ATCZ175 InterOP PROJECT

Measurement Campaign Report

Institute of Electrodynamics, Microwave and Circuit Engineering
Technische Universität Wien

Advisor:
Assoc. Prof. Dipl.-Ing. Dr.techn. Holger Arthaber

by
Proj. Ass. Dipl.-Ing. Christian Spindelberger

October 9, 2019

Contents

Abbreviations

AWGN additive white Gaussian noise

ED energy detector

ISM industrial, scientific, and medical

MF matched filter

NF noise figure

PDF probability density function

Q-Q quantile-quantile

VSA vector signal analyzer

WLAN wireless local area network

1 Introduction

In order to describe interference effects in the industrial, scientific, and medical (ISM) band, a measurement campaign was held in June of 2018. The goal of this campaign was to record traffic in wireless local area network (WLAN) channels localized in the 2.4 GHz and 5 GHz band, obtaining parameters such as traffic load and power level distributions. Further investigations will show if it is possible to abstract such a scenario by using different kinds of interference sources.

Section 2 introduces the measurement scenario, setup, and how the measured data is recorded. Subsequently, the observed records are analyzed utilizing several techniques, concerning typical traffic parameters (Section 3). Firstly, a rudimentary model is considered, describing the main issues regarding detection problems. Based on this point of view, a more complex method called energy detector (ED) is presented. This technique offers the opportunity to detect transmitted frames, characterizing them by their duration and mean power. This section is completed by determining the communication standard of detected events. At last, results utilizing the explained techniques are given in Section 4.

Figure 1: Lecture hall (left), measurement equipment (right)

2 Setup and Location

The measurement campaign took place during a lecture called "Computer supported Japanese" at the TU Wien. Due to a large number of attendees, a high device density of laptops, smart phones, and smart wearables, based on WLAN and BLE, was expected. Consequently, this scenario offered the opportunity to record multiple users, and an above average traffic load. Furthermore, the existing devices were distributed over the whole lecture hall, resulting in a varying power level of the received frames. Therefore, an omni-directional antenna with broadband receive characteristics was used to cover the desired

frequency range from 2.4 GHz to 5 GHz, capturing ISM band traffic. The described scenario is depicted in Figure 1. Because of space limitations, the measurement equipment was placed outside of the lecture hall. The antenna position is marked by the white arrow in the left subfigure. As the remaining equipment was placed outside, a 30 m long cable was utilized to connect the antenna.

Receiver Chain

Since the utilized RF cables introduce losses and, consequently, increase the noise figure (NF) as well, an amplifier chain was utilized to conquer these effects.

According to Figure 2, a low-noise amplifier was connected directly to the connector of the antenna to establish a calculated low total NF of ~ 1.53 dB. While the antenna was mounted on a tripod, an additional cable down to the floor was utilized to connect another power amplifier. The overall gain of these two amplifiers overcomes the insertion losses induced by the long transmission cable to the measurement equipment. Details about the used components can be found in Table 1.
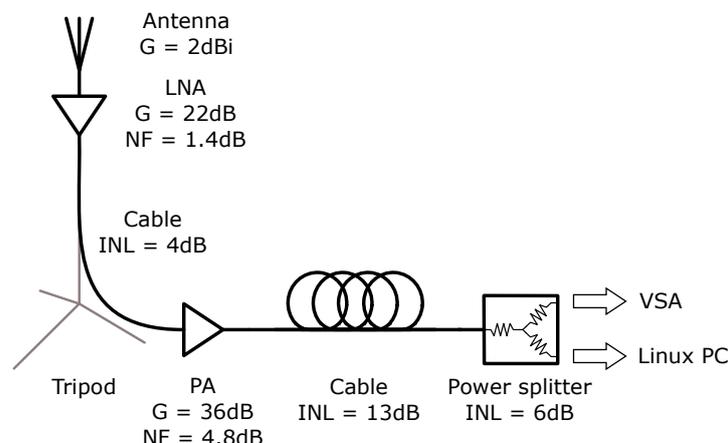
Figure 2: Receiver chain of the measurement setup, specified for 2.4 GHz (G...Gain, NF...Noise figure, INL...Insertion loss)

At the end of the 30 m long cable, a power splitter connects a vector signal analyzer (VSA) and a Linux PC to the receiver chain. The VSA records the received data with a sampling rate of 25.6 MSa/s up to a maximum length of approximately 10 s. In parallel, the PC demodulates all received WLAN frames and records details about each, which can be analyzed with software tools, such as Wireshark[1]. It will be verified if it is possible to

---

[1]Wireshark is a network protocol analyzer tool. It enables decoding frames on a microscopic level.

replace expensive measurement equipment (VSA) with a Linux PC for data recordings. At the beginning of the measurement campaign it was evaluated which WLAN channels were occupied most frequently. It turned out that channel 1 ($f_c = 2.412\,\text{GHz}$) and channel 36 ($f_c = 5.18\,\text{GHz}$) were the best choice in terms of traffic load. During the whole measurement scenario 40 records in the $2.4\,\text{GHz}$ and $5\,\text{GHz}$ band have been made.

Component	Manufacturer	Description
Antenna	Huber und Suhner	SENCITY Omni-S 1399.17.0224
LNA	Mini-Circuits	ZX60-83LN-S+
PA	Mini-Circuits	ZVE-8G
Power splitter	Mini-Circuits	ZFRSC-183+
Filter	K&L	CAV-01311
WLAN module (PC)	Atheros	WLE900VX
VSA	Keysight	UXA N9040B

Table 1: List of the utilized components corresponding to Figure 2
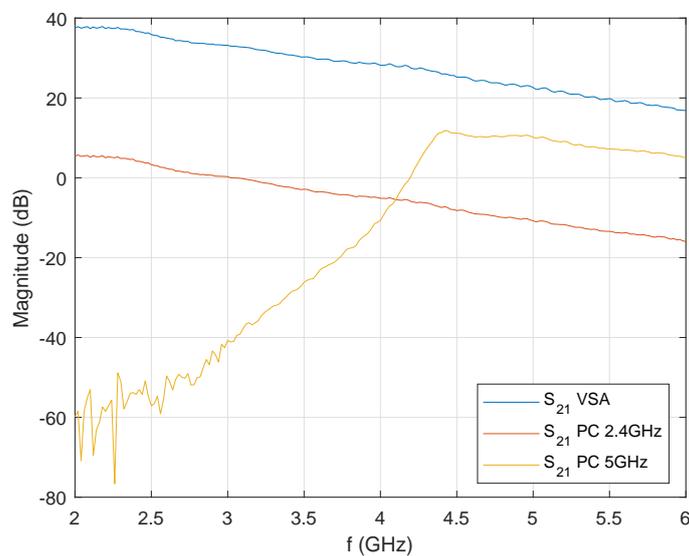
Figure 3: Gain of the receiver chain between antenna and: VSA (blue), PC 2.4 GHz (red), PC 5 GHz (yellow)

Regarding WLAN, it is possible to record live data exchanges, revealing details about MAC and PHY properties.

The receiver chain may cause blocking effects of the WLAN module implemented in the Linux PC. Therefore, an additional filter for measurements in the 5 GHz band was added. One can see the respective S-parameters in Figure 3. In order to further adjust the received power levels to an appropriate value, additional attenuators were applied.

3   Data Analysis

The VSA saves the data as a complex baseband signal with a sample rate of 25.6 MSa/s. The highest expected bandwidth is 22 MHz (WLAN DSSS), therefore, the whole measurement data is further processed by a filter with the desired frequency range. In order to create a statistical model of the observed scenario, the baseband signal is investigated in terms of channel on-/off-times and power levels. On- and off-times are specified as immediately consecutive samples with the respective decision state. The power levels are obtained by the mean value within the estimated on-times, recalculated by the measured S-parameters from Figure 3.
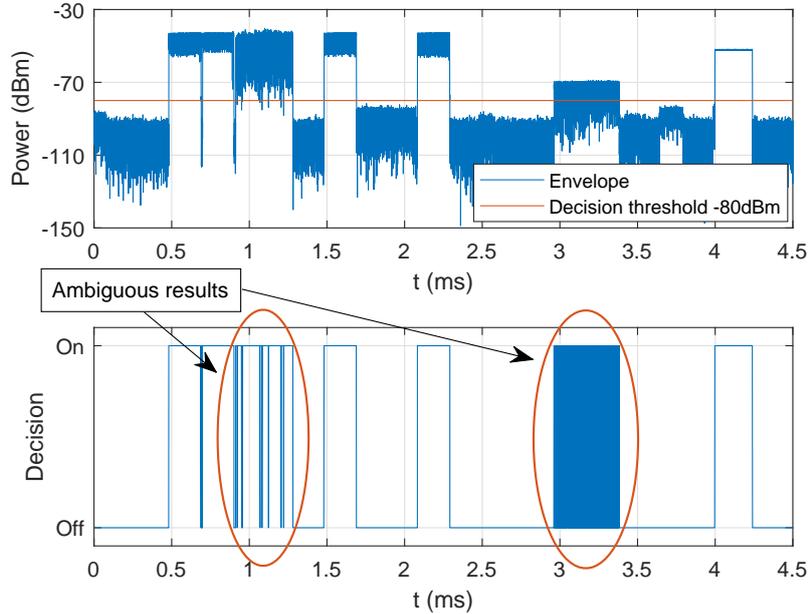
Figure 4: ISM band recording, WLAN channel 1, BW = 22 MHz: envelope (top), decision according to power level detection (bottom)

The detection of on- and off-times is a nontrivial problem for various reasons. First of all, the received signal modulation is unknown. For instance, OFDM has a noise-like amplitude

4

distribution. Because of strong amplitude variations, it becomes rather difficult to detect an OFDM modulated signal without interruptions by simply setting a power threshold. Figure 4 depicts the power envelope of a short sequence from the recorded ISM band data. Obviously, the envelope shows a high dynamic range for some frames, which makes them hard to detect correctly, even when the mean SNR is high enough ($> 10\,\mathrm{dB}$). The lower subfigure of 4 emphasizes the detection problem by a power threshold of $-80\,\mathrm{dBm}$. In this case, the mean SNR is $\geq 20\,\mathrm{dB}$ (theoretically calculated noise power of $-100.6\,\mathrm{dBm}$ for a bandwidth of $22\,\mathrm{MHz}$) and the decision is still ambiguous even for frames with a high power level. These uncertainties result in a wrong on- and off-time characterization. In the following, methods will be discussed to overcome these problems.

3.1   Time-Quantized Analysis

A simple approach to omit the detection problem is to slice the recorded power envelope in time domain into equally stepped intervals and determine the mean power of the respective slots. This model circumvents detection problems by neglecting an estimation of start- and end point of an event. Hence, it is possible that multiple interferer types occur in one slot, making it impractical to determine the respective communication standard. Therefore, only power levels are investigated.

Before proceeding with the data analysis of this technique, a statistical concept will be introduced. Quantiles are often used to describe random data sets by arranging the obtained samples according to their size. One popular quantile, for instance, is the median, dividing the sorted data set into two parts with an equal amount of samples ($50\,\%/50\,\%$). The separation of the data into groups of equal size can be arbitrarily extended. It is possible to examine similarities of two data sets regarding their probability distributions with a graphical method called quantile-quantile (Q-Q) plot. In order to obtain such a graph, the quantiles of the two sets are assigned to separate axes. Furthermore, the quantile separation has to be determined. A point on the Q-Q plot corresponds to an intersection between the same quantiles of the x- and y-axis. The points will lie approximately on the line $y = x$ if the two data sets being compared have the same probability density function (PDF).

Figure 5 depicts the power level distribution of two records from the $2.4\,\mathrm{GHz}$ ISM band with a reference step size of $1\,\mathrm{\mu s}$. The corresponding Q-Q plots show a comparison between different step sizes ($100\,\mathrm{\mu s}$ and $1\,\mathrm{ms}$) and the reference. Since the preamble of a WLAN frame lasts for more than $16\,\mathrm{\mu s}$, a resolution of $1\,\mathrm{\mu s}$ is assumed to be fine enough to neglect clipping effects of high power levels through averaging.
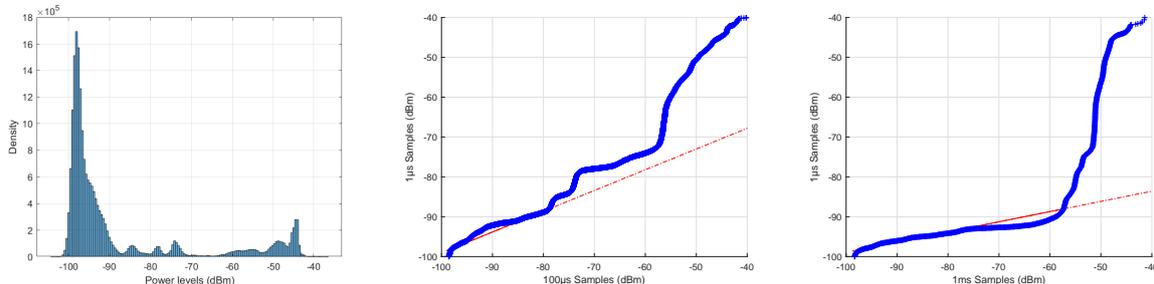
Figure 5: Histogram of measured power levels with a step size of $1\,\mu s$ (left), respective Q-Q plots with longer time intervals of $100\,\mu s$ (middle) and $1\,ms$ (right)

The Q-Q plots give an estimate on how good the power levels, averaged over longer step sizes, fit to the ones with a step size of $1\,\mu s$. If the plot appears to be linear, the compared samples tend to have the same distribution. It is noteworthy that the quantiles for a step size of $100\,\mu s$ deviate already for power levels greater than $-75\,dBm$, while the samples corresponding to the step size of $1\,ms$ seem to be linear up to a power of $-60\,dBm$. Furthermore, one can see that the red linear line has a smaller slope than $y = x$. This means, that the quantiles of the x-axis have a higher density of larger power levels than the quantiles of the y-axis ($1\,\mu s$). Hence, the power levels do not suffer from clipping effects caused by averaging for bigger step sizes, as it was expected.

3.2 Energy Detection

The previously presented technique is a simple approach to describe the power level characteristics of the captured ISM band records. Unfortunately, this model does not consider frame lengths and off-times of the channel. Thus, a more complex method must be utilized to conquer the explained detection problems.

Because of the large amount of different communication standards using the ISM band, it would cause an intolerable effort to demodulate each detected event and determine the frame length. In addition to this, interferers such as microwave ovens cannot be detected by this method. Therefore, another concept, optimized for single antenna receivers called ED is introduced. This so-called ED calculates the received energy over a time interval T and utilizes an appropriate threshold ξ to decide if a signal is present. The main parameters which determine the performance of the ED are the decision threshold, the number of samples ($N = Tf_s$), and the estimated noise power [1, 2].

The ED just decides if a signal is present or not. Therfore, this scheme can be broken down into a binary hypothesis-testing problem [1]:

Hypothesis 0 ($\mathcal{H}_0$:) *signal is absent*

Hypothesis 1 ($\mathcal{H}_1$:) *signal is present*

Hence, the received complex baseband signal is divided into two states over time index i (n_i...noise, x_i...transmitted signal, y_i...received signal):

$$y_i = \begin{cases} n_i & : \mathcal{H}_0 \\ x_i + n_i & : \mathcal{H}_1 \end{cases} \tag{1}$$

In order to be independent of absolute power levels, the ED estimates the mean SNR over time. Equation 2 defines the calculation of the ED by averaging the signal power (y_i^2) over N samples. Since the noise power is not perfectly known, a maximum likelihood estimation of $\hat{\sigma}^2$ is utilized over M samples. In conclusion, the ED is a sliding window that calculates the mean SNR and decides according to a given threshold (ξ) if a signal is present or not [2].

$$\Lambda(y) = \frac{1}{2\hat{\sigma}^2 N} \sum_{i=0}^{N-1} |y_i|^2 \overset{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \xi, \qquad \hat{\sigma}^2 = \frac{1}{2M} \sum_{i=0}^{M-1} |n_i|^2 \tag{2}$$

The desired ED performance will be described in terms of false alarm probability (P_f) and detection probability (P_d). P_f is the probability of falsely detecting a signal while it is actually absent ($\mathcal{H}_0$) and P_d is the probability of correctly detecting a present signal ($\mathcal{H}_1$). Assuming an additive white Gaussian noise (AWGN) scenario and a Gaussian signal, e.g., OFDM, to be detected, these probabilities may be expressed for large N and M by [2]:

$$P_f \cong Q\left(\frac{\xi - 1}{\sqrt{\frac{N+M}{NM}}}\right), \qquad P_d \cong Q\left(\frac{\frac{\xi}{1+\text{SNR}} - 1}{\sqrt{\frac{N+M}{NM}}}\right). \tag{3}$$

Eliminating the threshold ξ through the two equations from 3, it is possible to calculate the minimum achievable SNR, satisfying the desired P_f and P_d:

$$\text{SNR}_{\min} = \frac{1 + Q(P_f)^{-1}\sqrt{\frac{N+M}{NM}}}{1 + Q(P_d)^{-1}\sqrt{\frac{N+M}{NM}}} - 1. \tag{4}$$

The solution for the optimum decision threshold ξ_{opt}, meeting the desired probabilities (P_f, P_d), is then given by utilizing $\text{SNR}_{\min}$ in equation 2 (right):

$$\xi_{\text{opt}} = \left(Q(P_d)^{-1}\sqrt{\frac{N+M}{NM}} + 1\right)(1 + \text{SNR}_{\min}). \tag{5}$$

Figure 6 again shows the ISM band sequence from Figure 4, now based on the ED. Because of averaging effects of the sliding window and appropriate calculation of a threshold, the

decision states are unambiguous. If the size of the sliding window (N) is too large, the off-times of the channel tend to grow over and it is not possible anymore to define start and end points of two frames that are closer to each other than the defined window length. In the IEEE 802.11n standard, the shortest IFS is the so-called RIFS, which is $2\,\mu s$ long. Therefore, the sliding window size is set to $t = 2\,\mu s$. At a sample rate of $f_s = 25.6\,\text{MSa/s}$, this equals a number of $N = f_s t \approx 52$ samples.
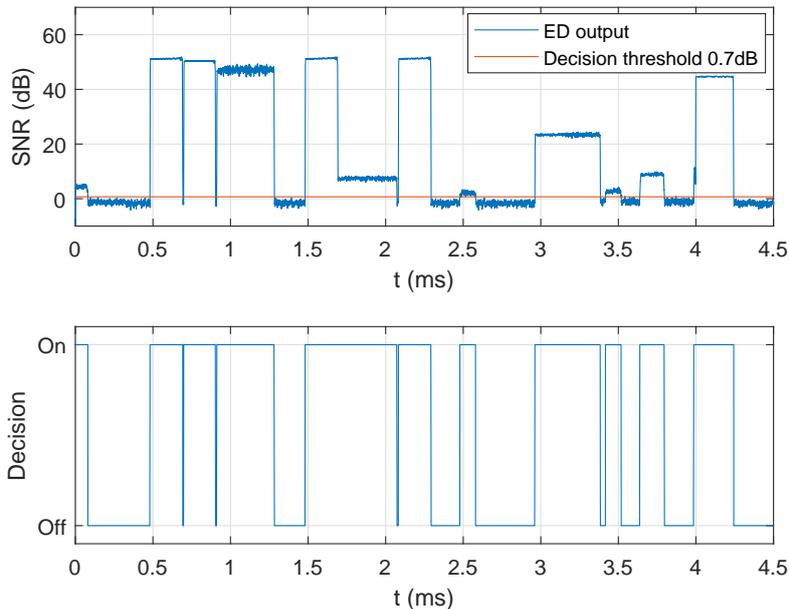
Figure 6: ED outptut in term of SNR over time of the sequence from 4 (top), decision according to threshold (bottom): $f_s = 25.6\,\text{MSa/s}$, $N = 52$, $M = 10^5$, $P_f = 0.1$, $P_d = 0.9$, $\xi = 0.7\,\text{dB}$

It is necessary to define a set of noise samples to calculate the ED output of Equation 2. Examining Equation 4 indicates that, for the desired probabilities P_f and P_d, the parameters N and M are available to optimize the $\text{SNR}_{\min}$. As the window size is $N = 52$ samples long, the noise sample size M can be chosen arbitrarily large to make the decision threshold as low as possible. One might think that the threshold can be made infinitely low. However, because of uncertainties caused by parameter estimation methods, the minimum achievable SNR becomes asymptotic for large values of M and N. Hence, a signal-free sequence in the recorded baseband data with $M = 10^5$ samples led to a satisfying noise power estimation ($\hat{\sigma}^2$). It must be mentioned that the noise power estimation suffers from several effects and influences the performance of the ED. The four most important factors are [2]:

1. temperature variation

2. change in low-noise amplifier gain due to thermal variations

3. initial calibration error

4. presence of interferers

It is possible to overcome the first three mentioned issues by using a sufficient amount of samples for noise power estimation. Nevertheless, the influence of interferers close to the noise floor, which makes them hard to detect, cannot be canceled out by increasing the amount of noise samples. Consequently, it must be taken into account that the ED performance may deviate from theoretical calculations.

3.3   Classifying Frames

With the ED, it is possible to identify channel occupations independent of standard and modulation. This approach can be further extended by classifying the detected frames in order to draw conclusions about different modulation schemes influencing the behavior of interfering signals. Hence, the amount of existing standards (BLE, WLAN-DSSS, WLAN-OFDM) in the captured ISM band data is estimated. This section treats the detection of WLAN frames with DSSS and OFDM modulation. The remaining signals, such as BLE, are classified as general interference.

WLAN DSSS

The DSSS modulation is only used in the $2.4\,\text{GHz}$ ISM band corresponding to the IEEE 802.11b standard. Even though this technique is obsolete by now, it is still frequently used to provide compatibility between old and new WLAN-standard revisions.
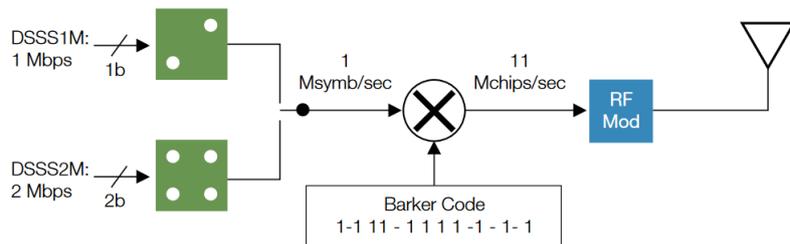
Figure 7: WLAN DSSS modulation scheme [3]

The legacy preamble has a bandwidth of $1\,\text{MHz}$ and is broadened by multiplying the baseband data with a spreading sequence up to $11\,\text{MHz}$. Figure 7 depicts the modulation

technique in detail. The spreading sequence is realized with a Barker code and ensures that the autocorrelation function has a minimum at off-peak values with the following BPSK symbols: $g_i = [1, -1, 1, 1, -1, 1, 1, 1, -1, -1, -1]$. Since the sequence stays the same for the whole preamble, it is possible to detect DSSS-modulated frames by exploiting this repeated structure. This can be realized, for instance, with a matched filter (MF) ($h_i = g^*_{N-i}$, N...impulse response length), which maximizes the SNR for the optimum decision point. The output of such an MF according to a DSSS preamble is given by Figure 8(a). One can see that the peaks are spaced equidistantly by $1\,\mu s$. Consequently, WLAN frames, which are modulated according to the IEEE 802.11b standard, can be classified with the explained timing constraint. It must be mentioned that channel variations may distort the MF output and peak distances deviate about a few samples.

WLAN OFDM

State-of-the-art WLAN systems utilize OFDM as a modulation scheme. The bandwidth is allowed to occupy a range of up to $160\,\mathrm{MHz}$ (IEEE 802.11ac). Such broadband transmissions are required for high data rate applications. As IoT focuses on wireless sensors and wearables, which are satisfied with low data rates, a single WLAN channel with a bandwidth of $20\,\mathrm{MHz}$ is analyzed.

(a) WLAN DSSS
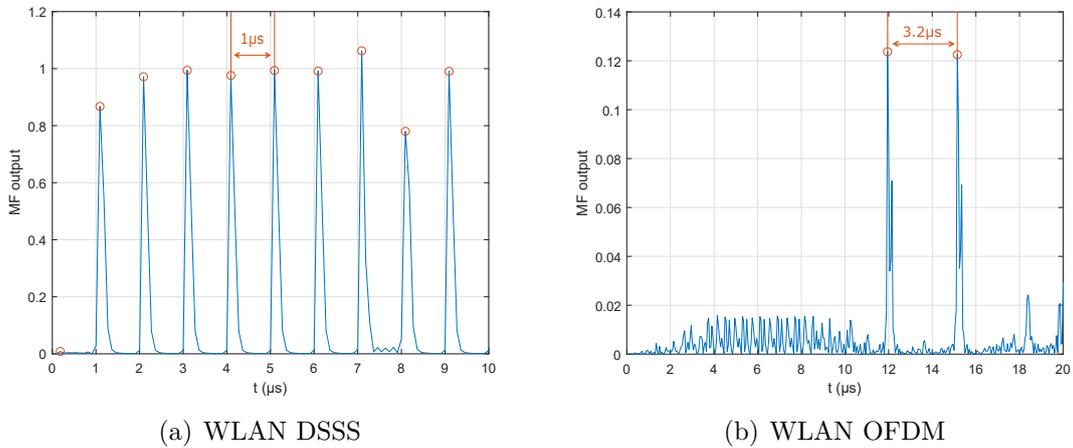
(b) WLAN OFDM

Figure 8: Matched filter output of corresponding WLAN preambles

The legacy preamble of every OFDM-modulated WLAN frame consists of a repeated structure, which makes them easy to detect. Figure 8(b) depicts the matched filter output of an OFDM-modulated WLAN frame from the measurement campaign, correlated with the long training sequence, separated by a time interval of $3.2\,\mu s$. Hence, a timing constraint can be set to identify such frames. Obviously, the main two peaks have a side lobe each,

which presumably appear due to multipath components. For low-SNR frames, the main peaks may not be detected correctly. Thus, a timing tolerance has to be taken into account. Further investigations showed that an empirically defined deviation of $10\,\%$ led to an appropriate functionality of the MF detection scheme. Therefore, an absolute deviation of up to $100\,\text{ns}$ for DSSS and $320\,\text{ns}$ for OFDM is accepted.

4 Results

Using the techniques which were presented in the previous section, the ISM band data is analyzed according to the ED scheme. In addition to this, the results are interpreted and parameter fittings of respective probability densities are investigated.
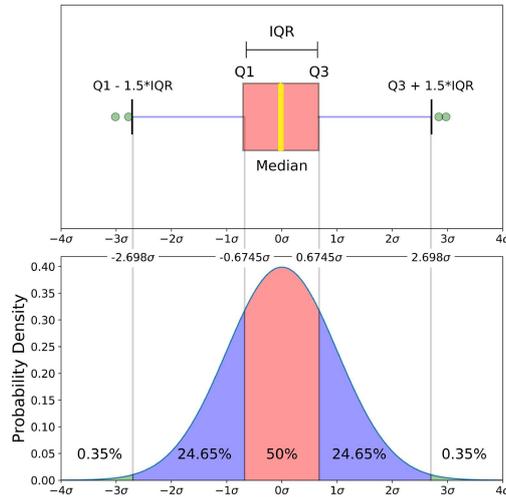
Figure 9: Example of a boxplot (top) according to a normal probability density function (bottom) [4]

The boxplot (Figure 9) is another statistical method for graphically describing numerical data utilizing quantiles. They are defined by the median (yellow), interquartile range $\text{IQR} = Q_3 - Q_1$ (box in red) and the upper $(Q_3 + 1.5 \cdot \text{IQR})$ and lower whisker $(Q_1 - 1.5 \cdot \text{IQR})$. Samples detected outside the range defined by the whiskers, which is $99.3\,\%$ for normal distributions, are outliers (green). Since the example relates to a normally distributed function, which is symmetric, the median is placed inside the IQR symmetrically as well. Furthermore, the length of the IQR indicates how the observed data is concentrated. For large IQRs, the samples are widespread and vice versa for short ranges. In conclusion, the boxplot offers an efficient way to describe data distributions graphically.

In this section, two 2.4 GHz ISM band records of WLAN channel 1 are considered for explaining the process of data analysis. Because of several aspects concerning interference characterization, the two recordings with the highest traffic load are analyzed (available on InterOp homepage). The data has been analyzed with the following parameters: $N = 52$, $M = 10^5$, $P_d = 0.9$, $P_f = 1 - P_d = 0.1$, $\xi = 0.7\,\text{dB}$. Through the false alarm probability of $P_f = 0.1$, outliers are likely to appear. In order to maintain a more stable decision output of the ED, an outlier detection has been implemented. The minimum off-times regarding WLAN systems last for 2 µs (RIFS). Because of several effects, such as channel variations, this minimum idle time may be violated. Thus, all off-times between two on states, which are smaller than 1 µs, have been dedicated to be on-times. Examining the preamble of BLE and WLAN standards indicates that preambles last for more than 8 µs. Therefore, all on-times which last shorter than 8 µs have been set to off state.

(a) Linear scale (b) Logarithmic scale
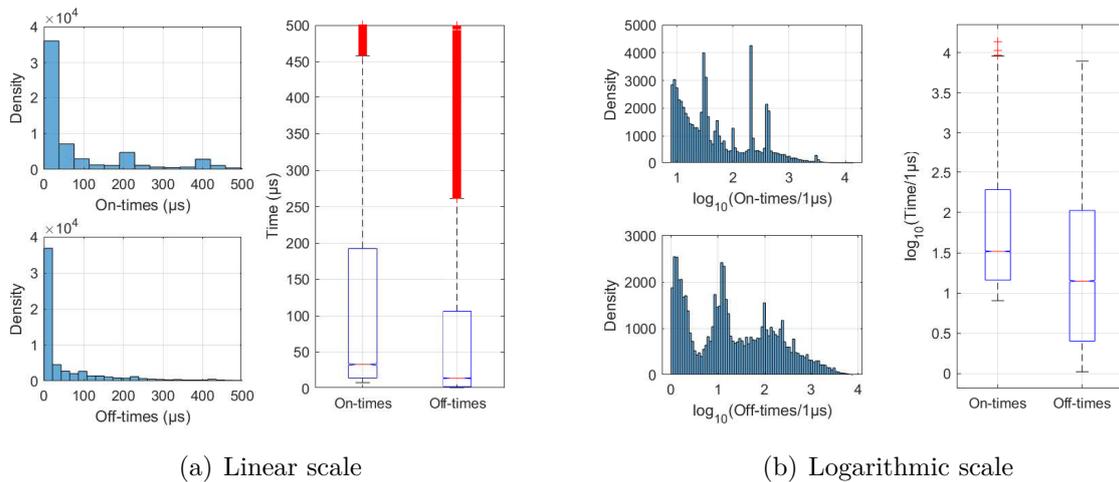
Figure 10: Histograms and boxplots of on- and off-times corresponding to their scale

Figure 10(a) depicts the histogram of the on- and off-times and the related boxplot. Samples that are not within the whiskers range are outliers and marked in red. One can notice the huge amount of outliers and that the median values of on- and off-times are between 15 µs and 33 µs. Furthermore, the histograms show a high density for low values and a nonsymmetric shape. In addition to this, the median value is close to the lower quantile Q_1 and the IQR is relatively large. Consequently, the data is spread over a wide range, which is also reflected in the big amount of outliers. Due to these properties, parametric distributions like gamma or beta may not fit sufficiently to cover the whole data range.

Since the on- and off-times spread over a wide range, it seems natural to transform the data logarithmically in order to gather the samples closer to each other. Figure 10(b)

shows the same data from 10(a), in a logarithmic scale. As expected, the data samples are closer to each other and the boxplot shows only two outliers for on-times. Nevertheless, it still must be clarified if it is possible to fit a distribution to the current histograms. Obviously, the structure of the densities is asymmetric and sharp spikes appear over the whole range. As a consequence, classical parametric distributions like Gaussian or Laplace will not fit. Therefore, a nonparametric kernel distribution is utilized to create PDFs of the respective random variables. The main parameters to describe such a PDF are the type of the smoothing function, e.g., Gaussian, triangle, and the width (variance) influencing the smoothness of the resulting approximation. Equation 6 describes the kernel estimator function [5]:

$$\hat{f}(y) = \frac{1}{n} \sum_{i=1}^{n} w(y - y_i; h). \tag{6}$$

The probability density w is the so called smoothing function and the corresponding variance is defined by parameter h. The kernel estimator function will be explained by a short example.
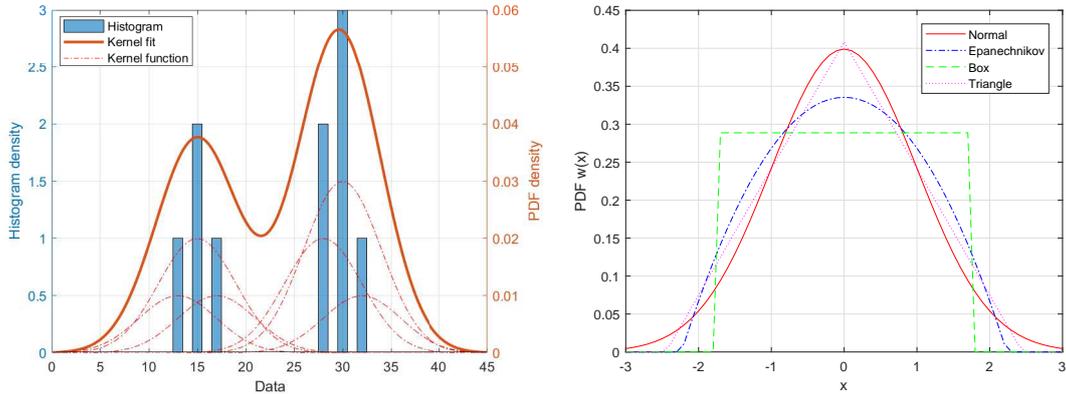
Figure 11: Kernel fit example: densities (left), smoothing functions (right)

Figure 11 depicts the histogram of some test data on the left. The kernel function is realized as a normal distribution with a variance of $\sigma = 4$. According to Equation 6, at every sample, a smoothing function is placed in red dashed lines. The overall sum of this ensemble, normalized by the amount of data points n, is the estimated kernel fit $\hat{f}(y)$, marked in red. On the right of Figure 11, a variety of smoothing functions is presented. Typically, the normal distribution is utilized, but for some use cases other shapes, such as the box, triangle, or Epanechnikov, yield better results.
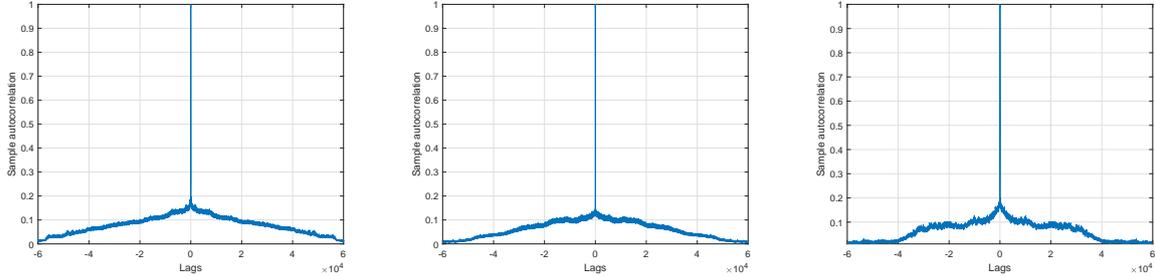
Figure 12: Autocorrelation of on- (left), off-times (middle), and power levels (right)

Before obtaining random variables for on-/off-times and power levels, dependencies in terms of periodicities must be considered. In order to describe such a behavior, the autocorrelation function can be examined through identifying periodic peaks. Figure 12 depicts the respective autocorrelation functions in the linear regime. As no periodic peaks appear in the plots, the data sets are assumed to be independently distributed. One must note that correlations between power levels and corresponding on-times have been neglected.
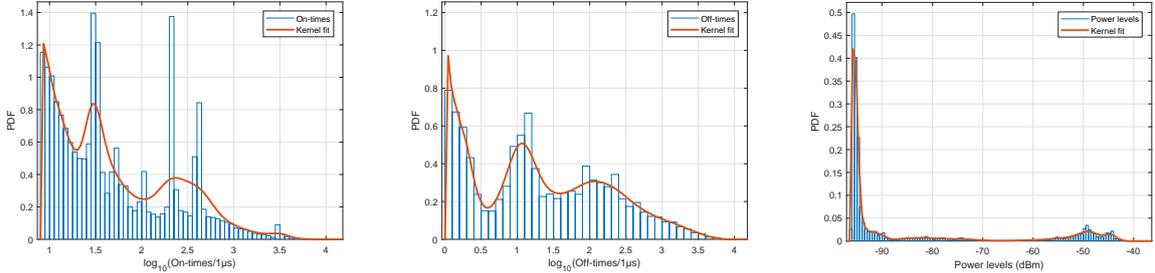
Figure 13: Parameter fitting: on-times (left), off-times (middle), power levels (right)

Since the data densities have been described by their respective autocorrelation functions, random variables can be created by utilizing kernel distributions. It turned out that the Gaussian smoothing function yields the best results for parameter fittings. Furthermore, this technique offers the opportunity to specify a desired range of the PDF. Figure 13 depicts the fitted PDFs of on-/off-times and power levels. It must be mentioned that the fitting strongly depends on the histogram resolution and smoothing function variance. If the bin width is chosen too large, significant values such as maxima and minima get lost. For very high resolutions, it becomes a challenging task to fit strong variations of data densities by a proper variance. Consequently, the bin width must be chosen carefully, but a lot of different rules for calculating the resolution exist. In this work, the Freedman-Diaconis rule led to satisfactory results and will be explained briefly [6]. The bin width

is defined by the interquartile range (IQR) of the observed data and is normalized by the amount of samples (n) as:

$$\text{Bin resolution} = 2\frac{\text{IQR}}{\sqrt[3]{n}}. \tag{7}$$

The interquartile range relates to the boxplot representation, explained previously. It is defined by the distance between the upper- and lower quantile, surrounding the median and covering $50\,\%$ of data (IQR $= Q_3 - Q_1$).

Unfortunately, the on-times density (Figure 10(b)) shows significant spikes through the data which appear noncontinuously. Even for a small variance, these spikes cannot be approximated ideally. One big advantage of the kernel fit is the opportunity to define the data range for parameter fitting. Unlike parametric functions, such as the beta or gamma function, exclusion rules can be set arbitrarily. Table 2 refers to the details about the bandwidth of the smoothing function and data ranges located on the x-axis of the respective histograms. Due to simulation settings, utilizing all recorded data sets would cause an intolerable increase of time intervals for measuring interference perturbations. Hence, only two data sets with the highest traffic density, using WLAN channels 1 and 36, are investigated.

	Channel 1 2.412 GHz		Channel 36 5.18 GHz	
Parameter	Bandwidth	Range	Bandwidth	Range
On-times ($\log_{10}(t/1\,\mu s)$)	0.18	$0.9 < x < 4.15$	0.15	$0.9 < x < 4.1$
Off-times ($\log_{10}(t/1\,\mu s)$)	0.18	$0 < x < 3.95$	0.11	$0 < x < 4.2$
Power levels (dBm)	0.13	$-96 < x < -37$	0.45	$-96 < x < -37$

Table 2: Kernel fit parameters: x-axes ranges (investigated x-axis range of the respective histogram) are in logarithmic scale

It has been mentioned previously that it is of interest, which modulations and standards are involved in the captured data. With the detection methods from Section 3.3, it is possible to distinguish between WLAN modulations (which are OFDM and DSSS). Otherwise, the detected frame is classified as general interference, such as BLE, Sigfox, or HiperLAN. Although these presented correlation techniques work out well, it may happen that multiple standards are detected in a single on-time, which will be classified as WLAN if a valid

preamble is detected. In addition to this, the observed statistical model utilizes the mean power of the whole frame. It probably occurs that the power level changes during a detected event for several reasons. Due to superposition of multiple users (hidden node) or, for instance, a BLE frequency hopping sequence, the power levels may change rapidly. As a consequence of averaging, strong power level variations will be clipped. Table 3 states the detailed amount of involved modulations and how many on-times have been detected in channels 1 and 36. The power threshold is empirically set to $-90\,\mathrm{dBm}$ to ensure a proper functionality of the frame classification techniques. This means that frames below this power level are not taken into account. It is noteworthy to see that the amount of IEEE 802.11b frames (WLAN DSSS) is still high compared to newer revisions, which are using OFDM. In addition to this, the percentage of interferers is in the same range as for the 5 GHz band. It must be emphasized that the true amount of interferers will be higher. Every on-time detected with a valid WLAN preamble is classified as such. Hence, frames will only be identified as an interferer when they are free of WLAN signals.

	Channel 1 2.412 GHz	Channel 36 5.18 GHz
WLAN DSSS	38.41 %	0 %
WLAN OFDM	52.4 %	92.74 %
Interference	9.19 %	7.26 %
Recordings	2	2
On-times	19,360	24,740

Table 3: Frame classification: power threshold $> -90\,\mathrm{dBm}$

In Section 2, the attempt of replacing a VSA by a Linux PC system is discussed. It turned out, for several reasons, that this is not feasible. The Linux system demodulates the received data by a WLAN card, which is not able to detect BLE or other kinds of standards. According to Table 3, this would cause a data loss of about 9 % in channel 1 and 7 % in channel 36. At an amount of more than 19,000 detected events, this seems to be bearable. One must note that the percentage of interferers changes up to 20 % if all records of 2.4 GHz band are investigated. This would cause an unacceptable data loss. Furthermore, Linux offers information about demodulated frames in terms of frame length, modulation, power level, and further details about MAC properties. As one important key parameter describing channel characteristics is the power level, inaccuracies have to be taken into account. There is no transparency about how power levels are measured and for which time period. In addition to this, capturing data by WLAN cards significantly suffers from crosstalk problems due to inadequate isolation. If an ideal receiver is examined, one would expect detecting data only when connected to a proper antenna. Unfortunately,

WLAN cards still receive signals when the respective inputs are left open or terminated by a load. As already mentioned, the Linux system was placed outside of the lecture hall. Consequently, the PC-demodulated traffic does not fit to the one from the VSA. In summary, it can be said that a real-time device for recording ISM band traffic cannot be replaced by a Linux PC system in this configuration.

References

[1] S. Atapattu, C. Tellambura, H. Jiang. *Energy Detection for Spectrum Sensing in Cognitive Radio*. Springer, 2014 6

[2] Andrea Mariani, Andrea Giorgetti, Marco Chiani. *Effects of Noise Power Estimation on Energy Detection for Cognitive Radio applications*. IEEE Trans. Commun., Vol 59, No. 12, December 2011. 6, 7, 8

[3] Tektronix. *Wi-Fi: Overview of the 802.11 Physical Layer and Transmitter Measurements*, 2013. `https://www.cnrood.com/en/media/solutions/Wi-Fi_Overview_of_the_802.11_Physical_Layer.pdf` 9

[4] Michael Galarnyk. *Understanding Boxplots*, 2018. `https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51` 11

[5] Adrian W. Bowman and Adelchi Azzalini. *Applied Smoothing Techniques for Data Analysis*. Clarendon press, Oxford, 1997. 13

[6] David Freedman, Persi Diaconis. *On the Histogram as a Density Estimator: L_2 Theory*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, Springer, 1981. 14