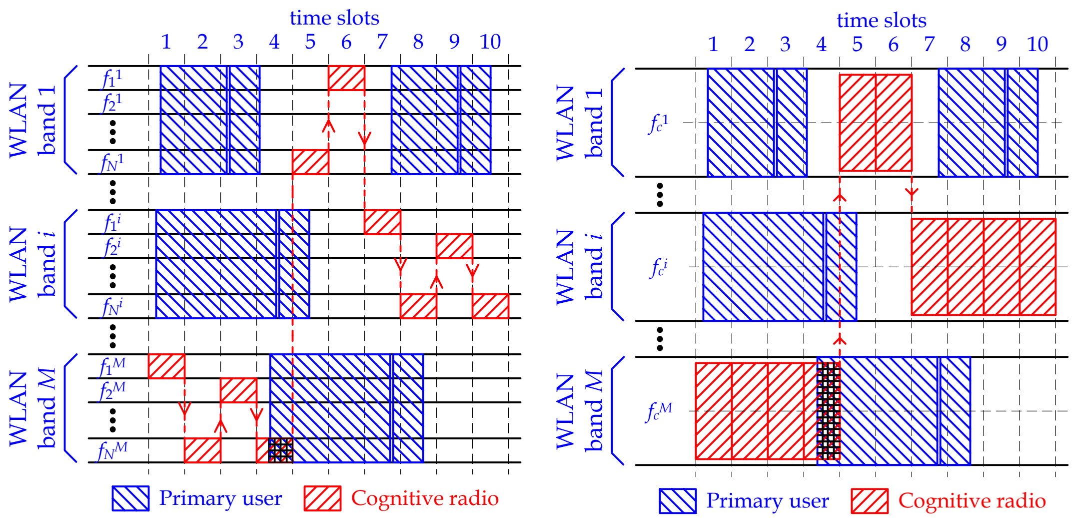
We consider the problem of designing a cognitive radio that can coexist with multiple parallel, independently evolving WLAN channels. We focus on a hierarchical setup [2] that requires the cognitive radio to abide by interference constraints, limiting the tolerable collision rate between both systems. Our contribution in designing the cognitive radio is twofold. First, we study by experiment the physical layer interaction between both systems. Second, we derive Cognitive Medium Access (CMA), a scheme that enables coexistence by ensuring that the above listed constraints are met. The derivation of CMA fundamentally relies on three experimental models, namely (i) the cognitive radio's impact on WLAN's carrier sensing, (ii) its impact on the packet error rate, and (iii) a continuous-time Markov chain model capturing the WLAN's bursty medium access. All three models are fundamental to proposing CMA and provide the foundation for our analysis. The experimental models described above were presented in [1] but due to space limitations we could not include all of our results. This technical report includes a thorough treatment of these models and substantiates our work's experimental component.

This report is organized as follows. After introducing the physical layer setup, Sec. 3 characterizes the physical layer interaction between both systems based on experimental studies. Specifically, we evaluate the cognitive radio's impact on the WLAN carrier sensing as well as its impact on the packet error rate. In Sec. 4 we review the stochastic WLAN model which allows us to predict the busy and idle durations of the different bands and ultimately constrain interference.

2 Physical Layer Setup

The physical layer setup considered in this paper consists of M parallel, independently evolving WLAN channels, as shown in Fig. 1. The special case of $M = 3$ is of primary practical interest since the ISM band at 2.4 GHz supports three such channels [3].

Although there are no restrictions in designing the cognitive radio, other than our ultimate goal of minimizing mutual interference, we shall focus on two specific setups. First, we consider the frequency hopping (FH) setup depicted in Fig. 1. Each of the M WLAN bands overlaps with N narrowband hopping channels. We note that this setup is representative of Bluetooth/WLAN coexistence. Second,



(a) Frequency Hopping Setup

(b) Direct-Sequence Spread Spectrum Setup

Figure 1: System setup. The cognitive radio is a time-slotted FH or DSSS system.

we consider the direct-sequence spread spectrum (DSSS) setup depicted in Fig. 1, where the cognitive radio uses the same frequency bands as the WLAN, but employs a different spreading code to reduce interference. Both the FH and DSSS setup are slotted with period T_s. In our analysis we chose $T_s = 625\,\mu$s, which corresponds to the Bluetooth slot duration.

The choice of these setups is practical for two reasons. First, WLAN is an *unslotted* system that performs medium access based on carrier sensing (CSMA/CA). A logical approach for enforcing an interference constraint is thus to sense the medium periodically, and transmit in a slotted fashion. Second, mutual interference is reduced by exploiting the fact that WLAN uses spread spectrum communications and thus has some inherent robustness to narrowband interference, or a DSSS system with different spreading code.

The FH setup is moreover supported by practical experience. The FH setup considered in this work can be seen as a "smart Bluetooth". While the physical layer setup is the same, we are free to design the system's hopping sequence based on sensing results at the beginning of every slot. This similarity enables us to use standard Bluetooth [4] as a benchmark in our numerical evaluation [1].

3 Measurement-Based Interference Model

Based on the physical layer setup introduced in the last section, we characterize the mutual interference between both systems. This will in turn provide a basis for deriving CMA. Our mutual interference model consists of two parts. First, we evaluate whether the cognitive radio impacts the WLAN's carrier sensing. Second, we obtain empirical results for the probability that a collision between both systems leads to a WLAN packet error.

3.1 Impact on WLAN's carrier sensing

The WLAN's behavior is altered *at the transmitter* if the cognitive radio impacts WLAN's carrier sensing. The WLAN uses CSMA/CA [3] for multi access communication. If the WLAN mistakes the cognitive radio's transmission for other active WLAN stations, it will defer medium access. This would not only undermine our paradigm of hierarchical DSA but also render our prediction model useless unless it included the WLAN's backoff behavior. The design of the cognitive radio thus needs to ensure that its transmissions remain transparent to the WLAN.

3.1.1 Measurement setup

We evaluated the cognitive radio's impact by measurement using the setup shown in Fig. 2. It consists of an 802.11b router and an RF signal source, generating the WLAN and the cognitive radio's signal, respectively. More precisely, we consider a static (non-hopping) FH signal with Bluetooth's modulation parameters [4] and a DSSS signal with 802.11b transmission parameters [5] except different spreading code[1]. Since the FH signal remains static in one of the hopping channels it is possible to examine the mutual interference resulting from this specific channel.

As shown in Fig. 2 the WLAN router and the signal source are connected via circulators, which couple signal generator and router while providing isolation in the reverse direction. A WLAN adapter is used to capture the received signal, and an Agilent vector signal analyzer is used to verify the correct operation of the setup. Please see Fig. 2 for names and model numbers of all devices.

The impact of the FH signal on the WLAN was assessed in the following way. The WLAN router continuously transmits packets and the WLAN receiver is used to capture these packets over long periods of time. In this way the rate of transmitted packets can be measured. In the presence of the interferer, some of these packets will not be transmitted since the channel is sensed busy at the transmitter, in turn decreasing the packet rate. It is this change in rate that is used to assess the interference caused by the WLAN.

3.1.2 Measurement result

The numerical results of our experimental study are shown in Tab. 1. The first column lists the signal generator's power level. This signal is present at the WLAN router with a small attenuation induced by the circulators (about $1.2\,$dB). The remaining columns in Tab. 1 specify the channel of the FH interferer with respect to the center frequency of the WLAN band. In our experimental study we focused on WLAN Channel 11 [3], which is located at a center frequency of $2.462\,$GHz. The 802.11b WLAN signal has a bandwidth of $22\,$MHz.

The result of our measurement can be interpreted as follows. The impact of the FH interferer on the WLAN's carrier sensing heavily depends on the channel index. In fact, the router seems to

[1] The standard prescribes the Barker code [1,0,1,1,0,1,1,1,0,0,0] (see [5]). We used [1,1,0,1,1,1,0,1,0,0,0].

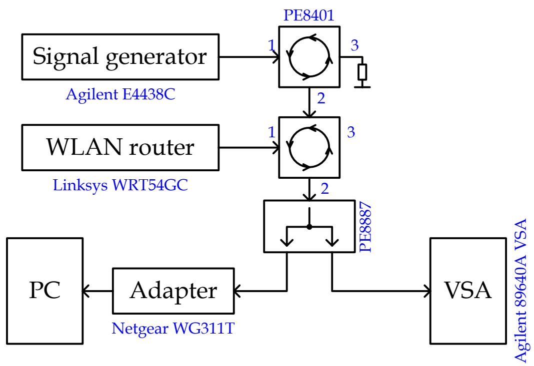


Figure 2: Measurement setup used to assess the impact on the WLAN's carrier sensing.

	Offset from center frequency in MHz				
P_1/dBm	-10	-9,. . . ,-1	0	1,. . . ,9	10
+10	1	–	1	–	1
-10	1	–	1	–	1
-30	41%	–	29%	–	40%
-50	–	–	28%	–	–

Table 1: Assessing the impact of a FH interferer on the WLAN's carrier sensing.

perform energy detection in narrow bands spaced about 10 MHz apart. Interference present in channels in between does not appear to impact the device's sensing behavior.

As a result, the WLAN's medium access will not be impacted by transmissions outside these narrowband channels. Furthermore, even if the cognitive radio transmits in those channels, no impact will occur as long as the interferers power level is small enough.

Given typical setups [6] and path loss models [7], we can thus conclude that the cognitive radio does not alter the WLAN's medium access. This solidifies our hierarchical approach and renders the stochastic prediction model applicable.

DSSS interferer In the case of a DSSS interferer, transmitting at the same center frequency as the WLAN but employing a different spreading code, we did not find any impact on the WLAN's carrier sensing up to a power level of $+10$ dBm. We conclude that a DSSS interferer does not have any impact on the WLAN system.

WLAN interferer Lastly, we consider interference from a DSSS type signal that uses the same spreading code as the WLAN. While such a signal design is not attractive from the coexistence viewpoint, we use this signal to validate our measurement setup. Ultimately, WLAN devices have to guarantee a certain sensing performance to ensure inter-operability between different WLAN manufacturers. There

P_0	-65	-70	-75	-76	-77	-80
CCA	98%	92%	73%	1%	–	–

Table 2: Assessing the impact of WLAN-type interference on the WLAN's carrier sensing.

are specifications for this so-called Clear Channel Assessment, which can be found in the standards [3, 5]. The standard mentions that sensing can be based on energy-detection or by identifying the WLAN's specific spreading code but does not mandate either method; only a required detection performance is specified. Using our measurement setup we confirm that this specification is met, which in turn substantiates our results.

A table with measurement results for this scenario is found in Tab. 2. It shows for an interferer's power level higher than -70 dBm the channel is continuously being sensed busy. For power levels smaller than -77 dBm on the other hand, the interferer does not impact the WLAN's carrier sensing. This result is in accordance with the standard, which specifies a sensitivity threshold of -76 dBm [5, p.58].

3.2 Impact on WLAN's packet error rate

3.2.1 Measurement setup

The second part of our mutual interference characterization focuses on the cognitive radio's impact on the packet error rate. Specifically, we measure the probability that a collision between both systems leads to a WLAN packet error. The measurement setup is shown in Fig. 3. It consists of a WLAN adapter card and a signal source generating the WLAN signal and the interferer, respectively. The signals are combined and captured via another WLAN card and the packet capture software "CommView for WiFi". A vector signal analyzer is used to verify the operation of the setup. Details on the used devices including manufacturers and model numbers are shown in Fig. 3.

The packet error probability is obtained in the following way. A continuous stream of packets is generated and captured at the receiver to determine the rate of packets with the interferer turned off. Subsequently, in the presence of interference the rate is measured again. The rate will be reduced since some packets will be too distorted to be captured by the adapter. Other packets will be captured but will show an invalid redundancy check. By comparing the number of successfully received packets with the interference-free case, we can calculate the probability of a packet error.

The measurement setup shown in Fig. 3 gives more details on the setup. Attenuators were used to obtain a configuration that was in accordance with the specification of all devices involved. The intermediate power levels are shown in the figure and correspond to the burst power of the WLAN (we triggered to the start of packets to obtain the appropriate values). We could thus select arbitrary values for the signal to interference value (SIR) *at the WLAN receiver*.

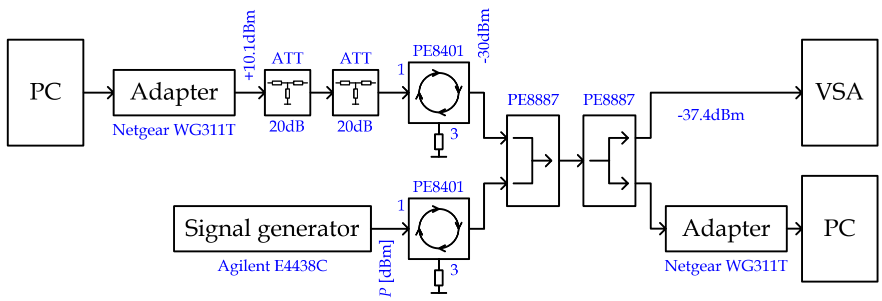


Figure 3: Measurement setup used to assess the impact on WLAN's packet error rate.

P	SIR	Offset from center frequency in MHz										
		0	1	2	3	4	5	6	7	8	9	10
-15	-12.7	100%	100%	100%	100%	100%	100%	100%	100%	97%	–	–
-20	-7.7	97%	100%	100%	100%	100%	100%	92%	12%	1%	–	–
-25	-2.7	13%	94%	94%	85%	55%	10%	1%	–	–	–	–
-30	2.3	2%	6%	4%	2%	1%	–	–	–	–	–	–

Table 3: Measurement results for the FH interferer's impact on the WLAN's packet error rate.

3.2.2 Measurement results

FH interferer The impact of a FH interferer on the WLAN's packet error rate is shown in Tab. 3. The leftmost column represents the power level selected at the signal generator. The SIR that results at the receiver is shown in the second column. The remaining part of the table shows the probability of a collision leading to a packet error with respect to the interferer's offset from the WLAN's center frequency. The results were symmetrical with respect to the center frequency, and thus Tab. 3 only shows results for positive offsets.

Our results illustrate that the interferer's impact is largest close to the center frequency. This is not surprising and has been reported previously [7]. In fact, this behavior can be attributed to the standard receive processing within WLAN adapter cards, specifically downconversion and IF filter stages [7, 8].

Clearly, the packet error probability heavily depends on the SIR available at the receiver, which in turn depends on typical propagation and path loss models. Instead of postulating a 'standard' setup, our derivation of CMA [1] focused on the worst-case assumption that every collision inevitably leads to a packet drop. This assumption is representative for scenarios were cognitive radio and WLAN transmitter are located close to each other. Otherwise it is a worst-case assumption.

DSSS-interferer We performed the same analysis for the DSSS-type interferer. The results are tabulated in Tab. 4. We can see that the results are similar compared to the FH interferer.

P	-25	-26	-27	-28	-29	-30	-35	-40
SIR	-3.1	-2.1	-1.1	-0.1	0.9	1.9	6.9	11.9
PERC	92%	80%	61%	34%	18%	10%	0.4%	–

Table 4: Measurement results for the DSSS interferer's impact on the WLAN's packet error rate.

4 Empirical WLAN model

The derivation of CMA is based on the empirical interference model presented in the previous section. In short, we have seen that while the WLAN's carrier sensing remains unaltered, a packet collision is likely to cause a packet error. For ease of analysis, and because such an assumption has frequently been made in other papers [9], we assume that every collision inevitably results in a packet error. This is a worst case assumption given our measurement results.

Since collisions cause packet errors we need to constrain the rate of collisions between both systems. The derivation of CMA is based on a previously established WLAN prediction model [10, 11]. This section reviews the model specifics, including the measurement setup that was used to gather the empirical data our model is based on.

4.1 Sensing testbed

The empirical data which forms the basis of this stochastic model was gathered via an 802.11b based WLAN operating in the 2.4 GHz ISM band. Different from related publications that capture packets by commerical WLAN adapter cards operating in a special mode we employ a vector signal analyzer to record raw complex baseband data which is subsequently processed to find the start and end times of packets. This approach guarantees an accurate and verifiable characterization of the channel's idle and busy periods.

For recording the baseband data we used an Agilent 89640A vector signal analyzer (VSA)[12] which internally downconverted the RF signals to an intermediate frequency and then was configured to sample at a rate of 44 MHz. We consider both a WLAN communicating via antennas, as well as an RF-isolated setup that guarantees our measurements to be free of interference from other devices operating in adjacent frequency bands. The setup is illustrated in Fig. 4 and Fig. 5, respectively.

4.1.1 Antenna-based setup

The antenna-based propagation setup consists of a Netgear WGT624 wireless router and three computers with wireless adapter cards (two Netgear WG311T and one WG511T; cf. Fig. 4). The setup operated in Channel 11, which represents a 22 MHz frequency band centered at 2.462 GHz. All the equipment was located in the same room, resulting in a high-SNR setup with no hidden terminals. Using the VSA, we verified that interference from adjacent channels was minimum although a completely interference-free setup could not be guaranteed.

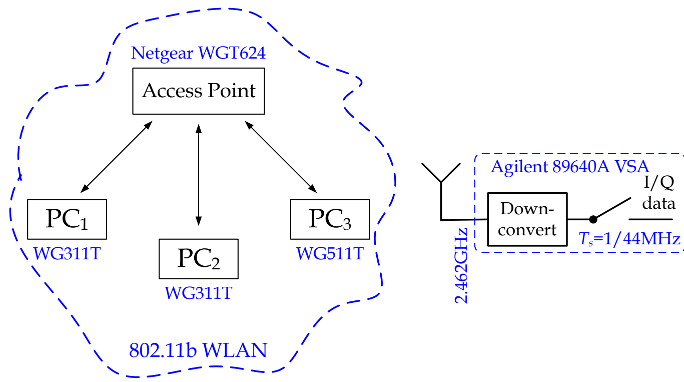


Figure 4: Antenna-based measurement setup.

4.1.2 Isolated RF-setup

Besides the antenna-based setup, we also considered the isolated RF-setup shown in Fig. 5. It consists of a Linksys WRT54GC wireless router and three workstations with Netgear WG311T wireless adapter cards. All the devices are connected to a Broadwave Technologies resistive power divider via RG174U coaxial cables and SMA connectors. The VSA is also connected to the divider resulting in a fully isolated setup. Strictly speaking there is still some residual interference that couples directly via the workstations into the wireless adapter cards. However, given that all devices are connected with coaxial cables this interference is small compared to the desired signal and can be neglected.

The Netgear router used for the antenna-based setup could not be used for the isolated measurements as well since its built-in antenna was non-detachable. The use of two different routers caused our setup to differ in terms of the type of synchronization preamble used. While the Netgear router could be configured to use only long-synchronization preambles, the Linksys router did not allow for specifying this option. As a consequence most of the time a short preamble was transmitted (given the high SNR setup). While this leads to slightly different packet durations, the qualitative behavior of our results remained unaltered.

4.1.3 Traffic generation

Each of the workstations was used to generate traffic using the Distributed Internet Traffic generator (D-ITG)[13]. The software allows for a flexible statistical characterization of the traffic, including varying packet lengths and inter-departure times. A detailed specification of the settings is provided with the measurement results in Sec. 4.3.

Additionally, we also investigate typical usage scenarios of WLAN by using the popular "Skype" voice-over-IP (VoIP) client to set up a conference call within the WLAN, using the traffic generator to simulate G.711 codec based voice communication, and using an SFTP client to download files from a central server. A detailed treatment of the results is again deferred to Sec. 4.3.

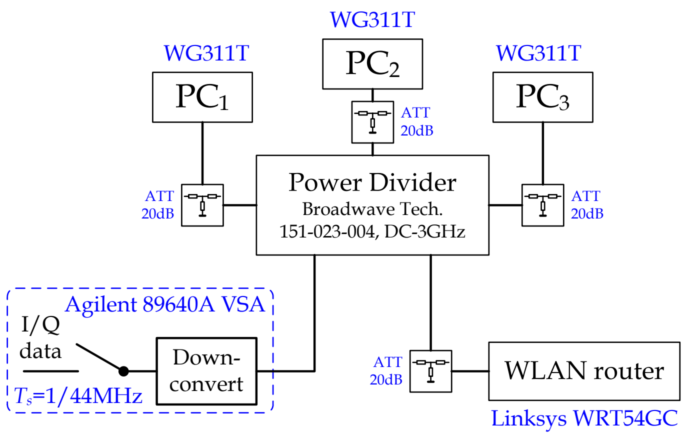


Figure 5: Isolated measurement setup.

4.2 Sensing methods

The two measurement setups described in the last section both yield time captures of the complex baseband signal. Given these data, we process the signals to determine the exact start and end of each packet. Clearly, this fully determines the channel's idle/busy durations.

We consider two different sensing strategies depending on whether the transmission standard of the primary user is assumed to be known. In the former case, the detection of the packets shall be based on energy. In the latter case we can exploit the standard-specifics to achieve better performance [14].

4.2.1 Energy-based detection

If the primary user's transmission standard is unknown, a natural approach for detecting the start and end of packets is based on the transmitted energy. In order to achieve satisfactory performance we consider blocks of N samples whose length is shorter than the smallest packet length [14]. The detection problem can then be formulated as

$$\mathcal{H}_0 \quad : \quad Y_i = V_i, \; i = 1, \ldots, N \tag{1}$$

$$\mathcal{H}_1 \quad : \quad Y_i = S_i + V_i, \; i = 1, \ldots, N, \tag{2}$$

where Y_i denotes the complex baseband samples, V_i are noise samples, $V_i \sim \mathcal{CN}(0, \sigma_0^2)$, and S_i denotes the signal samples drawn from a complex Gaussian, $S_i \sim \mathcal{CN}(0, \sigma_1^2)$. Lacking any information on the transmission standard of the primary user, the Gaussian assumption for S_i appears reasonable.

The hypothesis testing problem defined above is standard [15] and the optimal Neyman-Pearson detector is given by

$$T(\mathbf{y}) = \sum_{i=1}^{N} |y_i|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{3}$$

where the threshold γ is determined according to the probability of false alarm, which amounts to

$$\alpha = \Pr(T(\mathbf{y}) > \gamma | \mathcal{H}_0) = 1 - \tilde{\Gamma}_r(N, \frac{\gamma}{\sigma_0^2}), \tag{4}$$

where

$$\tilde{\Gamma}_r(N, \xi) = \frac{1}{\Gamma(N)} \int_0^{\xi} t^{N-1} e^{-t} dt \tag{5}$$

is the regularized gamma function and $\Gamma(N)$ is the complete gamma function. Similarly, the power of the detector is given by

$$\beta = \Pr(T(\mathbf{y}) > \gamma | \mathcal{H}_1) = 1 - \tilde{\Gamma}_r(N, \frac{\gamma}{\sigma_0^2 + \sigma_1^2}). \tag{6}$$

The above expressions show that the detection performance depends on the SNR $= \sigma_1^2/\sigma_0^2$ as well as the block length N. For our setup we chose $N = 44$ samples, which corresponds to 1 µs long blocks. If we demand $\alpha = 1 - \beta < 10^{-5}$ then we can see that we have to guarantee that the SNR is above 4.29 dB which is easily met in our setup.

Finally, it has to be noted that the Gaussian assumption for the noise V_i might not be appropriate if significant interference occurs. Indeed, this might be a limiting factor if we consider that the WLAN channels are partially overlapping. Suppressing this interference by a filter may thus be necessary in practice.

4.2.2 Feature-based detection

The energy-based detection scheme described in the last section is based on the assumption that the primary user's transmission standard is unknown. In some applications, however, it is reasonable to assume that the transmission specifics are known to the primary user. This knowledge can in turn be exploited to improve the detection of packets.

The layout of an 802.11b physical layer (PHY) frame is shown in Fig. 6. It consists of a PLCP preamble, split into a block of scrambled '1's ('0's for the short-preamble) and the start-of-frame delimiter (SFD) indicating the beginning of the PLCP header. The SFD can be used to precisely detect the start of the packet. The information provided in the header consists of a SIGNAL, SERVICE, and LENGTH field as well as a CRC protecting these three blocks.

From our viewpoint the SFD and the LENGTH field are most interesting; the former determines the start of the packet while the latter provides the duration (and thus the end) of the packet.

The receive processing for the feature-based detection scheme is depicted in Fig. 7. The complex baseband data collected at a rate of 44 MHz is first passed through a Gaussian pulse shaping filter with a bandwidth-symbol time product of $BT_s = 1/2$. In order to obtain chip-synchronization the filtered signal is correlated with the 11-sample Barker sequence specified by the standard [5]. The resulting signal shows periodic peaks whenever the spreading sequence lines up with the input signal. We detect these peaks and downsample the signal to the symbol rate of 11 Mbps. Subsequently, we

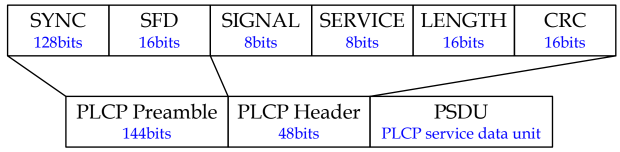


Figure 6: Physical layer preamble in 802.11b (long preamble).

despread and demodulate the DBPSK/DQPSK encoded preamble. The frequency offset at the receiver is noticeable but can be neglected since the signals are differentially encoded. After successful decoding, the resulting bit stream is descrambled and the start-of-frame delimiter (SFD) is detected. In the same way, the SIGNAL, SERVICE, and LENGTH field are extracted and the CRC check is performed to ensure that the extracted information is correct.

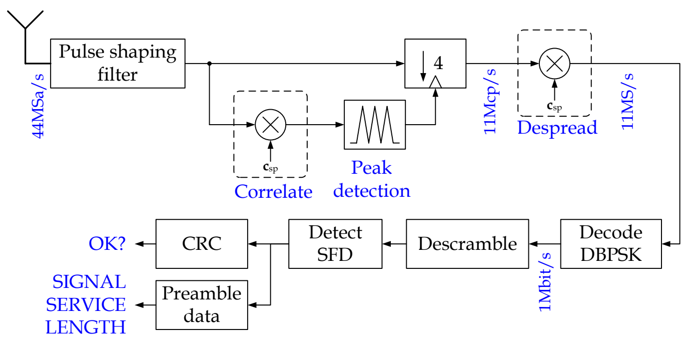


Figure 7: Receive processing for feature-based detection.

4.3 Measurement results

In this section we present the measurement results for the statistics of the busy/idle durations of the channel. We investigate different traffic scenarios as pointed out in Sec. 4.1.3. In particular, we first consider constant length UDP traffic with exponentially distributed inter-arrival times. This allows us to parameterize the 'business' of the channel by increasing the rate parameter σ of this distribution. Second, we consider FTP and Voice-over-IP traffic to investigate whether our idealized setup extends to practical traffic scenarios. For a better understanding of the results we start this section with a brief illustration of WLAN's medium access in order to keep this paper self-contained.

4.3.1 WLAN Medium Access Protocol

The 802.11 standard for WLAN [3, 5] uses the CSMA/CA protocol to control the station's access to the medium (cf.Fig. 8). This implies that before transmitting a packet, the station has to first sense the medium. If the channel is free, the station continues sensing for the distributed coordination function inter-frame space (DIFS). If the channel remains idle during the entire period, the station can go ahead and start transmitting.

After a packet transmission, the receiver has to confirm reception immediately by transmitting an acknowledgement. Only a short inter-frame space (SIFS) is necessary as to give priority to the (required) transmission of acknowledgements (cf. Fig. 8).

If the channel is busy in the first place the station has to defer access until the medium becomes idle again. Then, after a DIFS, a contention window is used to avoid collision between the multiple stations trying to access the medium. Specifically, each station generates a uniform random number $i \in \{0, \ldots, 31\}$ and defers transmission for $iT_{\text{slot}} = i \cdot 20\,\mu$s before accessing the channel (given that no other station has already started to access the channel before).

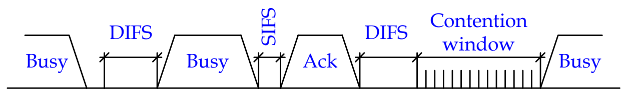


Figure 8: Medium access in an 802.11b-based WLAN.

The standard provides some more technical details that are not addressed above. In particular, if collisions occur the length of the contention window is increased. These specifics, however, do not manifest themselves in our measurement results and shall thus not be addressed here.

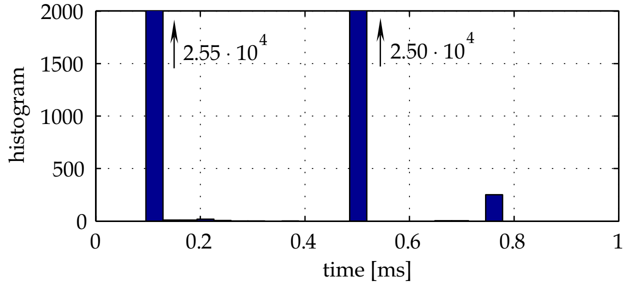
4.3.2 Measurement validation

We first look at a simple measurement scenario to further illustrate the specifics of the medium access and to validate our measurement setup. In particular, we consider the isolated measurement setup depicted in Fig. 5 with only one PC and the wireless router turned on (the other ports of the resistive power divider were terminated to eliminate reflections). The traffic generator was then used to generate UDP packets of constant length $512\,$B with constant inter-arrival times at a rate of 10^5pkts/s. This rate is too high to be transmitted across the channel but ensures that the workstation's transmit buffer is never empty.

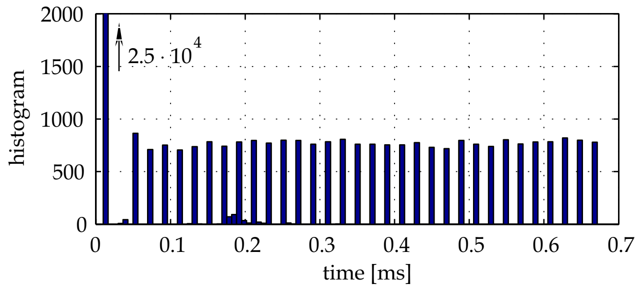
Using the setup described above we used the VSA to capture 100 blocks of complex baseband data, of duration 0.25 s each. The blocks were then processed using both sensing strategies discussed in Sec. 4.2. The results of energy- and feature-based detection match nicely leading to the histograms for the busy/idle durations shown in Fig. 9.

The histograms indeed reflect the characteristics of the standard. First, the histogram of the busy durations depicted in Fig. 9(a) shows only three components, corresponding to the transmission of acknowledgement packets ($t \approx 0.11\,$ms), data packets ($t \approx 0.51\,$ms) and beacon frames ($t \approx 0.76\,$ms), respectively. Given that we forced the data packets to be of constant length, this result is in accordance with our expectations.

The histogram of the idle durations reflects the standard as well. We see a discrete component at



(a) Histogram for the busy durations.



(b) Histogram for the idle durations.

Figure 9: Measurement validation using 1PC (cf. Sec. 4.3.2)

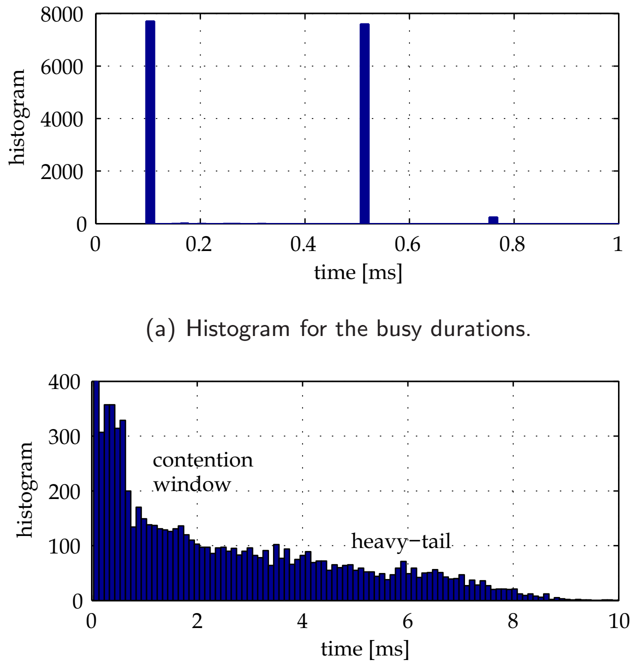
$t \approx 10\,\mu$s, which nicely corresponds to the SIFS. Furthermore, we see 32 discrete components, each spaced $20\,\mu$s apart. These correspond to the contention window as described in Sec. 4.3.1.

4.3.3 Constant Payload UDP Traffic

In the last section we have validated the measurement setup using a simplified traffic scenario. In this section we are now using all three workstations together with the wireless router (cf. Fig. 4 and Fig. 5). The traffic generator was used to generate UDP packets of constant length of 1024 B but the inter-arrival rates for each workstation were now drawn from independent exponential distributions with common but varying rate parameter σ. As σ increases, the number of transmitted packets per unit time increases and consequently the amount of whitespace decreases.

Exemplary histograms for busy and idle durations are shown in Fig. 10. In particular, the busy durations are again discrete as in Fig. 9(a) with the components corresponding to the acknowledgement packets, the data packets, and the router's beacons, respectively. The idle durations on the other hand allow for two preliminary conjectures. First, there is a significant component around $0.7\,$ms (corresponding to the effect of the contention window and the DIFS). Second, the tail of the histogram appears to decay slower than exponentially, suggesting that a heavy-tailed distribution might be a good fit.

Given the above observations as well as the standard specifics it makes sense to define the following



(a) Histogram for the busy durations.

(b) Histogram for the idle durations.

Figure 10: Histograms for the UDP traffic scenario (cf. Sec. 4.3.3)

set of states depending on the medium's usage.

DATA The channel is busy due to the transmission of a data packet. The sojourn time in this state is deterministic and amounts to the time required to transmit the 1024 B size packet.

SIFS The channel is idle due to the short inter-frame space required between a data packet and its subsequent acknowledgement. The sojourn time in this state is $10\,\mu$s.

ACK The channel is busy due to the transmission of an acknowledgement packet. The sojourn time is deterministic and amounts to 0.11 ms.

CW The channel is idle but there are primary users contending for the medium. The sojourn time in this state can be (approximately) derived from the standard. We assume a finite support from $[0, 0.7\,\mathrm{ms}]$ (the size of the contention window). The type of the distribution depends on how many terminals are contending for the medium at the same time. Given that we are mainly concerned with a lightly used channel a uniform distribution will turn out to be a good fit.

FREE The channel is idle since none of the primary users has packets to transmit. From the viewpoint of dynamic spectrum access the time spent in this state is essentially defining to what extent the channel can be reused. A generalized Pareto distribution will turn out to be a good fit for the sojourn time in this state.

The SIFS, CW, and FREE state each correspond to an idle medium. In our statistical analysis we will focus on the latter two since the SIFS duration is purely deterministic and too short to be used for dynamic spectrum access (only $10\,\mu s$).

While the histograms depicted in Fig. 10 give a first impression on the distribution of the idle durations, more insight can be gained by looking at the empirical distribution function, which is defined as the fraction of observations smaller than t [16]

$$F_e(t) = \frac{\# i : y_i \leq t}{n}, \tag{7}$$

where $y_i, i = 1, \ldots, n$ correspond to n independent samples. The empirical distribution function is shown in Fig. 11 for several values of the rate parameter σ. We normalized the rate parameter according to the maximum packet rate supported by the WLAN setup, that is $\bar{\sigma} = \sigma/\sigma_{\text{max}}$. We can make two important observations. First, the idle duration (whitespace) decreases with σ. Second, for $\bar{\sigma} \leq 0.5$ we can clearly see that the distribution of the idle times is a mixture of the contention window and the distribution of the truly 'free' channel (note the bend in the curves at 0.7 ms). Furthermore, the vertical line in Fig. 11 illustrates the finite support of the contention window's distribution. We can see that the slope of F_e within that region is approximately constant, suggesting that a uniform distribution as an appropriate fit (this is also suggested by the standard specifics). The tail distribution corresponding to the free channel shows heavy-tailed behavior and will be analyzed in detail in Sec. 4.4.

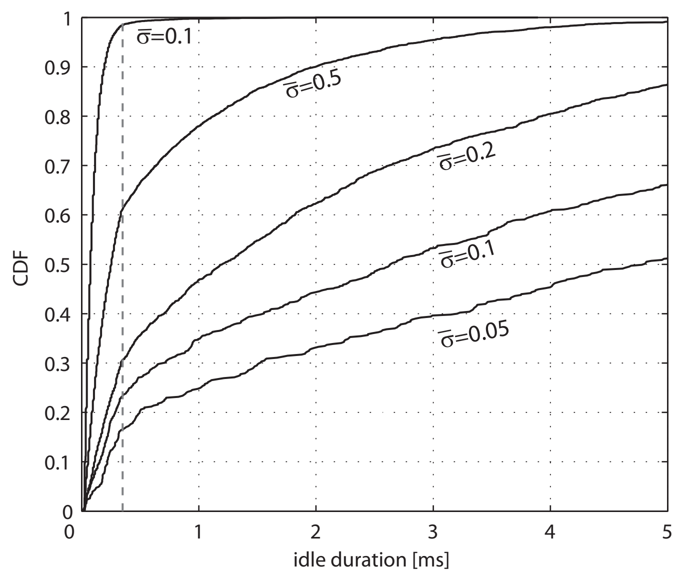


Figure 11: Empirical cdf for the idle durations. The rate parameter is normalized to the maximum traffic load supported by the setup, that is $\bar{\sigma} = \sigma/\sigma_{\text{max}}$.

4.3.4 Representative Traffic Scenarios

In addition to the UDP traffic, we have also looked at a variety of typical traffic scenarios, including file transfers and Voice-over-IP sessions over the WLAN. The resulting empirical cdfs for the idle durations are shown in Fig. 12 and are discussed separately in the following.

First, consider file transfer via secure-FTP from a remote server. In order to collect enough baseband data a text file of approximately 100 kB was transferred 1000 times using a secure-FTP client. The resulting curve shows that there is little remaining whitespace. The effect of the contention window is well-visible by the bend in the empirical cdf at 0.7 ms.

Second, we used D-ITG to generate traffic according to the G.711 codec (used in some VoIP clients). We consider the case of one and three codecs running simultaneously on each of the workstations. The resulting curves show an almost idle channel in the case of one active codec, while the channel appears quite busy in the case of three.

Finally, we used the popular "Skype" client to set up a conference call within the WLAN. A prerecorded audio sample was used to simulate the speech conversation on each of the workstations. The resulting empirical cdf shows that the channel is mostly idle.

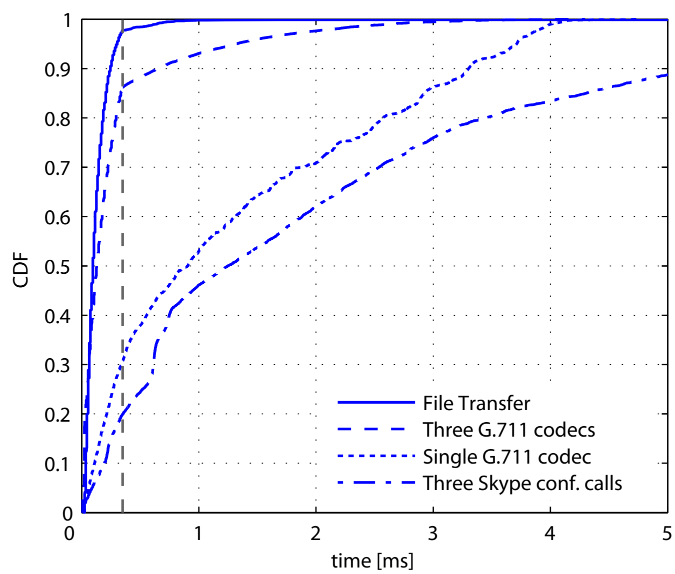


Figure 12: Empirical cdf for FTP and VoIP traffic.

In summary, the empirical cdfs for the traffic scenarios shown in Fig. 12 show a similar behavior compared to the UDP traffic considered before. Specifically, the tails of the distribution appear to be heavy-tailed again, and the influence of the contention window is again visible. We have considered goodness-of-fit for such nonstationary traffic scenarios in [10].

4.4 Semi-Markov Model

The definition of states given in Sec. 4.3.3 allows us to convert the processed measurement data to a sequence of states. Note however, that the states $\textsc{Cw}$ and $\textsc{Free}$ are not observable since we can only detect an idle medium but not conclude whether the system is in either of the states. We shall refer to the lumped version of $\textsc{Cw}$ and $\textsc{Free}$ as the $\textsc{Idle}$ state for brevity.

We have shown previously [11, 10] that a continuous-time semi-Markov model is an appropriate fit. A semi-Markov model can be viewed as an extension to a continuous-time Markov chain (CTMC) with separate statistical specification of the transition behavior and sojourn time within each state [17]. The transition behavior in a semi-Markov process retains the Markovian property with transitions from state i to j occurring with probability p_{ij}. In contrast to a CTMC though, given that a transition $i \to j$ occurs, the sojourn time t in state i (before transitioning to j) can be specified arbitrarily according to some cdf [17]. Recall that in a CTMC the sojourn times in all states need to be exponentially distributed [18].

For specifying the parameters of the semi-Markov process we treat these two parts separately. First, the transition behavior is estimated using the observation sequence and then distributions for the sojourn times are fit to each state.

4.4.1 Estimating transition probabilities

First, we need to find the transition probabilities given the sequence of states obtained by measurement. To this end we can use the well-known maximum likelihood estimator for the transition probability [19, 17]

$$p_{ij} = \frac{n_{ij}}{n_i}, \tag{8}$$

where n_{ij} is the number of transitions $i \to j$ in our observation sequence, and n_i is the total number of state i occurring in the sequence. Using the above estimator, we have shown in [14] that the sequence of states

$$\textsc{Data} \to \textsc{Sifs} \to \textsc{Ack} \tag{9}$$

is essentially deterministic since its transition probabilities are very close to one. In fact, this does not come as a surprise provided that our system is operating at high SNR and the above sequence simply corresponds to a successful transmission. It should be noted that while collisions still occur infrequently in our setup, their effect appears negligible. The transition diagram resulting from the above analysis is depicted in Fig. 13.

4.4.2 Specifying the sojourn times

So far, we have arrived at the transition diagram shown in Fig. 13. Since the transitions $\textsc{Data} \to \textsc{Sifs} \to \textsc{Ack}$ are deterministic and the sojourn time in each of these states is deterministic as well, we only need to fit the sojourn time the $\textsc{Idle}$ state (and the substates $\textsc{Cw}$ and $\textsc{Free}$).

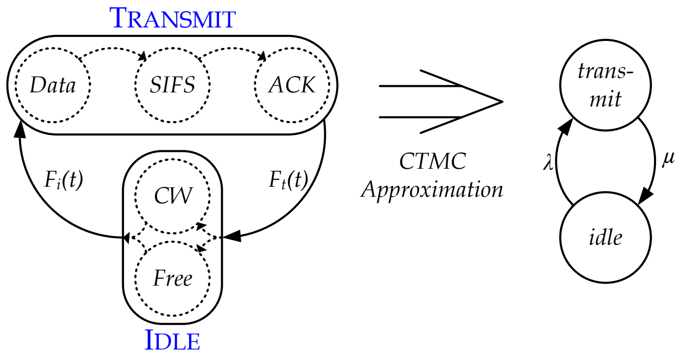


Figure 13: Transition diagram of the semi-Markov model.

The fact that the IDLE state consists of both the CW and the FREE state suggests a mixture distribution,

$$F(t;\boldsymbol{\theta}) = p_c F_c(t) + p_f F_f(t;\boldsymbol{\theta}), \tag{10}$$

where $F_c(t)$ is the cdf of the contention window (assumed uniform on $[0, T_c]$ and $F_f(t;\boldsymbol{\theta})$ denotes the generalized Pareto cdf of the unused channel depending on the unknown parameters $\boldsymbol{\theta}$. The transition probabilities p_c and p_f are also shown in Fig. 13.

In order to simplify the analysis we can exploit some structure in (10). In fact, we know that the support of $F_c(t)$ is limited to $[0, T_c]$ (cf. Sec. 4.3 and Fig. 11, $T_c \approx 0.7\,\text{ms}$). Hence, if we discard all observations $y_i \in [0, T_c]$ (whether or not they are really coming from $F_c(t)$) we are no longer dealing with a mixture but can estimate the parameters of $F_f(t;\boldsymbol{\theta})$ directly.

According to the above we are concerned with estimating the parameters of the generalized Pareto distribution from left-truncated data. Let the truncated data gained by discarding all idle times smaller than the threshold T_c be denoted by $\tilde{y}_i,\ i = 1, \ldots, N_t$. Assuming a generalized Pareto distribution we have the following expression for the pdf

$$f_f(t; k, \sigma) = \frac{1}{\sigma}\left(1 + k\frac{t}{\sigma}\right)^{-1-1/k}, \tag{11}$$

where k denotes the shape, and σ denotes the scale parameter [20]. The cdf is given by

$$F_f(t; k, \sigma) = 1 - \left(1 + k\frac{t}{\sigma}\right)^{-1/k}. \tag{12}$$

Provided that we can only use the left-truncated samples $\tilde{y}_i$ for estimating the parameters the maximum likelihood estimate of the parameter vector $\boldsymbol{\theta} = [k, \sigma]^T$ is given by [21]

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \prod_{i=1}^{N_t} \frac{f_f(y_i;\boldsymbol{\theta})}{1 - F_f(T_c;\boldsymbol{\theta})}, \tag{13}$$

where the term in the denominator is due to the left-truncation of the data at T_c. The maximization in the above formula was performed numerically, using an initial value obtained by a moment estimate for the non-truncated data [20].

Parameter	WLAN traffic load $\bar{\sigma} = \sigma/\sigma_{\mathsf{max}}$ 0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
	CTMC approximation										
λ^{-1} [ms]	15.9	9.10	4.48	2.90	1.98	1.39	1.01	0.74	0.55	0.36	0.21
μ^{-1} [ms]	1.11	1.08	1.05	1.03	1.02	1.03	1.03	1.02	1.03	1.03	1.03
	semi-Markov model										
σ [ms]	18.7	11.3	5.46	3.95	3.09	2.35	1.87	1.48	1.40	0.98	0.04
$k/10^{-2}$	-2.11	-2.50	2.47	1.51	2.61	1.69	7.9	13.3	7.55	11.2	50.1
p_c [%]	13.2	18.3	21.2	30.1	40.0	47.7	58.0	67.2	76.1	84.3	98.8
T_μ [ms]	1.11	1.08	1.05	1.03	1.02	1.03	1.03	1.02	1.03	1.03	1.03

Table 5: Measurement parameters for semi-Markov model and its CTMC approximation.

Given that we have estimated one of the terms in the mixture distribution (10), and realizing that $F_c(t)$ is a uniform distribution on $[0, T_c]$ we can find p_c and p_f, thus fully specifying the desired approximation to the empirical cdf. The fitted distribution as well as the empirical cdf are shown in Fig. 11 for $\bar{\sigma} = 0.05$, $\bar{\sigma} = 0.4$, and $\bar{\sigma} = 0.8$, respectively. The semi-Markov model with mixture distribution shows an excellent fit, which has been validated statistically in [11]. The fitted parameters are shown in Tab. 5 and were used in [1] for defining the simulation parameters.

4.5 CTMC approximation

The semi-Markov model presented above provides for an excellent fit with the empirical data. However, in deriving CMA the semi-Markov model is difficult to analyze since it does not possess the continuous-time Markov property. In order to simplify analysis we consider a CTMC approximation, which corresponds to fitting exponential (instead of mixture) distributions to the distributions of idle and busy periods as shown in Fig. 14. The exponential fit provides a good approximation although it is not strongly validated by statistical measures of fit. Nevertheless, we showed in [1] that CMA derived from the CTMC approximation is quite robust and performs well even if run on data generated using the semi-Markov model.

The parameters of the CTMC model are the exponential parameters λ and μ for idle and busy state respectively, leading to $F_t(t) = 1 - \exp(-\mu t)$ and $F_i(t) = 1 - \exp(-\lambda t)$. The estimated values for both parameters are tabulated in Tab. 5.

References

[1] S. Geirhofer, L. Tong, and B. M. Sadler, "Cognitive Medium Access: Constraining Interference Based on Experimental Models," *submitted to IEEE Journal of Selected Areas in Communications*,

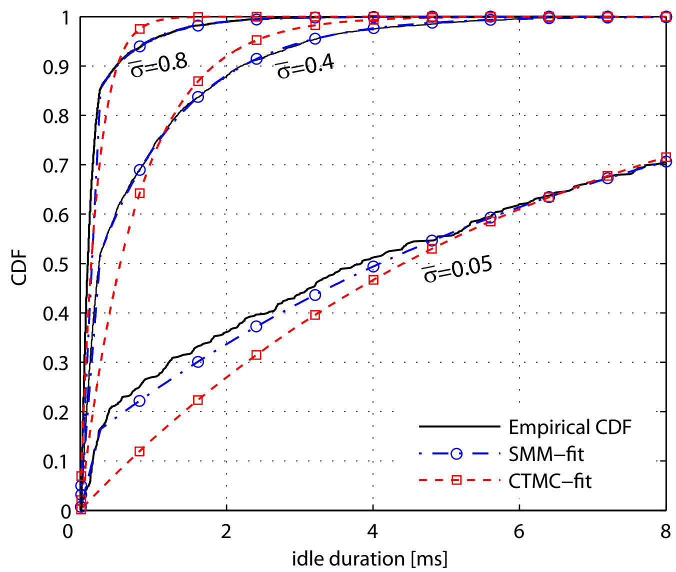


Figure 14: Goodness-of-fit for the semi-Markov model (mixture fit) and its CTMC approximation. The semi-Markov model shows an excellent match and is strongly validated by statistical measures of fit.

Mar. 2007.

[2] Q. Zhao and B. M. Sadler, "Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," *accepted to the IEEE Signal Processing Mag.*, 2006.

[3] ANSI/IEEE Standard 802.11, 1999 Edition (R2003), "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE/SA Standards Board, Tech. Rep., 1999.

[4] Bluetooth Special Interest Group, "Specification of the Bluetooth System," Nov. 2004.

[5] ANSI/IEEE Standard 802.11b-1999 (R2003), "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-Speed Physical Layer Extension in the 2.4GHz band," IEEE SA Standards Board, Tech. Rep., 1999.

[6] Golmie, N. and Van Dyck, R. E. and Soltanian, A. and Tonnerre, A. and Rébala, O., "Interference Evaluation of Bluetooth and IEEE 802.11b Systems," *Wireless Networks*, vol. 9, no. 3, pp. 201–211, May 2003.

[7] A. Kamerman, "Coexistence between Bluetooth and IEEE 802.11 CCK. Solutions to avoid mutual interference." Lucent Technologies, IEEE P802.11 Working Group Contribution IEEE P802.11-00/162, July 2000.

[8] J. Zyren, "Extension of Bluetooth and 802.11 Direct Sequence Interference Model," Harris Semiconductor, Study Group Contribution IEEE 802.11-98/378, Nov. 1998.

[9] N. Golmie, N. Chevrollier, and O. Rebala, "Bluetooth and WLAN Coexistence: Challenges and Solutions," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 22–29, Dec. 2003.

[10] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic Spectrum Access in the Time Domain: Modeling and Exploiting Whitespace," *submitted to the IEEE Communications Magazine*, 2006.

[11] ——, "Dynamic Spectrum Access in WLAN Channels: Empirical Model and Its Stochastic Analysis," in *Proc. First International Workshop on Technology and Policy for Accessing Spectrum*, 2006.

[12] Agilent Technologies, "Agilent 89611A 70MHz IF Vector Signal Analyzer," data sheet, Oct. 2001.

[13] S. Avallone, A. Botta, D. Emma, S. Guadagno, and A. Pescape, "D-ITG V.2.4 Manual," University of Napoli "Federio II", Tech. Rep., Dec. 2004.

[14] S. Geirhofer, L. Tong, and B. M. Sadler, "A Measurement-Based Model for Dynamic Spectrum Access," in *Proc. IEEE Conference on Military Communications (MILCOM)*, 2006.

[15] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer-Verlag, 1994.

[16] R. B. D'Agostino and M. A. Stephens, *Goodness-of-fit techniques*. Marcel Dekker, Inc., 1986.

[17] S. M. Ross, *Applied Probability Models with Optimization Applications*. Dover Publications, 1970.

[18] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw Hill Publishing Company, 2002.

[19] P. Billingsley, "Statistical Methods in Markov Chains," *The Annals of Mathematical Statistics*, vol. 32, no. 1, pp. 12–40, Mar. 1961.

[20] S. Kotz and S. Nadarajah, *Extreme Value Distributions. Theory and Applications*. Imperial College Press, 2000.

[21] A. C. Cohen, *Truncated and Censored Samples. Theory and Applications*. Marcel Dekker, Inc., 1991.