Quantifying the Spectrum Occupancy in an Outdoor 5 GHz WiFi Network with Directional Antennas

Michael Rademacher  
Hochschule Bonn Rhein-Sieg  
Sankt Augustin, Germany  
michael.rademacher@h-brs.de

Karl Jonas  
Hochschule Bonn Rhein-Sieg  
Sankt Augustin, Germany  
karl.jonas@h-brs.de

Mathias Kretschmer  
Fraunhofer FIT  
Sankt Augustin, Germany  
mathias.kretschmer@fit.fraunhofer.de

Abstract—WiFi-based Long Distance networks are seen as a promising alternative for bringing broadband connectivity to rural areas. A key factor for the profitability of these networks is using license-free bands. This work quantifies the current spectrum occupancy in our testbed, which covers rural and urban areas alike. The data mining is conducted on the same WiFi card and in parallel with an operational network. The presented evaluations reveal tendencies for various aspects: occupancy compared to population density, occupancy fluctuations, (joint)-vacant channels, the mean channel vacant duration, different approaches to model/forecast occupancy, and correlations among related interfaces.

Index Terms—Spectrum occupancy; spectrum sensing; U-NII band; IEEE 802.11; cognitive radio; LAA; LTE-U,

I. INTRODUCTION AND MOTIVATION

According to a recent report by the International Telecommunication Union (ITU) only one out of two people on the globe and only one out of seven people in the Least Developed Countries (LDC) are using the Internet [1]. To decrease the costs of broadband connectivity and thus make the Internet more widely available, so-called “Alternative Networks” have been evaluated [2], [3]. As one of these alternatives, WiFi-based Long Distance (WiLD) networks use Commercial Off-the-Shelf (COTS) high-gain directional antennas and WiFi cards to span multiple wireless links over distances up to 25 km. One crucial cost factor for “Alternative Networks” is the availability of license-free bands; however, according to [2], these bands face challenges due to overcrowding. The overall occupancy especially in the Unlicensed National Information Infrastructure (U-NII) band is likely to increase due to recent developments to utilize this band for LTE operations [4] or when modern WiFi standards such as IEEE802.11ac with their increased bandwidth become more ubiquitous. The motivation behind this work is to analyze the current U-NII band occupancy in our WiLD testbed. In particular, we are interested in using the findings to design an enhanced centralized channel allocation algorithm for our architecture. Like nearly all measurement campaigns, our analysis is limited both geographically and in time; however, we found viable tendencies.

The main contributions of this work are the following: We present a methodology where the data mining is conducted on the same WiFi interfaces in parallel with our operational WiLD network at 11 different locations. Using this data, we analyzed the spectrum occupancy in rural and urban areas, spectrum occupancy fluctuations, the number of (joint)-vacant channels, the mean channel vacant-time, different approaches to model/forecast occupancy and correlations among related interfaces.

II. RELATED WORK

A comprehensive and recent literature survey of spectrum occupancy measurement aspects is presented in [5]. The authors compare metrics, measurement targets, devices and previous measurement campaigns. They conclude that focused and carefully designed campaigns are of greater interest than more generalized ones and that the results must be carefully scrutinized since providing a simple spectrum occupancy value can lead to false conclusions.

To the best of our knowledge, limited previous work dealing with spectrum occupancy in the U-NII band exists. Only wide-band studies using high performance spectrum analyzers attached to omni-directional antennas have been conducted. In [6], an outdoor measurement on a rooftop in the city of Aachen (Germany) in 2007 is evaluated. The authors conclude that the U-NII band is vacant mainly due to popularity differences of 2.4 GHz and 5 GHz WiFi home networks. The authors in [7] report results from a spectrum occupancy measurement from two locations in Chicago (USA) and from a university campus in Finland. A mobile measurement deployment is used, and they discover that the occupancy in the U-NII band is well below 5%. However, one long-term measurement located on an exposed tower is also presented. For this setup, the channel occupancy significantly increases up to 50%. According to the authors, the main reason is that, in the US U-NII band, transmissions are mainly directed backhaul-links at exposed locations. Most related to our methodology, [8], [9] measured the spectrum occupancy in the 2.4 GHz band using 16 WiFi cards attached to omni-directional antennas in an urban parking lot. The authors propose a 3-state hypothesis model for dynamic channel selection.

III. METHODOLOGY

Our testbed is located in the Rhein-Sieg area of Germany around the Fraunhofer Campus in Sankt Augustin and covers rural areas as well as urban areas such as the Bonn Rhein-Sieg University of Applied Sciences. The testbed consists of
multiple point-to-point-links with distances ranging between a few hundred meters up to 11 kilometers. We focus on a practical build-up instead of relying solely on using radiotowers by mainly exploiting rooftops. Our WiLD nodes consist of a Single-Board Computer (SBC) with multiple (mini-)PCI-e slots holding the WiFi cards. Overall, we use three types of WiFi cards all based on the same chip and therefore on the same version of the Linux open-source driver (ath9k). The SBCs are placed in outdoor suitable enclosures mounted at the desired places. We use two types of cross-polarized antennas with 19 and 25 dBi gain. Figure 1b shows an example. The testbed is orchestrated by our WiLD network management software, WiBACK [10]. WiBACK manages the connectivity to each node on an operational channel using the same WiFi card on which we conducted the measurement process. The occupancy measurements are conducted on selected interfaces, which are embedded in multi-interface nodes. The location and name of all nodes is shown in Figure 1a. Overall, this measurement campaign lasted for one week from 05-Jul-2017 2:30 pm to 12-Jul 2:30 pm. The goal of our sensing approach was to evaluate how our network “sees” the current occupancy of the U-NII band.

The scanning process can be subdivided into different phases each associated with a specific duration [5]. We periodically allocated a specific amount of time for scanning other channels during which the respective interface was to be unavailable for packet-forwarding on the operational channel. The monitoring time \( T_m \) is the time spent by each WiFi card quantifying the occupancy on a channel. This duration is predetermined by the hardware as \( T_m = 69 \text{ ms} \). After each scan, the WiFi card returns to the operational channel to forward packets in the network for the duration of \( T_o \). We choose \( T_o = 1 \text{ s} \), which is significantly longer than \( T_m \), to avoid performance instabilities. Afterwards, the WiFi card tunes to the next channel, quantifying the channel occupancy. Overall, we scanned \( N_c = 19 \) different channels: 8 channels in the U-NII-2 band (5.180 GHz–5.320 GHz) and 11 in the U-NII-2e band (5.500 GHz–5.700 GHz) at a bandwidth of 20 MHz each. \( T_d \) describes a dead time to store the current occupancy in a database located at the centralized controller and transmitted over the operational channel. If a node had multiple WiFi cards \( (N_i) \), we sequentially conducted the same process with the next card. The duration of one scanning pass on one node is called the revisit/sweep time \( (T_s) \):

\[
T_s = N_i \times [N_c \times (T_m + T_o) + T_d]
\]  

(1)

We obtained for the complete campaign \( T_s = 50.3 \pm 7.8 \text{s} \). The variance is caused by two different factors. First, nearly all of our nodes have 2 WiFi cards while one node is also equipped with a third one. Second, some scanning inquiries were canceled by the Linux kernel driver.

Similar to [9], we calculate occupancy based on the duration the WiFi card senses the medium as busy. We found that the cards report two different values to indicate this duration. The first value describes the duration the radio spent receiving WiFi frames during \( T_m \), called \( T_u \). The second value is the duration the channel was sensed busy due to a signal above the current Clear Channel Assessment (CCA)-threshold, called \( T_uCCA \) [11]. We found that \( T_u \) and \( T_uCCA \) are almost equal in our testbed with one notable difference where \( T_uCCA \) was found significantly higher. This confirms our assumption that the U-NII band is currently mainly populated by WiFi transmission. We use \( T_u \) throughout this work since we obtained on one version of our WiFi card a low but constant signal for \( T_uCCA \), which needs further investigation. We define spectrum occupancy at a certain interface \( i \) on the \( c \)th channel measured at a time \( t \) as follows:

\[
o_{i,c,t} = \frac{T_u}{T_m}
\]  

(2)

Since \( T_s \) is significantly larger than the maximum expected packet transmission length (4 ms for IEEE802.11n), we averaged (down-sampled) multiple scan results \( (o_{i,c,t}) \) in a certain time period \( T_P \) to a single value \( O_{i,c,T_P} \). Unless stated otherwise, we use a period of \( T_P = 15 \text{ min} \), which reflects the mean of \( \approx 18 \) individual scans based on \( T_s \). The developed software components, a daemon written in C for the scanning
A third interface located on one of the nodes executes the artificial unidirectional UDP traffic until the link was saturated. Unoccupied channel and generating an increasing amount of simple indoor WiFi link using two of our nodes on a nearly with 15 min = 672/0661–26cc. The x-axis shows the overall time: Figure 2 shows a heatmap based on $O_{i,c,T_p}$, subplots) and all channels (y-axis of each subplot). Each interface is labeled using the associated nodes names in the form location-direction (cf. Figure 1a). For example, the first subplot (top left) shows $O_{i,c,T_p}$ at node 0661 using the interface attached to the antenna in direction to 26cc (0661–26cc $\neq$ 26cc–0661). The x-axis shows the overall time: one week/15 min = 672. We limit the heatmap to between 0 and 0.5 for a better visualization of the results since an occupancy of $O_{i,c,T_p} > 0.5$ was barely encountered. We excluded the operational channel in the plot by using a dedicated color (white). Most of the interfaces stayed on the same channel during the campaign; however, on certain inter-

IV. RESULTS AND ANALYSIS

Figure 2 shows a heatmap based on $O_{i,c,T_p}$ for all interfaces ($N_I = 14$, subplots) and all channels (y-axis of each subplot). Each interface is labeled using the associated nodes names in the form location-direction (cf. Figure 1a). For example, the first subplot (top left) shows $O_{i,c,T_p}$ at node 0661 using the interface attached to the antenna in direction to 26cc (0661–26cc $\neq$ 26cc–0661). The x-axis shows the overall time: one week/15 min = 672. We limit the heatmap to between 0 and 0.5 for a better visualization of the results since an occupancy of $O_{i,c,T_p} > 0.5$ was barely encountered. We excluded the operational channel in the plot by using a dedicated color (white). Most of the interfaces stayed on the same channel during the campaign; however, on certain inter-
A. Location and Free Channels

In particular, we are interested in differences between urban and rural areas. Therefore, Figure 4a shows a box plot of the mean spectrum occupancy for all channels on each interface:

\[
\tilde{O}_{i,c,T} = \frac{1}{N_C} \sum_{c=0}^{c=N_c} O_{i,c,T_P}
\]

We categorize specific interfaces if both associated nodes are clearly\(^\text{1}\) either in a rural (RR) or an urban location (UU). The two categories show a viable difference in spectrum occupancy: At our most rural interfaces (nodes 0661 and 26cc), the spectrum is vacant while all urban interfaces show a mean spectrum occupancy between 0.02 and 0.05. The numerous outliers (points) and large whiskers indicate again that the occupancy is fluctuating over time.

For channel allocation algorithms, the number of available vacant channels is a significant parameter. In this work, we use a binary decision model to determine the state of a channel:

\[
s_{i,c,T_P} = \begin{cases} 
1, & \text{if } O_{i,c,T_P} < \lambda \quad \text{(vacant)} \\
0, & \text{otherwise} \quad \text{(occupied)}
\end{cases}
\]

\(^\text{1}\)Node 6138 is challenging to categorize: While the population density is a rural place, the node is located on a tall radio tower. Node 1338 is located in a rural place but the antenna points with direct line-of-sight to a city.
we calculated the mean vacant duration for every channel at each interface. Given a sequence of occupancy states for each channel and interface \((s_{i,c,T_P})_{T_P=0}^{T_P=672}\), the mean size of the sets of concurrent 1s equals the mean vacant duration (e.g. \((010101) = 1; (110010) = 1.5; (111000) = 3\)). Figure 6 shows an Empirical Cumulative Distribution Function (ECDF) for all interfaces as well as for one rural (0661–fed4) and one urban (e3bc–a4f0) interface. The y-axis shows the fraction of channels that have mean vacancy duration less than or equal to the corresponding value on the x-axis. At the rural interface (0661–fed4), only 23% of the channels switch between occupied and vacant and 77% of the channels were vacant for one week (Point A). At the urban interface (e3bc–a4f0), channels are switching significantly more often and no channel has been found vacant for one week (Point B). The ECDF for all channels reveals that 22% are vacant for one week. However, up to this point, the function is steadily increasing - an interesting challenge for channel allocation in WiLD networks. If during the initialization of the network a channel is measured vacant, it may become occupied again shortly after. For example, 20% of the channels have a mean vacant duration of less than 3 periods and 50% of the channels have one less than 12 periods (Point C). However, for final conclusions regarding the possible time-scale, a faster revisit time \(T_s\) would be beneficial for this analysis, an aspect we regard as interesting but challenging for future work.

C. Time of the Day

Using our mined data, we were also interested in longer time-scale effects of spectrum occupancy: Is there a daily seasonality in occupancy depending on different times of the day? Therefore, we calculated the mean spectrum occupancy at every period for all channels at all interfaces:

\[
\bar{O}_{i,c,T_P} = \frac{1}{N_i N_c} \sum_{i=1}^{N_i} \sum_{c=1}^{N_c} O_{i,c,T_P}
\]

for two periods of \(T_P = 15\) min and \(T_P = 120\) min. Figure 5 reveals that there is a daily seasonality in our testbed. There are seven minimal turning points, which all occur in the early morning. In addition, the fluctuation of spectrum occupancy on the weekend is less than on weekdays.

D. Modeling

Inspired by the work in [9], [14], we tested different approaches to model or forecast occupancy. To initially reduce the amount of data, we considered interfaces with meaningful occupancy \((\bar{O}_{i,T_P} > 0.05)\) before considering the more challenging case of nearly vacant ones.

We tested the following hypothesis using the Kolmogorov-Smirnov (K-S) test: Is the occupancy at an interface on a certain channel equal to a common probability distribution? Overall, we tested 77 different distributions such as normal, chi and chi-squared, t and Weibull. However, we were unable to receive any statistically significant indication that spectrum occupancy is drawn from one of the evaluated distributions.

Afterwards, we used an autoregressive model to evaluate the possibility to model and forecast the spectrum occupancy on a channel. We followed the main steps of the Box-Jenkins approach. By evaluating the partial autocorrelation (PACF) plots, we found that AR(4) fits well for the majority of channels and that adding more factors does not significantly improve the accuracy (over-fitting). Similar to the results in [14], the AR(4) model seems to accurately follow the measured data, even reacting to spikes. In addition, we achieved pleasing values for the uncentered version of \(R^2\). However, when applying a more suitable metric for time-series, such as the Mean Absolute Scaled Error (MASE) [15], we found that the AR(4) model is not significantly superior to one-step errors from a naïve approach: taking the last measured value as a forecast \(O_{i,c,T_P} = \bar{O}_{i,c,T_P-1}\).

E. Correlation Among Links and Positions

Correlations for the spectrum occupancy among interfaces could reduce the needed sensing effort for channel allocation...
For the majority of interface combinations, we did not expect a correlation, and this expectation was confirmed by the results. However, by adding certain letters to Figure 7, we want to emphasize categories where correlations are more likely: “L” if the interfaces are part of the same link, “N” if both interfaces are on same node and “G” if the corresponding nodes are in proximity to each other. Indeed, the majority of cases with a higher mean correlation fall into these categories. For example, there is a strong correlation between e3bc–e344 and e3bc–a4f0. This situation occurs since multiple external WiFi Access Points (APs) are in proximity to the node so that the spectrum occupancy increases on both directional antennas (possibly due to side-lobes). However, this higher correlation is not persistent among all channels. Overall, only a weak mean correlation exists for two interfaces on the same link, which means that a mutual decision is necessary for channel allocation algorithms on WiLD links.

V. CONCLUSION AND FUTURE WORK

Our initial assumption was that spectrum occupancy in the U-NII band would be significantly higher than what we presented in this work. Especially in rural areas, we found that our WiFi cards sense numerous vacant channels and the mean spectrum occupancy in urban areas was measured as below 5%, which is less occupied than expected. If an intelligent channel assignment algorithm is assumed, there seems to be sufficient spectrum even for 40 MHz channels. Our methodology shows that it is a practicable approach to use the same WiFi interface for sensing and packet-forwarding. However, a more sophisticated approach should exploit specific idle times of the interfaces (e.g. inter-frame spaces) instead of dedicating a fixed amount of time continuously. In addition, to draw more reliable conclusions, we aim to investigate bounds for $T_s$. To declare a channel vacant, we chose a threshold of 0.01 occupancy based on our experience with WiLD networks. Increasing this value would have declared more channels as vacant but at a potential QoS loss: a trade-off that can be further investigated. If a channel becomes vacant, predicting the duration seems challenging. Based on previous work, we evaluated different approaches to model spectrum occupancy. Neither is our data drawn from one of numerous probability distributions nor does an autoregressive model provide a significant benefit, judged by a recent forecast metric. For certain channels there is a weak correlation among distributed interfaces in our network. However, this correlation is too weak to exploit. Although the results of this work reflect our testbed in Germany, we found a great potential for “Alternative Networks” operating in a license free band, especially for our desired scenario, which implies directional antennas and rural areas.

REFERENCES