Measurement 199 (2022) 111573

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

On the use of sniffers for spectrum occupancy measurements of Bluetooth low energy primary channels

A. Valenzuela-Pérez [a], M. García-Lozano [a,*], J.L. Valenzuela [a], D. Pérez-Díaz-de-Cerio [a], Á. Hernández-Solana [b], A. Valdovinos [b]

[a] Universitat Politècnica de Catalunya (UPC)., Campus del Baix Llobregat, C/ Esteve Terradas, 7, 08860 Castelldefels (Barcelona), Spain
[b] Aragon Institute for Engineering Research (I3A), University of Zaragoza (UZ), C/ María de Luna, 1, 50018 Zaragoza, Spain

ARTICLE INFO

Keywords:
Spectrum occupancy measurement
Channel occupancy
Sniffer measurement
Bluetooth
BLE

ABSTRACT

The methods usually employed to measure channel occupancy show limitations in the context of Bluetooth Low Energy (BLE) advertisements. We propose and analyze the use of BLE sniffers as light and portable low-cost spectrum occupancy meters to be used in scenarios where real time signal analyzers are not adequate. For the measurement technique to be successful, several low-level effects must be considered. The paper argues about on-air time, receiving blind times due to processing and intra system interference, buffer saturation and frequency anchoring. Hence, a compensation procedure based on collision rate estimation is proposed. Results with the refined method show that occupancies of 40% can be measured with an overestimation error whose percentile 95% is 5 percentage points. This is reduced to 1.9 points when the occupancy is 15%. The sniffers perform in real time and are shown to correctly track short term load variations. The strategy has been successfully used to characterize occupancy in highly variable and loaded scenarios such as subway platforms and a shopping mall. Values up to 25% have been observed, which implies a relevant packet error rate. Hence, the tool can be used to make agile audits and configure the parameters that control communication redundancy in new or existing networks.

1. Introduction

The 2.4 GHz industrial, scientific, and medical (ISM) band is massively used to provide wireless communications under a high variety of standards and proprietary systems. Among them, Bluetooth, and Bluetooth Low Energy (BLE) in particular, is an omnipresent technology and a dominant actor in consumer electronics. Annual shipments of BLE devices have grown from 3.6 to 4.5 billion in the past five years. And despite a one-year shift in forecasts due to the pandemic, annual shipments of Bluetooth enabled devices is expected to reach 6.4 billion in 2025 [1].

BLE is characterized by the use of advertising messages that broadcast the presence of devices. Such *primary advertisements* are transmitted over three carrier frequencies labeled 37, 38 and 39, one after the other. They can be used to transmit data (advertising mode) or to start a connection between devices (connected mode). When a connection is established, the carrier frequency is hopped among the remaining 37 channels, completely covering the 2.4 GHz band. Such 37 frequencies can also be used to extend the primary advertising data capacity. Indeed, BLE advertisements and their broadcast messaging capability have become enormously popular.

Bluetooth is present in a vast variety of contexts, vehicular industry, consumer electronics, lightning systems, white goods… Most of these devices continuously broadcast advertising messages. Primary advertisements are also widespread to create *beacons* [2] that transmit short packets for proximity marketing, information points or location services. The latter include guidance and indoor positioning [3], asset tracking [4], antitheft systems [5] and so on. Remarkably, Bluetooth beaconing is one of the proposed technologies for drone remote identification which will be mandatory in 2023 in some countries [6]. Bluetooth 5.1 specifically promotes the development of high precision positioning applications by allowing to estimate the signal angles of departure and arrival [7]. Primary advertisements are also the basis of Bluetooth mesh networks. They allow direct communication without the need of handshake

* Corresponding author.
E-mail addresses: ana.valenzuela@estudiantat.upc.edu (A. Valenzuela-Pérez), mariogarcia@tsc.upc.edu (M. García-Lozano), jose.luis.valenzuela@upc.edu (J.L. Valenzuela), david.perez.diaz.de@upc.edu (D. Pérez-Díaz-de-Cerio), anhersol@unizar.es (Á. Hernández-Solana), toni@unizar.es (A. Valdovinos).

https://doi.org/10.1016/j.measurement.2022.111573
Received 16 February 2022; Received in revised form 7 June 2022; Accepted 27 June 2022
Available online 30 June 2022

or association to share messages. This is a notable feature for Industry 4.0 and spots Bluetooth mesh as a very interesting capillarity network [8].

Primary channels enjoy some protection against interference from other systems. This is due to their spectral location, at the edges of the band and out of the main channels of prevalent standards such as those based on IEEE 802.11 and 802.15.4, as shown in Fig. 1. This permits to minimize the overlap and thus, to maximize the signal to interference ratio (SIR). Among the three channels, number 37 (2402 MHz) is almost free of interference. On the other hand, number 38 (2426 MHz) is the most exposed since it might collide with non-primary IEEE 802.11 channels and presents adjacent interference with IEEE 802.15.4 g in its channel 15 (centered in 2425 MHz). Channel 39 (2480 MHz) collides with channel 26 of 802.15.4. However, channel 26 is not allowed in several countries and the presence of systems based on this standard is still much lower than that of 802.11 networks. Also, some 802.15.4-based systems do not use it (e.g. WirelessHART) or define it as optional (e.g. ISA 100.11a). For this reason, interference on channel 39 will generally be low. Nevertheless, if an advertising frequency is blocked, the other channels might well be free, given that they are several MHz apart.

On the other hand, intra-system interference can be problematic in BLE crowded environments and considering the forecasted technology penetration this might become an issue for advertising applications. As BLE devices share the access to spectrum, collisions are inevitable. Such reliability reduction has effects over the connection delay, the discovery process and over the performance of mesh networks. In this case, it is highly advisable to perform an audit to measure the occupancy of the primary BLE channels. From here, radio engineers could make a correct configuration of the parameters that control the communication redundancy and/or the mesh layout itself [8].

Throughout this paper, we examine the methods usually employed to measure channel occupancy and argue about their limitations in the context of BLE advertisements. We propose and analyze the use of BLE sniffers as a light and portable means to measure occupancy with no limits in measurement time while being able to capture very short-term variations. We examine low level effects such as processing and blind times, induced by intra system interference, and how they even affect the occupancy characterization itself. Errors are quantified to establish the limits of this approach. Finally, several measurement campaigns have been carried out and are analyzed as channel occupancy examples to investigate the potential saturation of primary BLE channels.

This objective is developed along the rest of the paper. Related works are commented through the next section. Next, section 3 describes and quantifies the required improvements to usual methods and presents sniffers as a potential solution. Drawbacks and solutions are also investigated. Obtained results are presented in section 4. The last section closes the work with our main conclusions.

2. Related work

Spectrum surveys on the 2.4 GHz band have been well-reported in the scientific literature. The usual methodology relies on sweeping and Fast Fourier Transform (FFT) based spectrum analyzers or software defined radios acting alike. However, most works relying on this methodology focus on IEEE 802.11 Wireless Local Area Networks (WLANs) channels. Relevant measurement campaigns are presented in [9]. The authors characterize WLAN channels intensively by performing two one-week measurement campaigns, taking up to 1250 spectral sweeps per minute (a total time of 48 ms per sweep). Another exhaustive investigation is done in [10], where eight different locations were covered at ten different measurement moments. The occupancy of the WLAN channel is analyzed, in this case, with a sampling period of 10 ms. The work presented in [11] also investigates several environments in the city of Melbourne using a sweep time of 4 ms. The authors of [12] analyze spectrum occupancy at 2.4 GHz for cognitive radio applications by taking measurements during 24 h with a time resolution of 1 s. The researchers in [13] highlight the importance of improving the temporal resolution to make a correct characterization of channel occupancy. Specifically, they perform a study on WLAN channels by capturing the entire band every 205 μs in an unspecified indoor environment. Specific BLE spectrum occupancy is hardly addressed. Most works deal with connected mode [15,16] with a special focus on inter system interference and packet error rate (PER) [17]. When performing spectral measurements, time resolutions are found to be similar to the previous works, for example, [18] utilized a time sweep of 4 ms.

Sweep and processing times determine the time resolution to revisit a certain frequency to be evaluated. Such sampling period has a direct impact on the quality of the occupancy estimation. If all packets are to be captured, the time resolution should be set considering that BLE primary advertising messages can be as short as 128 μs. On the other hand, the sampling period is directly proportional to the bandwidth to scan. Also, it is inversely proportional to the resolution bandwidth in sweeping analyzers or the FFT subcarrier spacing in digital ones. Hence, it must be noted that BLE channels occupy 1 MHz, well below other systems operating at 2.4 GHz.

Moreover, the exact detection of the instants when the channel switches from occupied to idle (and back) is not possible. This leads to a systematic error [19]. Such error is aggravated when the channel is occupied by many packets of short duration, which indeed is the case of

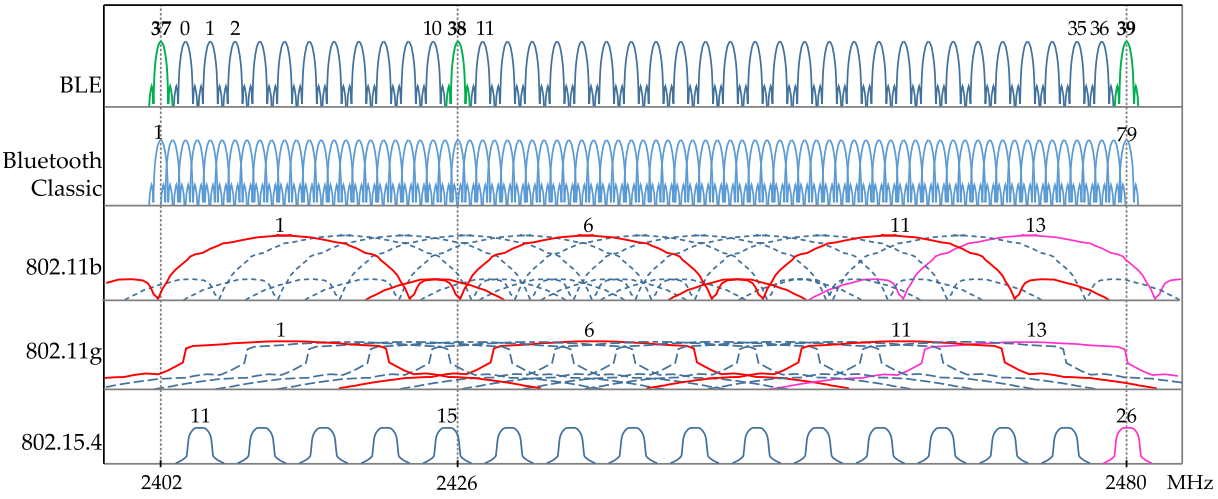


Fig. 1. BLE channel distribution with respect to the main channels of IEEE 802.11 and 802.15.4.

BLE. For all these reasons, an accurate measurement and characterization of BLE occupancy imposes more stringent requisites and requires a careful study.

3. Materials and methods

Many research works that study channel occupancy rely on spectral analysis and employ time resolutions that are too large in relation to the size of BLE advertising packets. In some cases, such times are several orders of magnitude higher than the BLE packets. This implies that many of the BLE packets would not be captured. Nevertheless, it is not strictly necessary to capture each and every transmission to get the mean occupancy. Since it is a statistical measure, accuracy can be improved by accumulating a sufficient number of samples.

Let us take some of the previous time resolutions as an example, 4 ms, 50 ms and 1 s. If the actual occupancy is 10% and the estimate must have a relative error of less than 5% with a confidence of 99.73% (three sigmas), the minimum number of measures required would be 90,000 (refer to Appendix A). This implies measurement times of 6 min, 75 min and 25 h respectively. These values are even higher if we consider that samples are not independent. For example, in certain types of advertisements that can trigger responses on other devices, or in a BLE mesh network with high redundancy, where sending a packet involves new subsequent transmissions. These times are excessive for a correct characterization of the occupancy of BLE primary channels. This is because much of the traffic is generated by mobile devices, and levels are not kept constant in such time windows in most environments. Take, for instance, concentrations of people on a subway platform or in a shopping mall. Time cycles of occupancy are in the order of seconds. For example, if load remains constant during 5 s, the required sampling time would be of just 56 μs under the premises indicated. Hence, it is concluded that time resolution must be reduced far beyond the typically used values.

Our experiments integrate the FSW26 real-time signal analyzer from Rohde & Schwarz. This device allows taking I/Q samples with a sampling time of 0.8 μs and can store up to 10^8 samples (complex numbers). Hence, it must be noted that the number of samples to store rises as a new limiting factor. With the previous sampling time, this means measurements of up to 80 s. After this, the data must be cleared and loaded to another storage medium to continue with the measurement process. So, there is a compromise between accuracy and required measurement interruptions, which might well be not acceptable for a correct characterization of occupancy variations along time. Additional drawbacks of real time signal analyzers are their large volume, weight, transport difficulties and the need for an energy source. This makes them an impractical tool for many out-of-lab situations.

3.1. On the use of sniffers for occupancy measurements

The proposed approach relies on the use of sniffers, devices that are often used as protocol analyzers. Sniffers detect, capture, and demodulate packets of a certain technology. They are typically supplied with a small form factor, as a USB dongle. They obtain relevant information such as the arrival time of the packet, its duration, the frequency used, its signal level, etc. Having the exact moments of arrival and the duration of the advertisements opens the door to use the sniffer as a tool to measure the channel occupancy in near-real time and, recording for hours, without interruptions. Our investigation is done with the nRF52840 and the nRF52832 Bluetooth SoCs from Nordic Semiconductor and with the CC2652 from Texas Instruments, which show a resolution equal to 1 μs, 1 BLE symbol time. Also, it is remarkable their low cost, low power consumption, excellent portability, and ease of use:

- Cost: The price of a development kit that includes a Bluetooth chip with sniffer capacity is less than 50 Euros, to which we must add the price of a laptop. On the other hand, a signal analyzer capable of making lossless captures in real time, and considering the main vendors in the market, is around 80,000 Euros and depending on the options included.
- Power consumption: A sniffer and a laptop storing the data would not exceed 60 W. While an analyzer can reach peaks of around 1 kW when operating in real time mode. This implies that the analyzer always requires an available energy point.
- Portability: The weight of the sniffer-based solution will be determined by the laptop. Therefore, it is generally less than 2 kg. However, an analyzer like the one mentioned usually weighs around 20 kg. Its dimensions are also much larger than a laptop with a USB dongle, they generally exceed 450 mm (of width and depth) and 200 mm of height.
- Ease of use: The sniffer-based solution is relatively plug-and-play. It just requires the use of Wireshark, the most used protocol analyzer. It is needed to add a plug-in to configure the specific channel to monitor. The use of a signal analyzer is not trivial, the number of parameters to adjust is much higher. Nevertheless, it is usually a knowledge that most electrical engineers acquire during their training.

On the other hand, there are some specific problems that must be addressed for this approach to be successful. They are a matter of study along the rest of this section.

- **Exact occupancy of a given packet**

Sniffers can read the structure of the packet and obtain the number of carried payload bits. Thus, they are able to indicate the duration of a packet. Of course, this matches the Bluetooth specifications, which set a preamble, access address, header, payload and a checksum. However, it is important to note that this duration does not match the occupancy time. Transceivers in commercial devices occupy the channel for an additional time. The rationale behind this is that they must first ramp-up in transmission mode (and ramp-down into idle state). So, they generate a *pre advertisement sequence* and after that time, the device is ready to initiate a packet transmission. Transmission is ended with a *post advertisement sequence*. They transmit a carrier, which is the result of modulating several '1s' just before and after the packet structure (Fig. 2). The length of this extra time varies depending on the manufacturer and version of the hardware. We have characterized it in the lab for over ten different Bluetooth vendors and it ranges from 4 to 60 μs.

- **Dealing with post-processing and overlapped packets**

The sniffer cannot properly handle overlapped or too close packets. This leads to a number of missed packets that is proportional to the number of devices resulting in occupancy underestimation. The different causes/situations are explained in detail next.

Firstly, the sniffer requires a processing time right after decoding an advertisement [20]. During that time, channel listening is interrupted. This means a potential loss of packets, not only of those that are overlapping, but also those that are too close in time. It is important to remark that such blind times happen every time a new packet is detected and received. This implies that the underestimation error is to be directly proportional to the channel occupancy.

This is illustrated in Fig. 3, where the blind time for the nRF52840 is characterized. This is done by transmitting two consecutive advertisements with a variable controlled separation between them. They are generated in Matlab and their I/Q components are fed into a vector signal generator (Rohde & Schwarz SMW200), pre and post packet carrier sequences of 6.5 μs are included. The spacing is reduced until the second advertisement is no longer detected by the sniffer. Then, the exact interframe space is measured with a real-time signal analyzer (Rhode & Schwarz FSW26). A picture of this set-up is given in Fig. 3a. For this device, it is necessary a space of at least 50 μs between both

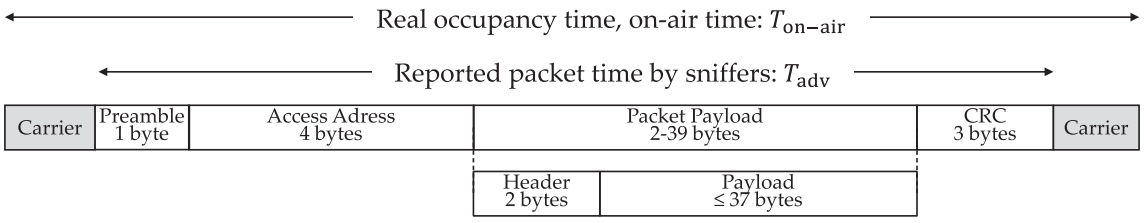


Fig. 2. BLE packet structure defined by the standard and initial and end carriers that increase occupancy.

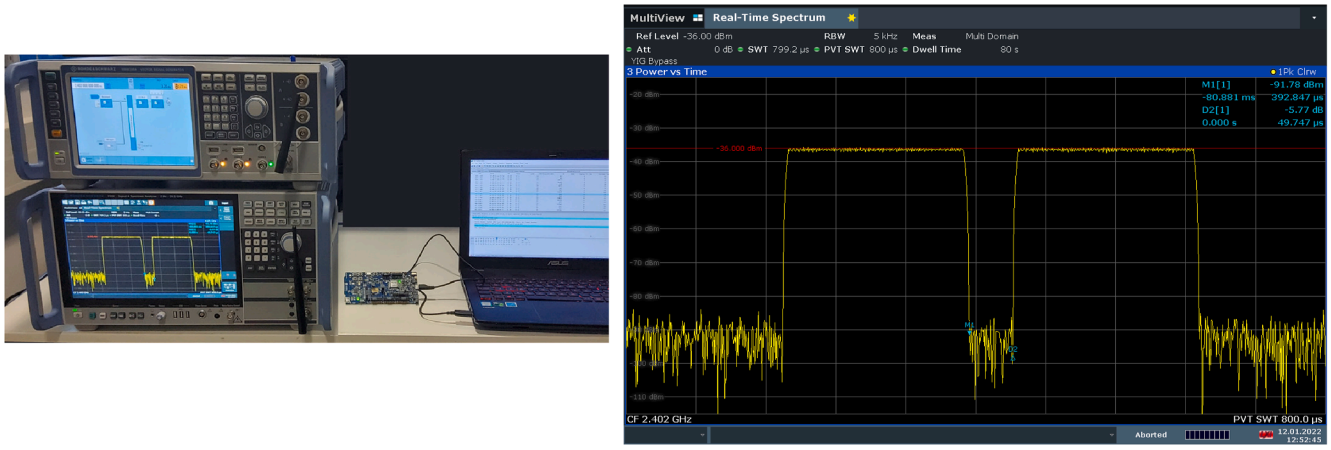


Fig. 3. Blind-time due to processing time in sniffer. **(a)** Measuring set-up. **(b)** Received power versus time, it can be observed the minimum space between consecutive advertisements (50 µs).

advertisements (Fig. 3b). This means that the gap between actual data transmissions is 63 µs (6.5 µs + 50 µs + 6.5 µs).

The second reason for underestimation is packet loss due to specific overlap situations when collisions occur. Collisions may lead to insufficient SIR, which results into corrupt packets that are impossible to decode. In this case, the sniffer would not report them, and a falsely null occupancy would be assumed. On the other hand, collision/overlap situations with sufficient SIR may yield different outcomes. It is necessary to understand that the sniffer initiates the receipt of an advertisement by retrieving its preamble and access address. These are the first five bytes of the packet (see Fig. 2) and are used to recover timing and frequency synchronization. With this in mind, it is possible to identify several overlapping cases with sufficient SIR. Fig. 4 shows three different situations captured in lab and leading to different probabilities of packet error and hence, occupancy underestimation as depicted next. The figures represent the received power versus time and the colored boxes represent the packets themselves and help to read the plots.

1. Fig. 4(a): The first packet shows a high received power and it is correctly synchronized, though potentially malformed due to the overlapping. Its probability of not detection matches its packet error rate, $PER \approx 1-(1-BER)^b$, where BER is the bit error rate and b is the number of overlapped bits. Note that the BER is a function of the experienced SIR. The second packet is lost with probability 1 because the sniffer is busy with the first one as already explained. This contributes to the occupancy underestimation.
2. Fig. 4(b): The sniffer starts synchronizing with the first packet, then a second advertisement with high relative power overlaps and greatly degrades the SIR. All overlapped bits show a $BER = 0.5$, so the packets in that situation are lost with probability $PER \approx 1-0.5^b$. The second packet will nearly always be missed, but there is a possibility for detection. This happens when it overlaps with the sync sequence of the first packet. In this case, the sniffer will get synchronization with the second advertisement, which can be recorded thanks to its favorable SIR.

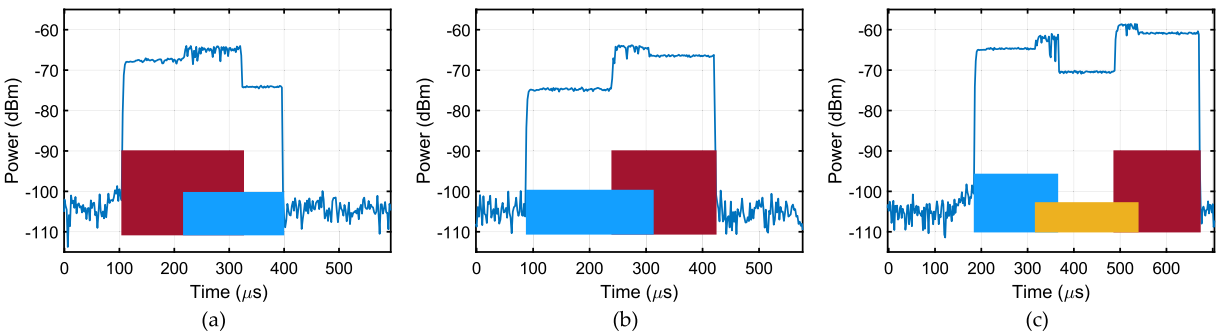


Fig. 4. Overlapping possibilities for advertising packets of different power and sufficient SIR.

3. Fig. 4(c): The last case shows overlapping of three different packets. The sniffer can demodulate the first one, which has a power level that overcomes interference from the second. When it has finished, the channel is still occupied by the second packet, which is lost because the receiver misses its synchronization sequence. The third packet is received because the sniffer detects sync bytes with good SIR. Advertisements like the first and last show a $PER \approx 1-(1-BER)^b$, while the second has a $PER = 1$.

- **Buffer saturation and USB capacity**

Nowadays, there are different Bluetooth devices that can be used as sniffers but not all of them would be suitable to perform occupancy measurements. It has been observed that the basic hobbyist models saturate their buffers quickly. That is, packets are lost due to the device's inability to store and transmit a high volume of packets through the USB port. Hence, they would not be adequate to measure on highly loaded scenarios. Information about this capability is not present in usual specifications. The modules considered in this investigation (nRF52840, nRF52832 and CC2652) were able to report occupancies of 40% successfully and after applying countermeasures to mitigate underestimation errors, as described later.

- **Anchoring to a specific radio channel**

Since BLE advertisements are repeated on all three primary frequencies, Bluetooth occupancy will be the same on all of them. Indeed, sniffers must remain anchored to one of the three primary channels for the entire measurement. Otherwise, if the reception frequency was changed periodically, packet losses would occur. This is because commercial devices take a time to switch frequencies, which causes a blind time in which no frequency is heard. This gap has been determined to range from several microseconds up to several milliseconds [20].

3.2. Underestimation errors: Quantification and compensation

To have a first evaluation of sniffers performance and provide a means for compensation of underestimation errors, a controlled scenario is implemented initially. An increasing number of devices is deployed, and the channel occupancy is evaluated both with the sniffer and a real time signal analyzer for comparison purposes. Also, it is possible to compute the theoretical occupancy in this controlled scenario, so this is included in the study to check accuracies as well.

To isolate the devices from other external actors, the measurements are done in a Faraday box (see Fig. 5). Moreover, all BLE devices are configured to generate advertising packets with a fixed length of $T_{\text{adv}} = 304$ μs with an advertising interval $T_{\text{advInt}} = 100$ ms. Note that the time between two advertising events is T_{advInt} plus a random delay (τ_{advDelay}) uniformly distributed between 0 and 10 ms. Thus, the mean time between advertisements is ($\overline{x}$ denotes mean of x):

$$\overline{T_{\text{advInt}}} = \mathrm{T}_{\text{advInt}} + \overline{\tau_{\text{advDelay}}} = 105 \text{ ms}. \quad (1)$$

Due to pre and post packet sequences, real on-air time (occupancy time for one advertisement) is $T_{\text{on-air}} = 360$ μs in this setup.

Hence, we create a predictable scenario with computable theoretical occupancy. When n devices are generating advertisements, the occupancy probability is one minus the probability that all of them are not occupying the channel:

$$P_{\text{Oc}n} = 1-(1-P_{\text{Oc1}})^n, \quad (2)$$

where P_{Oc1} is the channel occupancy probability of an advertising device, computed as the quotient of its activity (on-air) time and the mean inactive time:

$$P_{\text{Oc1}} = \frac{T_{\text{on-air}}}{\overline{T_{\text{advInt}}}}. \quad (3)$$

Fig. 6 shows the theoretical percentage of channel occupancy for an increasing number of devices (labeled as *Math model*). It can be seen that the real time signal analyzer shows an almost perfect match.

On the other hand, the curve denoted as *Sniffer Nordic/Texas raw* represents the occupancy provided by the sniffers. Both chips from Nordic (nRF52840 and nRF52832) performed identically so a single line is plotted for them (*Sniffer Nordic raw*). Anyways, results with the Texas sniffers are almost identical with just 1 percentage point of difference when 100 devices are active. It can be observed that sniffers perform with an absolute error of 1.1 percentage points for occupancies of 5% and smaller. But the result tends to diverge from the real value as the load increases. An underestimation error is present, and it is directly proportional to occupancy. This is due to the effects previously analyzed. To overcome such underestimation a compensation procedure is proposed:

The occupancy probability measured by the sniffer $P_{\text{Oc,sniffer}}$ is

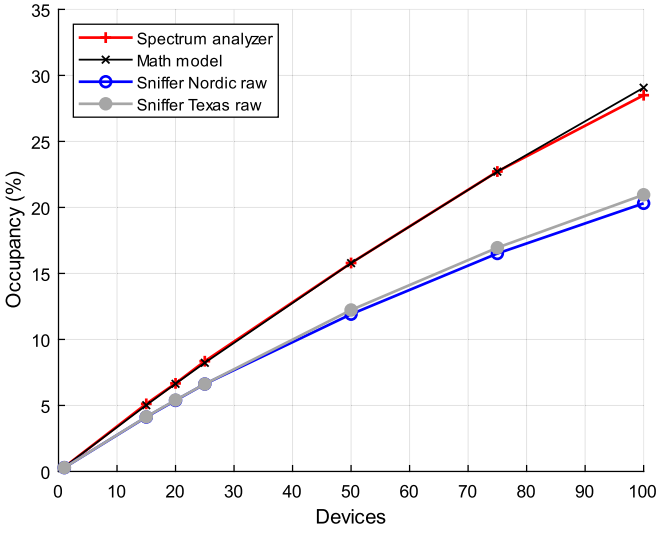


Fig. 6. Analytical and experimental occupancy for an increasing number of devices.

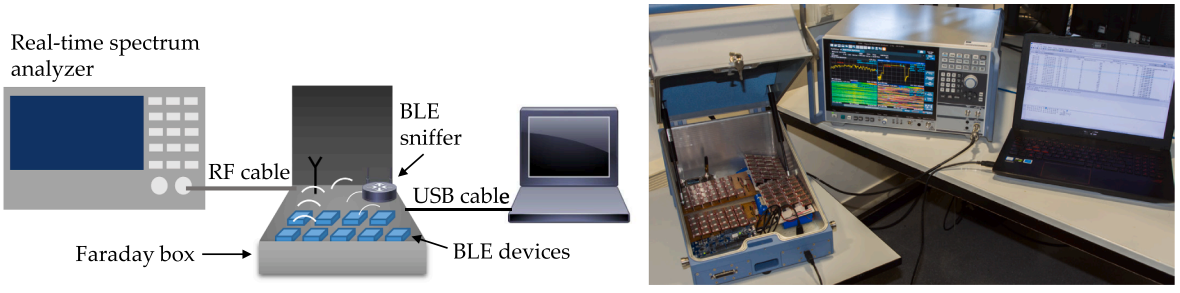


Fig. 5. In lab measurements are performed in a controlled environment in a Faraday box.

computed from the captured packets and their duration T_{adv} during an observation time, T_{obs}. For example, if load variations happen in a scale of seconds, T_{obs} could be settled to 1 s and overlap the observed intervals to offer a temporal resolution of microseconds. In any case, the observation time must be chosen according to the desired level of variability to be characterized. Note that the sniffer does not provide the real on-air time but just the net packet time, T_{adv}.

From $P_{Oc,sniffer}$, it is possible to obtain the *number of equivalent devices*, n_1, that would generate such occupancy if their T_{adv} was the average duration of the captured packets, $\overline{T_{adv,obs}}$, and their $\overline{T_{advInt}}$ was the observation time (T_{obs}).

$$P_{Oc,sniffer} = 1 - \left(1 - \frac{\overline{T_{adv,obs}}}{T_{obs}}\right)^{n_1} \rightarrow n_1 = \frac{\log\left(1 - P_{Oc,sniffer}\right)}{\log\left(1 - \frac{\overline{T_{adv,obs}}}{T_{obs}}\right)}. \quad (4)$$

Note that n_1 is affected by the sniffer occupancy underestimation due to collisions and blind times. So, by estimating the *PER*, it is possible to obtain a more precise number of equivalent devices, n_2. The *PER* equals the collision rate (*CR*) when collisions imply high interference and so insufficient SIR. In this sense, given a device that sends an advertisement at a time instant t, right after the pre advertisement sequence, it will collide with another, if the latter transmits in the interval $[t - T_{on-air}, t + T_{adv}]$. Hence, for the equivalent n_1 devices:

$$CR = 1 - \left(1 - \frac{T_{adv} + T_{on-air}}{T_{obs}}\right)^{n_1 - 1} \quad (5)$$

However, as previously explained, some of the advertisements can be captured depending on the SIR and the final collision overlap. Then, a more realistic perceived *PER* could be defined as in (6) where ξ is a reduction factor which is commented later.

$$PER = CR \cdot \xi \quad (6)$$

From here we can now obtain n_2,

$$n_1 = (1 - PER) n_2 \rightarrow n_2 = \frac{n_1}{1 - PER} \quad (7)$$

And, thus, the estimated occupancy probability $P_{Oc,est}$ can be calculated as:

$$P_{Oc,est} = 1 - \left(1 - \frac{T_{on-air}}{T_{obs}}\right)^{n_2} \quad (8)$$

This is labeled as *Sniffer processed* in Fig. 7 (for both the Nordic and Texas case). Results are displayed for a ξ range from 0.5 to 0.8. The shaded area precisely indicates the region where the actual occupancy is located. To facilitate the comparison, all the curves in Fig. 6 are also plotted herein. For occupancies of 30% an error of 3 percentage points is made by excess and an underestimation error of 2.5 percentage points for the worst case. On the other hand, for occupancies of 10% the absolute error committed is under ±0.6 percentage points.

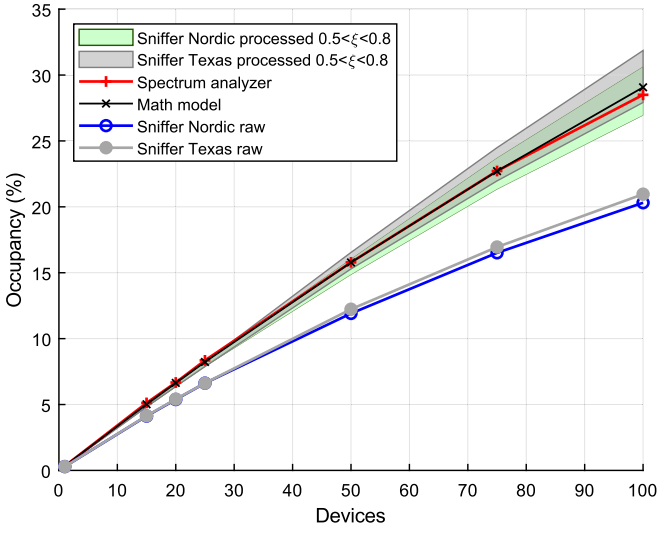


Fig. 7. Analytical and experimental occupancy (raw and processed) for an increasing number of devices.

Regarding the reduction factor ξ, it would be possible to provide an exact value that would obtain the actual occupancy, but this number would be scenario specific. Given the randomness in actual packet sizes, number of devices and the eventual occupancy of each device, it is considered more advisable to handle a range for ξ. Also, given the slightly lower underestimation error of the Texas sniffer, the range could be smaller for this second case. Nevertheless, as a rule-of-thumb, a ξ of 0.8 allows to have a slight overestimation and can be taken as a reference value that allows to play safe. For example, in cases where occupancy measurements are needed to configure a newly added Bluetooth mesh network.

The use of this correction procedure based on collision rate estimations validates the utilization of sniffers as occupancy meters to be employed in scenarios where real time signal analyzers are not adequate, as previously argued.

4. Results

4.1. Measurement campaigns in campus.

Initially, occupancy measurements have been made at different locations on a university campus. Areas with different levels of close to constant load and with variations have been identified, see subfigure titles in Fig. 8. Note that in this and the following results, the nRF52840 chip from Nordic has been used, which is the case that showed slightly more underestimation error.

Given the perfect matching between the real time signal analyzer and the theoretical results previously observed, the device results are used as benchmark. However, the measurements must be limited to 80 s, since this is the maximum time allowed by the analyzer when operating with maximum temporary resolution, as discussed above.

Plots show both the raw result provided by the sniffer itself and the occupancy value after applying the proposed processing. The results are satisfactory. As previously observed, major errors appear with higher occupancy. For instance, with high loads (Fig. 8(a), occupancy around 40%), the absolute error of the upper bound ($\xi = 0.8$) is less than 4 percentage points in 71% of the cases and it is less than 5 points in the 95% of the cases. With average load (Fig. 8(b), occupancy around 27%), the 85% of the errors are now under 3 percentage points and the percentile 95% reduces to 3.5 points. Finally, with occupancies of 15% (not plotted), percentile 85 reduces to 1.5 percentage points and percentile 95 to 1.9 points.

Fig. 8(c) and (d) show how the measurement tracks variations correctly and how the result becomes more accurate when load decreases. For the lowest loads, it is possible to observe that real occupancy crosses with the lower bound estimation. This means that a value of ξ of 0.5 may be insufficient when very low loads happen and the lower bound is indeed higher than real occupancy. However, the error is so small that it is not considered interesting to suggest a ξ range starting at 0.4 or 0.45, as this would result in larger errors with high load, where the method is more imprecise.

4.2. Measurement campaigns in real environments.

Once the strategy has been validated, this last section shows measurements campaigns carried out in out-of-campus environments. The first set of measures has been carried out on two subway platforms. In those scenarios, load variations can be very abrupt in short intervals of time. However, short term alterations and occupancy peaks can be captured thanks to the high resolution allowed by sniffers.

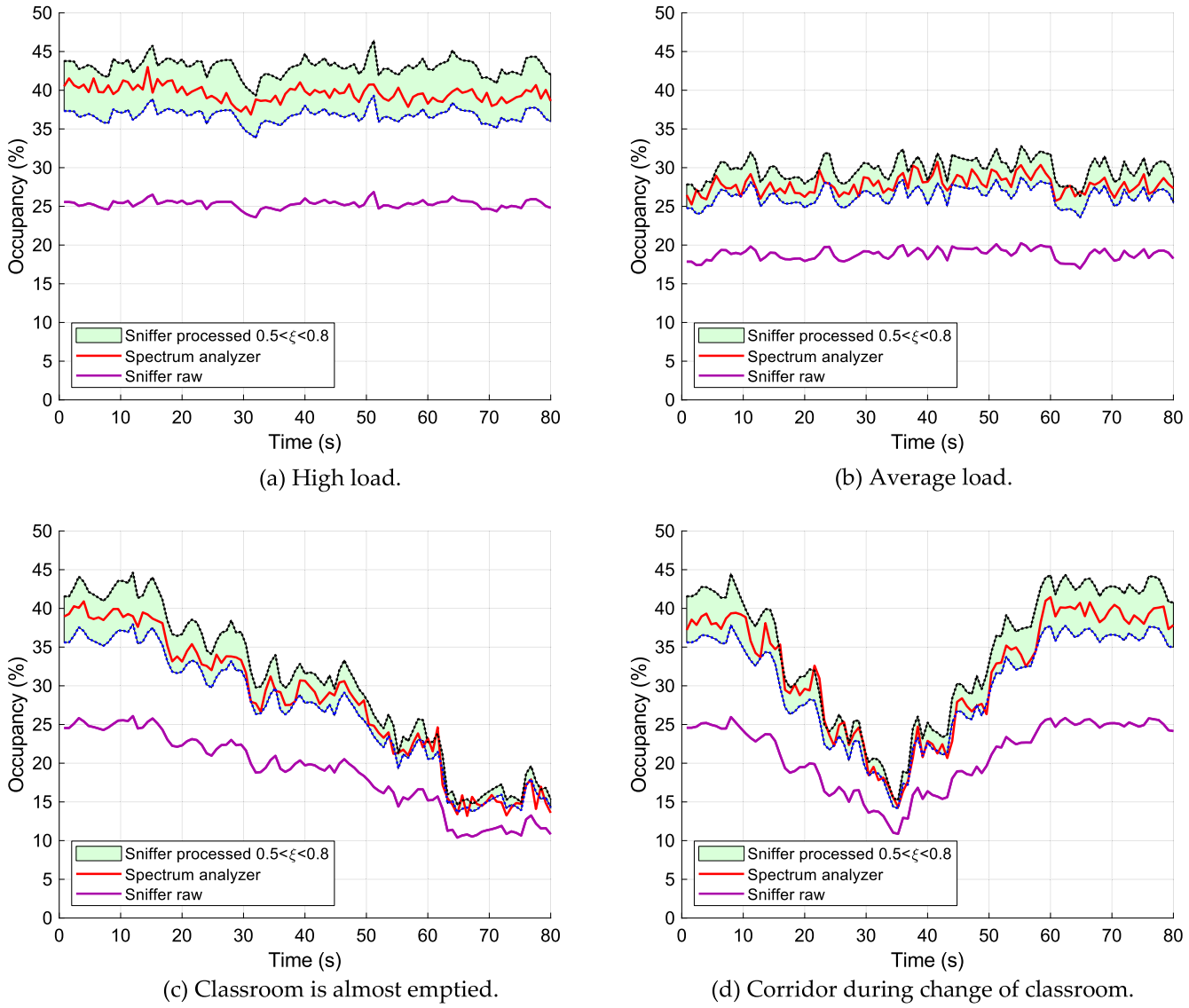


Fig. 8. Occupancy measurements made with sniffer and signal analyzer in a campus environment. Measurements are shown in places with different load levels and levels of variability.

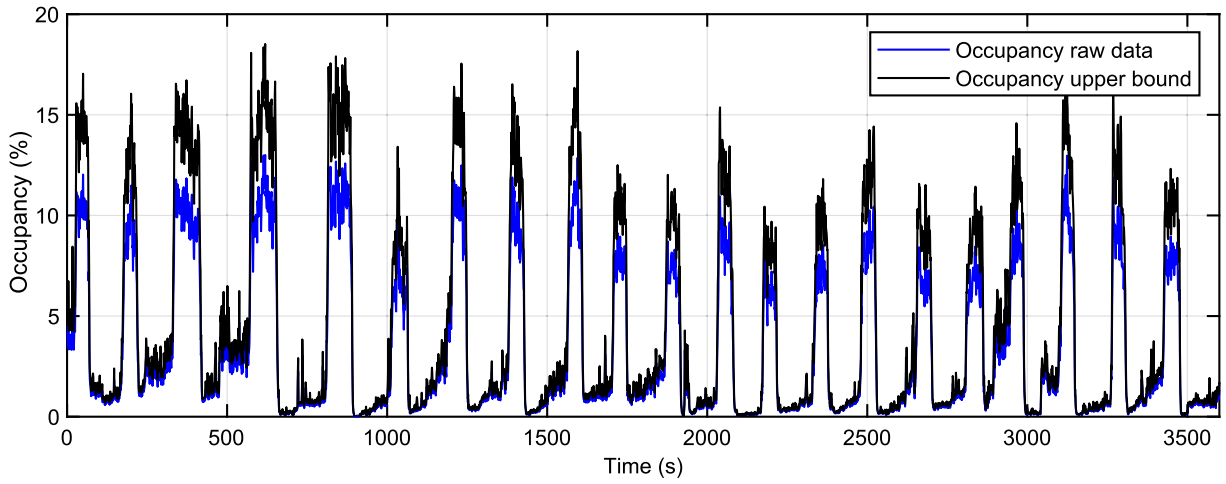
Fig. 9 shows occupancy estimation for the two platforms, recorded during 3600 s. Both the raw and processed data are plotted. Resolution allows to identify the load peaks every time a new train arrives and how long it remains in the platform. Once the platform is emptied, it can be seen a progressive accumulation of people waiting until a new load peak (new train) arrives. Occupancies can reach 18% so it can be claimed that advertisements are not saturating the allocated spectrum yet. However, this value implies that a very relevant PER is to be experienced. Indeed, the occupancy value establishes a lower bound estimation of the PER. This is because the most favorable situation happens when the SIR is sufficient to recover a packet in a collision. Here, packets are only lost when the receiver is busy decoding a previous advertisement. In these circumstances, the PER is directly determined by the probability of occupancy of the channel itself, $PER_{n,\min} = P_{\mathrm{Oc}n} = 1-(1-P_{\mathrm{Oc}1})^n$.

Occupancy was also measured in a shopping mall. This case shows much slower variations, but periodicities can be found in a daily basis. Results were recorded for the same hour at different weekdays, showing a weekly pattern. This is depicted by Fig. 10, which shows the mean, maximum and minimum values that were recorded during the experiment. Note, in the horizontal axis, that the two sets of data correspond to two different weeks (indicated by two separate lines). For the sake of readability, only $\xi = 0.8$ is represented. In this scenario, a mean occupancy around 14% happens on Sundays when the shops in the mall are closed, and the maximum recorded value is close to 35% on Saturdays.

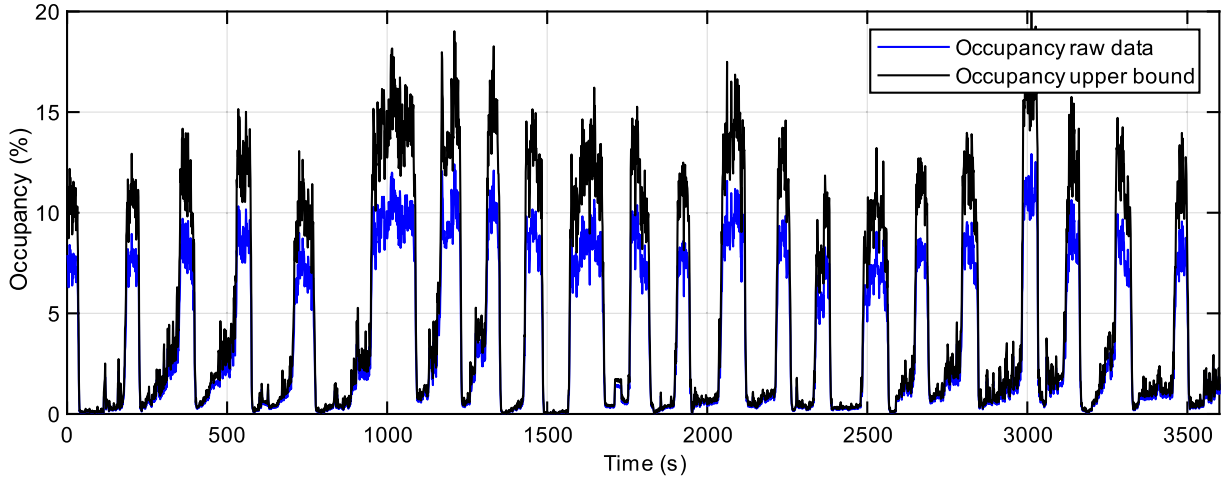
As a note aside, it is interesting to observe that social behaviors can be analyzed from occupancy patterns [21]. For example, accumulation of people in train platforms at different moments of the day and depending on train frequency and/or special events. Another example, during the measurement campaigns at the mall, three days rained, and occupancy was reduced (Wed, Thu, Fri points in the second week). Hence, it is feasible to investigate the impact of different weather phenomena or other external causes on the influx of people to shopping malls.

5. Conclusions

An accurate measurement and characterization of BLE channel occupancy imposes more stringent requisites to the methods that are usually employed for other technologies. Many works that study channel occupancy rely on spectral analysis and employ large time resolution relative to the size of BLE advertising packets. This implies too long measurement campaigns and not being able to track short-term



(a) Subway platform 1.



(b) Subway platform 2.

Fig. 9. Occupancy evaluation in subway platforms by means of post-processed sniffer data.

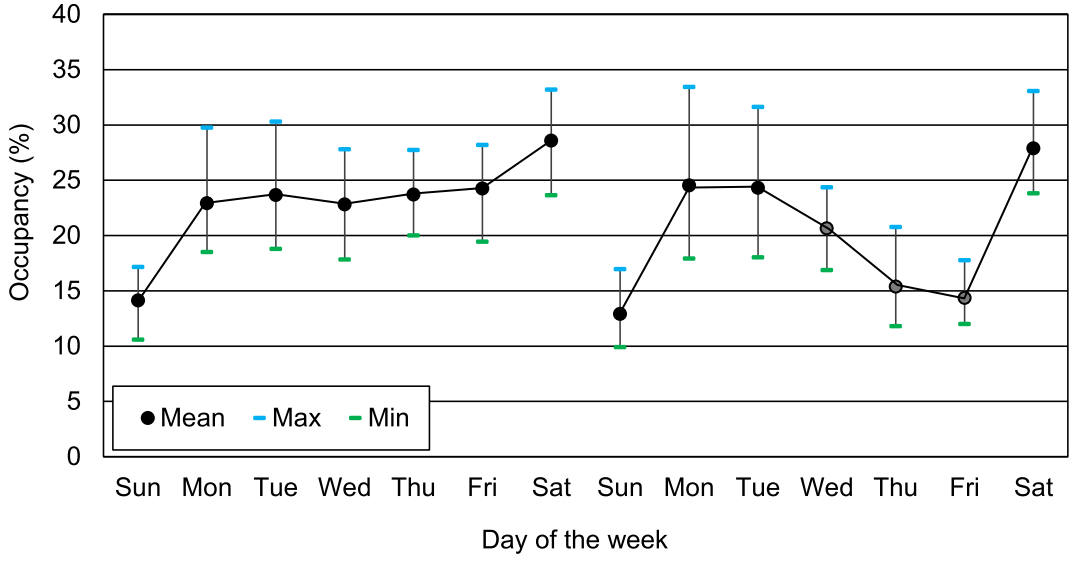


Fig. 10. Mean, maximum and minimum occupancy in mall for the same hour and for two weeks.

variations. Real time signal analyzers are a valid option but show a compromise between accuracy and required measurement interruptions. Also, their large volume, weight, cost, transport difficulties and the need for an energy source, make them an impractical tool for many out-of-lab situations.

BLE sniffers are proposed to measure occupancy in near-real time, thus being able to capture very short-term variations, recording for hours, without interruptions. Also, it is remarkable their low cost and excellent portability when combined with a laptop. The paper argues about specific problems that must be addressed for this approach to be successful, namely buffer saturation, channel anchoring and more importantly, the exact on-air time of a given packet and the existence of receiving blind times due to processing and packet overlapping.

Results in a controllable lab scenario allow to quantify underestimation errors made with *raw* sniffer data. The error is directly proportional to occupancy, and sniffers just perform correctly for low loads, with an absolute error of 1.1 percentage points for occupancies of 5%. But the error is too large for higher occupancies. Hence, a compensation procedure is proposed to extend the use of sniffers to higher load levels being based on collision rate estimations.

Results with the refined measurement procedure show that occupancies of 40% can be measured in real time with an overestimation error whose percentile 95% is 5 percentage points. This is reduced to 1.9 points when the occupancy is 15%. The sniffers are shown to perfectly track short term occupancy variations. The strategy has been successfully used to characterize occupancy in highly variable and loaded scenarios such as subway platforms and a shopping mall. Occupancies of 18%-35% are experienced, and those values establish a lower bound of the *PER*. Thus, the measurements show that BLE technology is not saturating the spectrum yet, but a relevant *PER* is to be experienced. Hence, the tool can be used to make agile audits and configure the parameters that control communication redundancy in new or existing BLE meshed networks and/or the mesh layout itself. The measurement proposal allows to extend its usefulness beyond identifying communication problems. Examples are the analysis of social behavior, performance of services, and so on.

CRediT authorship contribution statement

A. Valenzuela-Pérez: Investigation, Formal analysis, Data curation, Writing – review & editing. **M. García-Lozano:** Conceptualization, Investigation, Formal analysis, Data curation, Methodology, Writing – original draft, Funding acquisition, Supervision. **J.L. Valenzuela:** Conceptualization, Investigation, Formal analysis, Data curation, Methodology, Writing – review & editing. **D. Pérez-Díaz-de-Cerio:** Investigation, Formal analysis, Methodology, Writing – review & editing. **Á. Hernández-Solana:** Investigation, Validation, Writing – review & editing, Funding acquisition. **A. Valdovinos:** Investigation, Validation, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The work by UPC has been funded by MCIN/ AEI /10.13039/501100011033 and by ERDF A way of making Europe, with the grant RTI2018-099880-B-C32 and PID2021-125799OA-I00.

The work by I3A-UZ has been funded by MCIN/ AEI /10.13039/501100011033 and by ERDF A way of making Europe, with the grants RTI2018-095684-B-I00 and RTI2018-099063-B-I00, and by the Government of Aragon (Reference Group T31 20R).

Appendix A

When the sampling time is longer than the duration of a BLE advertising message, some packets are lost. The required number of samples to obtain an acceptable estimate of channel occupancy is calculated next.

In this situation, and especially for very high sampling times, the state of the channel can be considered to be independent between samples. Then, the occupancy X is a random variable with Bernoulli distribution and characterized by a mean value $\mu = p_{\text{occ}}$ and a standard deviation $\sigma_{\text{occ}} = \sqrt{p_{\text{occ}}(1-p_{\text{occ}})}$, being p_{occ} the probability that the channel is occupied. When the measurement campaign is performed, a single mean occupancy m is obtained. But, according to the central limit theorem, the means of multiple experiments exhibit a normal distribution whose mean value is precisely μ and its deviation is $\sigma_{\text{occ}}/\sqrt{n}$, where n is the number of samples. Hence, the probability C that the mean occupancy of an experiment is in the interval $(\mu-\delta, \mu+\delta)$ can be expressed as:

$$P(\mu-\delta < m < \mu+\delta) = C = \text{erf}\left(\frac{z}{\sqrt{2}}\right). \tag{A1}$$

Where erf() is the error function and z is the standardized value (z-score) of $\mu+\delta$, i.e., $z = \delta\sqrt{n}/\sigma_{\text{occ}}$. Equivalently, it is possible to rewrite the expression as follows:

$$P(m-\delta < \mu < m+\delta) = P\left(m - z\frac{\sigma_{\text{occ}}}{\sqrt{n}} < \mu < m + z\frac{\sigma_{\text{occ}}}{\sqrt{n}}\right) = C. \tag{A2}$$

This indicates that the actual mean occupancy μ is within the range $(m-\delta, m+\delta)$ with a confidence level equal to C. By setting a value for the maximum desired error δ, it is possible to obtain the minimum number of samples needed:

$$\delta = z\frac{\sigma_{\text{occ}}}{\sqrt{n}} = z\sqrt{\frac{p_{\text{occ}}(1-p_{\text{occ}})}{n}}, \tag{A3}$$

$$n = \frac{z^2}{\delta^2}p_{\text{occ}}(1-p_{\text{occ}}). \tag{A4}$$

The occupancy is the parameter to be estimated, therefore, the worst case is calculated by considering the maximum possible uncertainty, $p_{\text{occ}} = 0.5$. In addition, taking a typical confidence level of $C = 0.9973$ (three sigmas) implies $z = 3$, therefore:

$$n \geq \frac{2.25}{\delta^2}. \tag{A5}$$

For example, if it is needed to estimate an occupancy of 10% with a relative error of 5% (absolute error, $\delta = 0.005$), $n \geq 90000$.

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.measurement.2022.111573.

References

[1] Bluetooth SIG. Bluetooth 2021 Market Update. Available online: https://www.bluetooth.com/wp-content/uploads/2021/01/2021-Bluetooth_Market_Update.pdf (accessed on 15/02/2022).

[2] K.E. Jeon, J. She, P. Soonsawad, P. Chet Ng, BLE beacons for internet of things applications: survey, challenges and opportunities, IEEE Internet Things J. 5 (2) (2018) 811–828, https://doi.org/10.1109/JIOT.2017.2788449.

[3] P.J. García-Paterna, A.S. Martínez-Sala, J.C. Sánchez-Aarnoutse, Empirical study of a room-level localization system based on bluetooth low energy beacons, Sensors 21 (2021) 3665, https://doi.org/10.3390/s21113665.

[4] R. Belka, R.S. Deniziak, G. Łukawski, P. Pięta, BLE-based indoor tracking system with overlapping-resistant IoT solution for tourism applications, Sensors 21 (2021) 329, https://doi.org/10.3390/s21020329.

[5] N. Papadakis, N. Koukoulas, I. Christakis, I. Stavrakas, D. Kandris, An IoT-based participatory antitheft system for public safety enhancement in smart cities, Smart Cities 4 (2021) 919–937, https://doi.org/10.3390/smartcities4020047.

[6] Federal Aviation Administration, United States Department of Transportation. UAS Remote Identification Overview. Available online: https://www.faa.gov/uas/getting_started/remote_id/ (accessed on 15/02/2022).

[7] G. Pau, F. Arena, Y.E. Gebremariam, I. You, Bluetooth 5.1: an analysis of direction finding capability for high-precision location services, Sensors 21 (11) (2021) 3589, https://doi.org/10.3390/s21113589.

[8] A. Hernandez-Solana, D. Perez-Diaz-De-Cerio, M. Garcia-Lozano, A.V. Bardaji, J.-L. Valenzuela, Bluetooth mesh analysis, issues, and challenges, IEEE Access 8 (2020) 53784–53800, https://doi.org/10.1109/ACCESS.2020.2980795.

[9] L. Mucchi, R. Vuohtoniemi, H. Virk, A. Conti, M. Hamalainen, J. Iinatti, M.Z. Win, Spectrum occupancy and interference model based on network experimentation in hospital, IEEE Trans. Wireless Commun. 19 (9) (2020) 5666–5675, https://doi.org/10.1109/TWC.2020.2995116.

[10] J. Kokkoniemi, J. Lehtomäki, Spectrum Occupancy Measurements and Analysis Methods on the 2.45 GHz ISM Band. Proceedings of the 7th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Stockholm, Sweden, 2012. https://doi.org/10.4108/icst.crowncom.2012.248307.

[11] A. Al-Hourani, V. Trajković, S. Chandrasekharan, S. Kandeepan. Spectrum Occupancy Measurements for Different Urban Environments, in: Proceedings of the 2015 European Conference on Networks and Communications (EuCNC), Paris, France, 2015, pp. 97-102. https://doi.org/10.1109/EuCNC.2015.7194048.

[12] M. Cardenas-Juarez, M.A. Diaz-Ibarra, U. Pineda-Rico, A. Arce, E. Stevens-Navarro, On spectrum occupancy measurements at 2.4 GHz ISM band for cognitive radio applications, in: Proceedings of the 2016 International Conference on Electronics, Communications and Computers (CONIELECOMP), Cholula, Mexico, 2016, pp. 25-31. https://doi.org/10.1109/CONIELECOMP.2016.7438547.

[13] A.A. Cheema, S. Salous, Spectrum Occupancy Measurements and Analysis in 2.4 GHz WLAN, Electronics 8 (2019) 1011, https://doi.org/10.3390/electronics8091011.

[15] Q.D. La, D. Nguyen-Nam, M.V. Ngo, H.T. Hoang, T.Q.S. Quek, Dense deployment of BLE-BASED BODY AREA NETWORKS: A COEXISTENCE STUDY, IEEE Trans. Green Commun. Netw. 2 (4) (2018) 972–981, https://doi.org/10.1109/TGCN.2018.2859350.

[16] A. Ancans, J. Ormanis, R. Cacurs, M. Greitans, E. Saoutieff et al., Bluetooth Low Energy Throughput in Densely Deployed Radio Environment, in: Proceedings of the 23rd International Conference Electronics, Palanga, Lithuania, 2019. https://doi.org/10.1109/ELECTRONICS.2019.8765577.

[17] R. Natarajan, P. Zand, M. Nabi, Analysis of coexistence between IEEE 802.15.4, BLE and IEEE 802.11 in the 2.4 GHz ISM band, in: Proceedings of the 42nd Annual Conference of the IEEE Industrial Electronics Society (IECON), Florence, Italy, 2016, pp. 6025-6032, https://doi.org/10.1109/IECON.2016.7793984.

[18] M.O. Al Kalaa, W. Balid, N. Bitar, H.H. Refai, Evaluating bluetooth low energy in realistic wireless environments, in: Proceedings of the 2016 IEEE Wireless Communications and Networking Conference (WCNC), Doha, Qatar, 2016. https://doi.org/10.1109/WCNC.2016.7564809.

[19] S.V. Kizima, V.A. Kozmin, A.B. Tokarev, The advantages of using the absolute measurement error when estimating the occupancy of the radio-frequency spectrum, Meas. Tech. 55 (5) (2012) 568–573, https://doi.org/10.1007/s11018-012-0003-2.

[20] D. Perez-Diaz de Cerio, Á. Hernández, J.L. Valenzuela, A. Valdovinos, Analytical and experimental performance evaluation of BLE neighbor discovery process including non-idealities of real chipsets, Sensors 17 (2017) 499, https://doi.org/10.3390/s17030499.

[21] A. Pratama, W. Widyawan, A. Lazovik, M. Aiello, Multi-user low intrusive occupancy detection, Sensors 18 (3) (2018) 796, https://doi.org/10.3390/s18030796.