Revealing Spectrum Allocation Scheme and Temporal Transmission Behavior of IoT Devices using Passive Packet Sniffing

Ahmed Abdelghany, Bernard Uguen, Christophe Moy, Jérôme Le Masson, Francois Marie

Revealing Spectrum Allocation Scheme and Temporal Transmission Behavior of IoT Devices using Passive Packet Sniffing

Ahmed Abdelghany, Bernard Uguen, Christophe Moy, *Senior Member, IEEE*, Jérôme Le Masson, François Marie
Univ Rennes, CNRS, IETR - UMR 6164
F-35000, Rennes, France
{ahmed.abdelghany, bernard.uguen, christophe.moy, jerome.lemasson, francois.marie}@univ-rennes.fr

***Abstract*—The growing prevalence of Internet of Things (IoT) technologies in many outdoor applications has led to an increase in the deployment of popular LoRaWAN (Long Range Wide Area Network) networks. As a result, there is increasing interest in tracking end nodes and detecting device transmission behavior. In this paper, a passive packet sniffer is demonstrated to receive LoRa (Long Range) packets in the Campus Beaulieu area of Rennes, France, over a period of two months. After processing the acquired packets, the inter-arrival times between packets of each identified device are estimated properly using a proposed technique, i.e., based on Kernel Density Estimation (KDE). Another algorithm is also employed to detect any pattern in the inter-arrival times. The study is then extended to reveal the spectrum allocation technique of a device and detect any periodic structure in the sequence of the used frequencies. Over and above, the spectrum allocation scheme is modelled as a Markov chain, allowing for the prediction of the next selected frequency using the estimated trajectory. The proposed algorithms are validated through statistical analysis of the inter-arrival times, temporal patterns, and spectrum allocation schemes' characteristics from the actual acquired data.**

***Index Terms*—Internet of Things (IoT), Low Power Wide Area Network (LPWAN), LoRa (Long Range), Transmission Behavior, Frequency Allocation, Pattern Detection, Packet Sniffing, Smart City, Network Traffic Monitoring.**

I. Introduction

Internet of Things (IoT) is increasing within different market segments, such as smart homes, security applications and remote sensing [1]. For that, Low Power Wide Area Network (LPWAN) paves the way to ensure the challenging communication requirements and energy constraints of IoT devices. Here, LoRaWAN (Long Range Wide Area Network) is considered one of the major techniques due to its ease of deployment and operation in the unlicensed frequency bands, i.e., $868\,\mathrm{MHz}$ in Europe and $915\,\mathrm{MHz}$ in the USA [2]. As a result, LoRa devices can function for multiple years without requiring a battery replacement.

Monitoring the LoRaWAN is required to use the network efficiently and plan for optimal utilization [3]. This monitoring process involves receiving the LoRa packets, acquiring their transmission parameters and identifying the active devices. Additionally, it is necessary to associate the traffic activity for estimating the transmission behavior of the IoT devices at a given geographical location. For instance, a packet sniffer can derive information by leveraging the temporal feature of the received packets. Besides the temporal behavior, the spectrum allocation technique of a device can be revealed from its transmission frequencies. Accordingly, this analysis can help an eavesdropper to have vital information about all the nearby nodes and then enable the identification of the received packets from a targeted node based on its unique transmission behavior.

Recently, some studies have examined actual LoRa packets received through practical measurements. For example, the authors in [4] use their software tools to perform an exploration of a dataset composed of LoRa packets. To provide an in-depth analysis of the landscape of commercial IoT deployments, [5] conducts a measurement study at eight key locations distributed over a large city. While in [6], an approach, named DEVIL (DEVice Identification and privacy Leakage), is introduced to detect the temporal patterns arising in the transmitted packets by LoRa devices. However, these previous works do not fully utilize the privileged information in the received packets and then do not investigate the entire transmission behavior of each individual end node. Otherwise, they are more focused on a single aspect such as estimating only the periodic temporal pattern of transmitted packets, while ignoring the frequency allocation procedure of the device. Over and above, they do not have well-defined criteria for identifying the packets belonging to the same device, relying solely on the non-unique Device Address (DevAddr) [7]. It is important to note that when an end node intends to join a network using the Over-the-Air Activation (OTAA) method, it sends a join request packet that contains a 64-bit Device Extended Unique Identifier (DevEUI), after which the network server assigns it a non-unique 32-bit DevAddr for all subsequent transmissions.

In this paper, a passive packet sniffer is implemented using readily available tools and a commercial LoRa gateway in the Campus Beaulieu area of Rennes, France. Over two

months, the network activity and the transmission parameters of the packets are explored. For classifying the packets transmitted by the same end node, a Sequence of Contiguous Packets (SCP), i.e., proposed in [8], is utilized as a device sub-identifier beyond the DevAddr. Consequently, the inter-arrival times between packets are properly estimated for every single device and then statistically analyzed. Furthermore, the temporal patterns and the spectrum allocation algorithm of the devices are detected, as detailed in the upcoming sections.

The remainder of this document is organized as follows. Section II gives an overview of the system setup. Section III provides the detection of the inter-arrival times and temporal patterns. Then, the frequency allocation is investigated and modelled in Section IV. In Section V, the experimental results of the proposed algorithms are presented and commented on. Finally, the work is concluded in Section VI.

II. Proposed Passive Packet Sniffer

A. Overview and System Configurations

In a LoRaWAN network with a star network topology, a standard LoRa end node sends an uplink message that can be received by all gateways within its coverage area. Taking advantage of this, the proposed experiment uses a gateway to passively log all LoRa packets sent by nearby end nodes, as illustrated in Figure 1.

To accomplish this, a Tektelic KONA Macro Gateway is connected to a fixed antenna on the university building's roof [9]. On this gateway, the ChirpStack Gateway Bridge service, i.e., part of the ChirpStack open-source LoRaWAN Network Server stack [10], is configured to publish the packets' logs to an Message Queuing Telemetry Transport (MQTT) broker [11]. To receive immediately the packets through the internet, a desktop computer, which has Node-RED software installed on it [12], is utilized. Here, a Node-RED flow is implemented that starts by subscribing to the aforementioned MQTT broker, and then the packets are streamed to a Python script. This Python script uses the Lora-Packet decoder library, i.e., introduced in [13], for decoding the LoRa physical payload. Accordingly, the packet is timestamped and appended to a data file for further analysis, as detailed in the following sections. Additionally, these data are made available to the research community on this online repository [14].

B. Data Extraction

The gateway provides the transmission parameters for each packet $\mathbf{m}[n]$, identified by its packet index n. Additionally, decoding the Lora physical payload gives essential information from the frame header about the device identifiers, message type, packet counter, etc. Accordingly, Table I summarizes the main acquired parameters. Once these parameters are extracted, SCPs are determined and then information about the characteristics of each SCP can be obtained such as the SCP length (number of packets), Effective Signal Power (ESP) statistics (mean, variance, etc.) at each frequency band $f[n] \in$

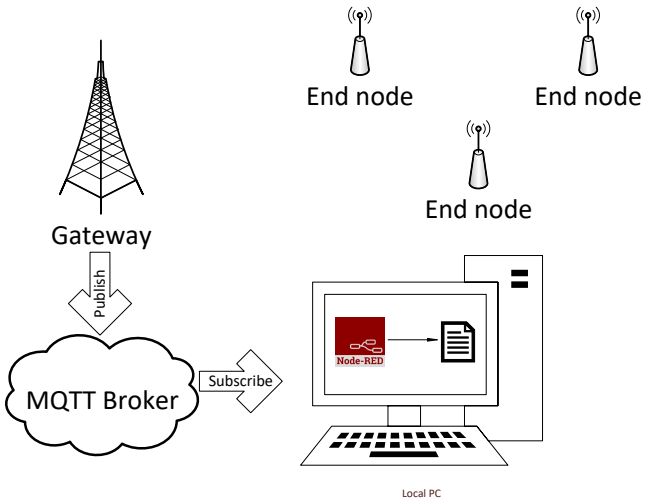


Fig. 1: Setup of the proposed packet sniffer.

$\{867.1, 867.3, 867.5, 867.7, 867.9, 868.1, 868.3, 868.5\}$ MHz [1]. Additionally, the following sections demonstrate how to acquire critical features related to the monitored devices' temporal behavior and frequency allocation techniques.

TABLE I: Description of some fields for a received packet $\mathbf{m}[n]$

Field	Type	Description
$t[n]$	float	Time at which the packet is received by the gateway
$f[n]$	float	The center frequency of the received packet
$ESP[n]$	float	The Effective Signal Power of the received packet
$a[n]$	str	The Device Address (DevAddr) of the received packet
$FCnt[n]$	int	Counter of the transmitted packets, which corresponds to the total number of transmitted packets (by a node) from the beginning until the received packet

III. Temporal Behavior

A. Inter-arrival Time

The temporal analysis of the received packets is essential for disclosing the nature of application and then identifying a targeted device. For that, let $t[i]$ be the timestamp for a packet $\mathbf{m}_a^l[i]$, assigned to DevAddr a and belongs to SCP index l, with a frame counter $FCnt[i]$. Then, the time between two consecutive packets, named inter-arrival time, is computed as:

$$\Delta[i] = t[i+1] - t[i]. \tag{1}$$

Accordingly,

$$\mathbf{\Delta} = \left[\Delta[0], \ldots, \Delta[i], \ldots, \Delta[N_a^l - 2]\right], \tag{2}$$

where N_a^l is the total number of packets in a given SCP index l.

Based on the experimental observations, it is noted that the inter-arrival times $\mathbf{\Delta}$, which are colored by red in Figure 2,

are very often composed of multiple distinct values rather than one value. Hence, the inter-arrival times can be modeled as a multimodal distribution. Based on that, this distribution is estimated using Gaussian Kernel Density Estimator (KDE) $f(x)$ as:

$$f(x) = \frac{1}{(N_a^l - 1)h\sqrt{2\pi}} \sum_{i=0}^{N_a^l - 2} e^{-0.5(\frac{x - \Delta[i]}{h})^2} \tag{3}$$

with

$$h = 0.9 \min\left(\sigma_{\mathbf{\Delta}}, \frac{IQR}{1.34}\right)(N_a^l - 1)^{-\frac{1}{5}}, \tag{4}$$

where h is a smoothing parameter, called the bandwidth, which is calculated using Silverman's rule-of-thumb [15]. Accordingly, $\sigma_{\mathbf{\Delta}}$ and IQR are the standard deviation and the interquartile range of the entire inter-arrival times $\mathbf{\Delta}$, respectively. Figure 2 illustrates an actual example of a device with three main inter-arrival times. This reveals that this device sends a packet every $\approx 20\,\text{min}$, $\approx 40\,\text{min}$ or $\approx 1\,\text{h}$. Thus, the position of each peak is detected using a peak detector, then, the statistical distribution (Mean, variance, etc.) is obtained for each significant inter-arrival time.

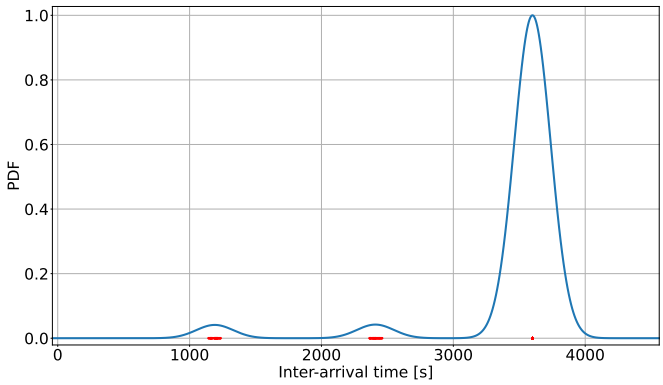


Fig. 2: Three distinct inter-arrival times are estimated using KDE.

B. Pattern of the Inter-arrival Time

After focusing on the elapsed time between a packet and the next one transmitted from the same device, it is noted that many devices send their packets with a periodic temporal pattern. As depicted in Figure 3, the period length L of a pattern is not known a priori. As illustrated in [6], to acquire the period length L, the sequences of inter-arrival times are first extracted as follows:

$$\mathbf{d}[j] = \left[\Delta[j], \Delta[\hat{L} + j], \Delta[2\hat{L} + j], \ldots\right] \forall j \in \{0, \ldots, \hat{L} - 1\} \tag{5}$$

For each of the sequence $\mathbf{d}[j]$, the standard deviation $\sigma_{\mathbf{d}[j]}$ is calculated. Then, all the estimated standard deviations are averaged as:

$$\epsilon_{\hat{L}} = \frac{1}{\hat{L}} \sum_{j=0}^{\hat{L}-1} \sigma_{\mathbf{d}[j]}, \tag{6}$$

where $\epsilon_{\hat{L}}$ is the estimation error for the period length $\hat{L}$. Accordingly,

$$\mathbf{E} = \left[\epsilon_1, \epsilon_2, \epsilon_3, \ldots\right]. \tag{7}$$

Here, the estimation error $\epsilon_{\hat{L}}$ gives a low value for the correct estimated period length $\hat{L}$, and it is high otherwise.

In this paper, a shuffled version of $\mathbf{\Delta}$ is used as a reference to compare its estimation error $\mathbf{E}^s = \left[\epsilon_1^s, \epsilon_2^s, \epsilon_3^s, \ldots\right]$ versus the original unshuffled $\mathbf{\Delta}$, as shown in Figure 4. One can observe that the error values of the unshuffled $\mathbf{\Delta}$ overlap with the shuffled version, as depicted in Figure 4a. Otherwise, when a periodic structure exists, the error values of the unshuffled $\mathbf{\Delta}$ are lower than shuffled ones at the multiples of the period length, as the case in Figure 4b. Consequently, the estimation of this period length $\hat{L}$ can be described as follows in Algorithm 1.

Data: $\mathbf{E}, \mathbf{E}^s$
Result: $\hat{L}$
$\mathbf{P} = []$
for $l = 1$ *to* $length(\mathbf{E})$ **do**
  C_1 = $(\epsilon_l < \epsilon_l^s)$
  C_2 = $(\epsilon_l < \min \mathbf{E}^s)$
  if $(C_1 \wedge C_2)$ **then**
    $\mathbf{P}.append(l)$
  end
end
$\mathbf{PP} = \mathbf{P}[(\mathbf{P} \mod \mathbf{P}[0]) == 0]$
C_3 = $(length(\mathbf{PP}) == \lfloor \frac{length(\mathbf{E})}{P[0]} \rfloor)$
if C_3 **then**
  $\hat{L} = \mathbf{P}[0]$
else
  $\hat{L} = 1$
end

Algorithm 1: A heuristic of the period length $\hat{L}$ estimation

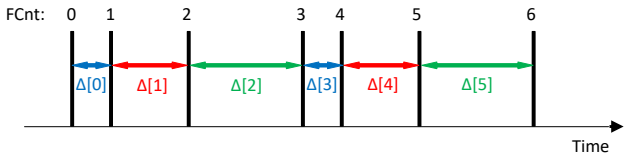
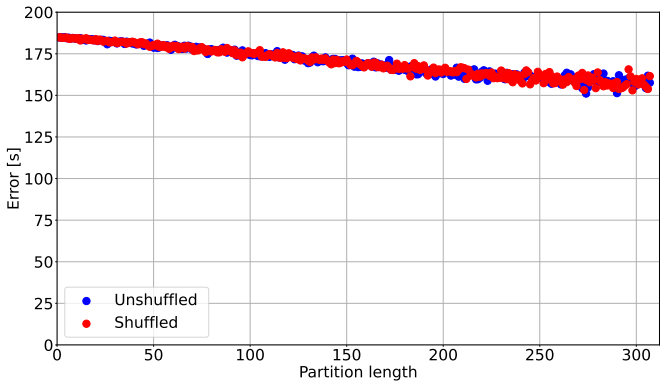


Fig. 3: An example of pattern of inter-arrival times with period length L of 3, whereas the packet counter is indicated above each packet.

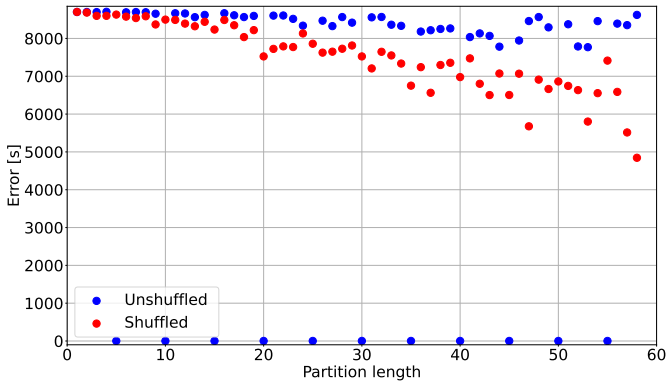
IV. Frequency Allocation Scheme

A. Pattern of the Frequency Allocation

The work here is extended to detect any pattern in the spectrum allocation of a device. Hence, to acquire the period

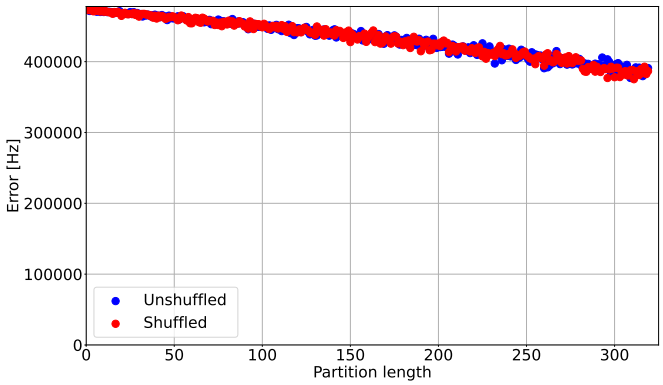


(a) No pattern is detected.

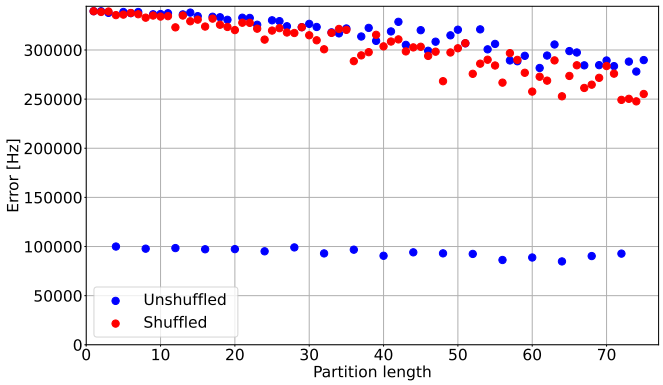


(b) A pattern is detected with period length of 5.

Fig. 4: Pattern detection in the inter-arrival times.



(a) No pattern is detected.



(b) A pattern is detected with period length of 4.

Fig. 5: Pattern detection in the frequency allocation.

length L, the sequences of the used frequencies are first extracted as follows:

$$\mathbf{f} = \left[f[0], \ldots, f[i], \ldots, f[N_a^l - 1]\right]. \tag{8}$$

Hence,

$$\mathbf{c}[j] = \left[f[j], f[\hat{L} + j], f[2\hat{L} + j], \ldots\right] \forall j \in \{0, \ldots, \hat{L} - 1\}. \tag{9}$$

At this point, almost the same technique, i.e., previously demonstrated in Section III-B, is applied to reveal any periodic structure in the sequence of selected frequencies. Figure 5 shows that the estimated error is small when a pattern is detected. Nevertheless, the error values are not as small as those in the case of temporal pattern detection, where the periodic structure is very regular and the error values are close to zero, as depicted in Figure 4b. Based on that, it can be concluded that a detected pattern in the frequency allocation is likely to have a pseudo-periodic structure, as shown in Figure 5b.

B. Modelling the Frequency Allocation

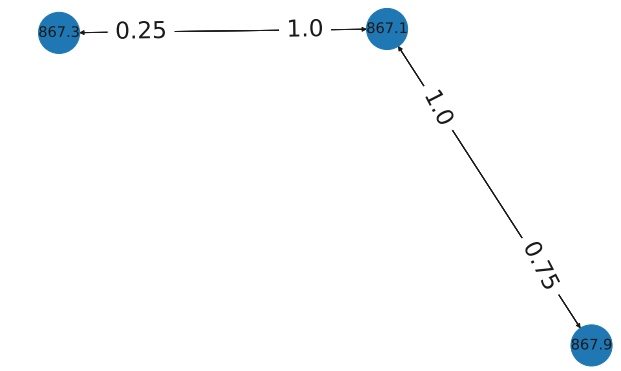
For jamming and security purposes, it is advantageous to create a model of a device's frequency allocation scheme, especially when the targeted device has a frequency allocation pattern. Subsequently in this paper, the stream of the assigned frequencies $\mathbf{f}$ is modelled as a Markov chain, which describes the potential sequence of events (frequencies) where the probability of each event depends only on the previous state achieved (frequency) [16], as illustrated in Figure 6. This enables the prediction of the next allocated frequency of a targeted device, which helps to optimize the accuracy of selective jamming [2], as the demonstrated examples in Figure 6.

V. Evaluation Results

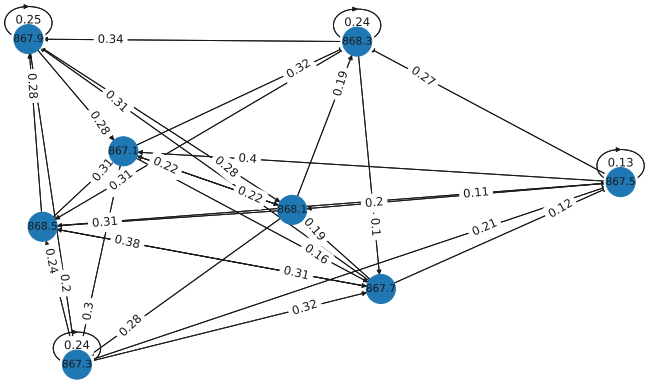
A. Analysis of Inter-arrival Times

To assess the practicality of using the demonstrated algorithms on real data, data obtained from the implemented packet sniffer is utilized as an example. The histogram in Figure 7, obtained using the proposed algorithm for inter-arrival time estimation, indicates that approximately 50% of the analyzed SCPs have only one inter-arrival time, While more than 40% of the SCPs indicates a device sending packets with two inter-arrival times. Then, this percentage decreases as the number of inter-arrival times increases.

Figure 8a shows that more than 20% of the mean values of the inter-arrival times are $\approx 1\,\mathrm{h}$, and roughly 10% of the inter-arrival times are about $\approx 10\,\mathrm{min}$, which are commonly used



(a) A Markov chain with 3 states.



(b) A Markov chain with 8 states.

Fig. 6: Modelling the frequency allocation as a Markov chain.

by humans. On the other hand, Figure 8b demonstrates that the inter-arrival times' standard deviations have concentrated distributions. Here, the most concentrated distribution is for the inter-arrival time of $\approx 1\,\mathrm{h}$, with a standard deviation of $\approx 3\,\mathrm{min}$.

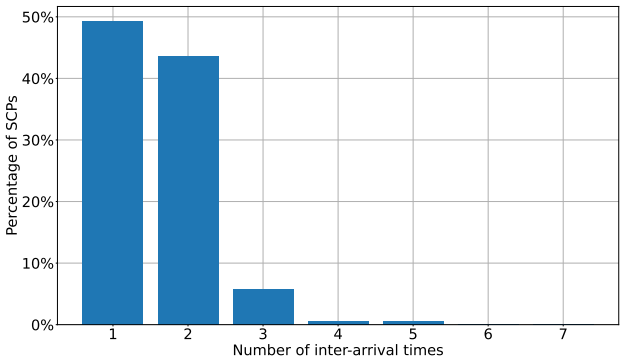
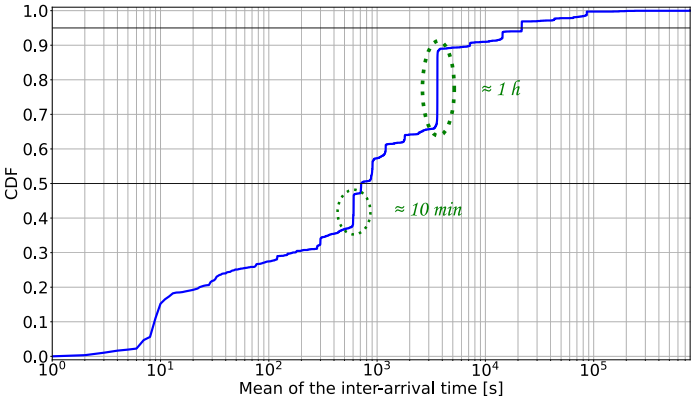


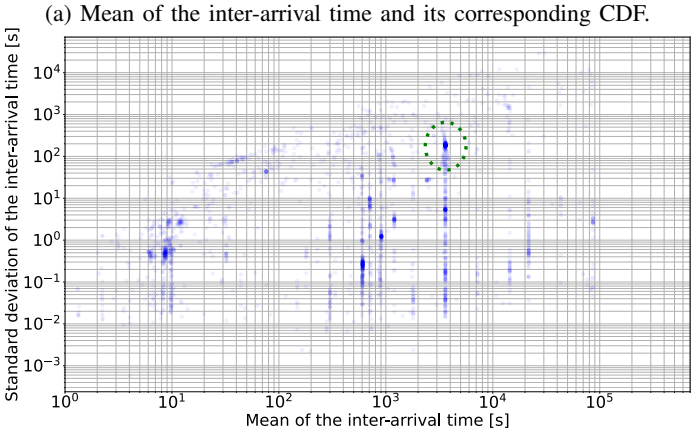
Fig. 7: Percentage of the SCPs versus the number of inter-arrival times.

B. Analysis of Temporal Patterns

Table II shows that most of the analyzed SCPs ($\approx 99.459\%$) don't have a temporal pattern in their inter-arrival times.



(a) Mean of the inter-arrival time and its corresponding CDF.



(b) Mean of the inter-arrival time versus its standard deviation.

Fig. 8: Statistical analysis of the acquired inter-arrival times.

Otherwise, the rest are having a temporal pattern with different period lengths.

TABLE II: Temporal pattern according to the period length

Period length	Count	Percentage
1	24805	99.459%
2	98	0.393%
3	7	0.028%
5	1	0.004%
6	1	0.004%
7	1	0.004%
8	1	0.004%
13	1	0.004%
24	1	0.004%
25	18	0.072%
26	1	0.004%
73	3	0.012%
145	2	0.008%
Total	**24940**	**100%**

C. Analysis of Frequency Allocation Patterns

Table III shows that most of the analyzed SCPs ($\approx 99.238\%$) don't have a pattern in their spectrum allocation algorithm. Otherwise, less than 1% are having a periodic structure with different period lengths.

TABLE III: Frequency pattern according to the period length

Period length	Count	Percentage
1	24750	99.238%
2	16	0.064%
4	84	0.337%
8	89	0.357%
16	1	0.004%
Total	**24940**	**100%**

D. Analysis of Frequency Allocation Model

To demonstrate the feasibility of using the proposed modelling of the frequency allocation scheme, the fairness of the transitions between the frequencies (states of the Markov chain) is evaluated using Pearson's chi-squared test [17]. While considering only the SCPs with a length of more than ten packets, the CDF vs the p-value is plotted, as shown in Figure 9. The results show that almost 55% of these SCPs have low p-value (≤ 0.05), indicating that these devices are biased in their spectrum allocation algorithm (null hypothesis is rejected). In other words, they are not fair in the sequence of the chosen frequencies and then their trajectory can be predicted by an eavesdropper, as the demonstrated example in Figure 6a whose p-value is ≈ 0.

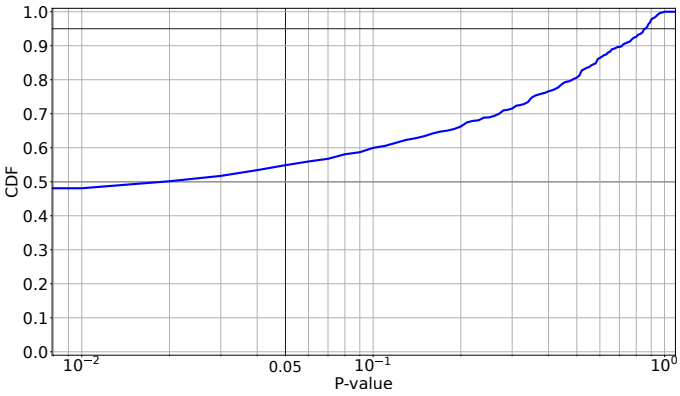


Fig. 9: P-value and its corresponding CDF.

VI. Conclusion

This paper proposes how to uncover crucial information about the nearby active LoRa devices using a passive packet sniffer. This proposed sniffer is made up of common tools and a commercial gateway, and it processes received LoRa packets in the city of Rennes, within two months. An estimator is proposed for estimating the prominent inter-arrival times between packets and then the distribution of these values is statistically analyzed. Moreover, the pattern of the inter-arrival times is detected and inspected, before the work is shifted toward exploring the spectrum allocation technique of the devices. Hence, the periodic structure in the stream of the selected frequencies is examined. Furthermore, a Markov chain model is proposed to lay out the trajectory of the sequence of used frequencies by each device. The feasibility of the proposed model is demonstrated while having most of the analyzed devices with significant biased transitions between the states (frequencies) of their Markov chain model.

For future work, the introduced packet sniffer can be used to monitor the nearby active devices and estimate their temporal information. Additionally, by using the projected path of a frequency allocation scheme for a targeted device, a jammer could anticipate and enhance the effectiveness of selective jamming. The proposed algorithms can also be used for Research & Development purposes to cluster the devices based on their transmission characteristics. Accordingly, an actual physical device can be identified based on its distinct features, then, the work can be extended to determine its location.

Acknowledgment

The authors wish to give recognition to CREACH LABS, and also express their gratitude to Region Bretagne for providing financial support through the SAD program. Moreover, the authors would like to thank Jean François Legendre from Gwagenn company for borrowing from him the gateway [18].

References

[1] A. Abdelghany, B. Uguen, C. Moy and D. Lemur, "Decentralized Adaptive Spectrum Learning in Wireless IoT Networks Based on Channel Quality Information," *IEEE Internet of Things Journal*, vol. 9, no. 20, pp. 19660-19669, 15 Oct.15, 2022.

[2] E. Aras, N. J. Small, G. S. Ramachandran, S. Delbruel, W. Joosen and D. Hughes, "Selective Jamming of LoRaWAN using Commodity Hardware," *14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, New York, USA, pp. 363–372, 2017.

[3] N. Blenn, and F. Kuipers, "LoRaWAN in the wild: Measurements from the things network," *arXiv preprint*, arXiv:1706.03086, 2017.

[4] P. Spadaccino, F.G. Crinó and F. Cuomo, "LoRaWAN Behaviour Analysis through Dataset Traffic Investigation," *Sensors 2022*, vol. 22, no. 7, p. 2470.

[5] K.N. Choi, H. Kolamunna, A. Uyanwatta, K. Thilakarathna, S. Seneviratne, R. Holz, M. Hassan and A.Y. Zomaya, "LoRadar: LoRa sensor network monitoring through passive packet sniffing," *ACM SIGCOMM Computer Communication Review*, pp. 10-24, 2020.

[6] P. Spadaccino, D. Garlisi, F. Cuomo, G. Pillon and P. Pisani, "Discovery privacy threats via device de-anonymization in LoRaWAN," *Computer Communications*, vol. 189, pp. 1-10, 2022.

[7] Addressing & Activation | The Things Network. [Online]. Available: https://www.thethingsnetwork.org/docs/lorawan/addressing/.

[8] A. Abdelghany, B. Uguen, C. Moy and J. Le Masson, "Aggregation of Contiguous Packets in an Actual LoRaWAN Passive Packet Sniffer," *IEEE 97th Vehicular Technology Conference: VTC2023-Spring*, Florence, Italy, 2023. [Online]. Available: https://hal.science/hal-04016140.

[9] Tektelic KONA Macro IoT Gateway. [Online]. Available: https://www.tektelic.com/uploads/Brochures/Kona%20Macro.pdf.

[10] ChirpStack open-source LoRaWAN Network Server. [Online]. Available: https://www.chirpstack.io/.

[11] MQTT - The Standard for IoT Messaging. [Online]. Available: https://mqtt.org/.

[12] Node-RED. [Online]. Available: https://nodered.org/.

[13] LoRaWAN packet decoder. [Online]. Available: https://github.com/anthonykirby/lora-packet.git.

[14] Measurement Data. [Online]. Available: https://gitlab.com/ahmednagy/lorawan-beaulieu-measurement-january-2023.git.

[15] B. W. Silverman, "Density Estimation for Statistics and Data Analysis," *Chapman and Hall*, London, 1986.

[16] J. Norris, "Markov Chains," *Cambridge University Press*, 1997.

[17] P. E. Greenwood and M. S. Nikulin, "A Guide to Chi-Squared Testing," *Wiley*, New York, USA, April 1996.

[18] Gwagenn company website. [Online]. Available: http://www.gwagenn.com/en/home/.