CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Technology independent targeted interference detection for wireless IoT
Authors	Morillo, Gabriela;Roedig, Utz;Stankovic, John
Publication date	2024-08-12
Original Citation	Morillo, G., Roedig, U. and Stankovic, J. (2024) 'Technology independent targeted interference detection for wireless IoT', 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Madrid, Spain, 8-11 July 2024, pp. 44-50. https://doi.org/10.1109/MeditCom61057.2024.10621167
Type of publication	Conference item
Link to publisher's version	10.1109/MeditCom61057.2024.10621167
Rights	© 2024, IEEE. For the purpose of Open Access, the author has applied a CC BY public copyright licence to this Author Accepted Manuscript version.
Download date	2026-11-03 16:07:00
Item downloaded from	https://hdl.handle.net/10468/16213

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Technology Independent Targeted Interference Detection for Wireless IoT

Gabriela Morillo
School of Computer Science and IT
University College Cork
Cork, Ireland
g.morillo@cs.ucc.ie

Utz Roedig
School of Computer Science and IT
University College Cork
Cork, Ireland
u.roedig@cs.ucc.ie

John Stankovic
Department of Computer Science
University of Virginia
Virginia, USA
jas9f@virginia.edu

***Abstract*—The expansion of Internet of Things (IoT) applications, particularly in critical systems, demands developing effective measures against targeted radio interference attacks. To remove such threats it is essential to identify attacks and to distinguish them from natural occurring interference. Furthermore, it is desirable to develop a communication technology-independent detection method as IoT deployments such as smart cities or factories are usually heterogeneous. This research introduces a technology-independent method for detecting targeted interference in IoT deployments capable of operating on resource-constrained IoT devices. Packet loss rates and packet loss patterns independent of the specific physical layer technology are analysed to determine if a targeted attack is present. The proposed approach is validated through comprehensive evaluations of two example technologies, Narrowband-Internet of Things (NB-IoT) and IEEE 802.15.4 Guarantee Time Slot (GTS), demonstrating its effectiveness in detecting attacks. Our evaluation shows that the detector can distinguish between targeted interference attacks and the impact of naturally occurring interference.**

I. Introduction

Many IoT applications are considered critical systems, and it is important to guarantee that such deployments are resilient to cyber-attacks. An attacker may use radio interference to disrupt communication while minimising the risk of detection. The attacker will selectively interfere with specific elements of the communication protocol in order to maximize disruption while minimising interference duration. Instead of a simple continuous and overpowering jamming signal an attacker will use a highly selective jamming signal with just sufficient power to cause disruption.

It is essential to identify such an attack in order to remove the threat. A particular challenge in this context is to distinguish naturally occurring interference from a deliberate attack. It is important to distinguish these two interference scenarios as they require different solutions. In case of natural interference, a response might be to step up transmission effort. In case of a targeted attack it might be better to stop transmission temporarily and put effort on identifying and removing the attacker. There is a need to devise methods for targeted interference detection.

IoT deployments may use a variety of wireless communication technologies. Depending on the specific use case a different wireless technologies may be beneficial. In some cases a number of wireless technologies might be used together in a heterogeneous setup. It is therefore desirable to devise a targeted interference detection that can be useful to a large set of wireless technologies.

Targeted interference detection may be executed at different points of a network and with the help of different tools and equipment. For example, it may be possible to execute interference detection centrally at a base station with specialist equipment. On the other hand, interference detection may be carried out on IoT devices only using information available from the existing communication transceiver. Depending on the approach, different cost and coverage will be associated with interference detection.

The presented work extends our previous effort to construct an interference detector for the NB-IoT protocol based on monitoring of loss rates [1]. We describe a novel method for targeted interference detection that can execute on resource-constrained IoT devices and can be used across different wireless technologies. The interference detection is independent from the lower-layer (physical and MAC) specifics of the wireless technology. The detector monitors loss rates and loss sequence of different protocol elements. A targeted interferer will cause loss rates and/or patterns to specific protocol elements while natural interference will impact on all protocol elements the same way. To perform targeted interference detection for a wireless protocol, protocol elements need to be defined, which are then observed technology-independent by the detector. In this work, we show how IoT nodes using two example technologies, NB-IoT and 802.15.4 can perform targeted interference detection. The specific contributions of our work are:

- *Technology Independent Targeted Interference Detector:* We provide a design of a targeted interference detector monitoring *loss rates* and/or *loss patterns* of protocol elements in a technology-independent way.
- *Example Detection Scenarios:* We show how the detector can be used in two scenarios: NB-IoT and 802.15.4 GTS. For both scenarios we describe example attacks and how monitoring is performed.
- *Evaluation:* We provide a comprehensive simulation evaluation of the detector for the two use cases NB-IoT and 802.15.4 GTS.

The remainder of the paper is organised as follows: in Section II, we describe the existing related work. In Section III, we provide a description of our threat model. In Section IV, we present the IoT case studies subjected to targeted jamming attacks. Section V presents the proposed Technology Independent Targeted Interference Detector. In Section VI, we present the experimental evaluation of the model and discuss the results. Section VIII concludes the paper.

II. Related Work

In an effort to characterise, model, and mitigate interference in wireless IoT communication networks, numerous studies have been undertaken. Jamming interference attacks have been the subject of a comparatively small number of studies, with even fewer works giving their attention to jamming attack detection. To the best of our knowledge, no prior work has proposed a technology independent jamming attack detection method. Thus, we focus in the next paragraphs on work in the IEEE 802.15.4 and Low Power Wide Area Network (LPWAN) space as we use these technologies as examples for our work.

For IEEE 802.15.4 MAC layer jamming attacks, extensive research works have been proposed [2]–[6]. Likewise, a range of studies have addressed the issue of detecting jamming in IEEE 802.15.4 networks [7]–[9].

Wood et al. [10] propose DEEJAM, a MAC-layer protocol for defeating energy-efficient jamming in networks based on IEEE 802.15.4. Combining four defensive mechanisms together, the protocol defeats the effectiveness of jamming from attackers with the same capabilities as other IEEE 802.15.4 network nodes. This work focuses on IEEE 802.15.4 jamming attacks and it differs from our research as it focuses on the techniques to reduce/eliminate most of the impact of jamming instead of detecting ongoing attacks.

Mahony et al. [11] propose an Intrusion Detection System (IDS) for ZigBee networks that is based on machine learning. The model formulates an interference detection focused entirely on analyzing received in-phase and quadrature-phase samples from the received WSN signal. In contrast to our research, this study employs a machine learning approach to detect interference which would be difficult to execute on resource constrained devices.

In a similar fashion, jamming interference in LPWANs has been the subject of numerous studies, including those on LoRa [12], [13], Sigfox [14] and NB-IoT [1], [15]–[17].

Ionescu et al. [17] describe energy depletion attacks on NB-IoT devices using malicious interference. The model considers jamming focused on the initial unprotected downstream communication, MIB-NB, SIB1-NB, or SIB2-NB information, and two different jamming approaches: message jamming and message injection. This work is complementary to ours as it describes attacks that we use in our evaluation. The work focuses on the attack but does not consider detection of interference based attacks.

III. Threat Model

We assume an attacker is equipped with an IoT device with similar capabilities to the other devices in the network. The attacker uses the device to emit a jamming interference signal, and it is able to monitor the communication between the transmitter node and the target node. The attacker knows the protocol and aims to use the jamming signal such that communication between the nodes is still possible to evade detection. The attacker attempts to use interference for the shortest duration possible to reduce energy usage (the attacker may operate on a battery-powered device) and to reduce the likelihood of detection as the attacker does not want to be found. Thus, the attacker will aim interference at specific protocol elements to achieve maximum impact with a minimum interference effort. We called that a targeted attack. We assume further that the attacker aims to disrupt the communication channels with the attack. The attacker performs a targeted attack by emitting a jamming signal to disrupt (parts) of a specific physical channel. As a result, the attacker will prevent the successful reception of specific Protocol Data Unit (PDU) at the end device. The attacker may not be able to determine the content of all protocol elements, as some are transmitted encrypted. However, not all channel elements use encryption (e.g. MIB in NB-IoT), and interference can also be applied to encrypted signal parts.

IV. IoT Case Studies and Jamming Attacks

A. Narrowband-Internet of Things (NB-IoT)

NB-IoT is a standard introduced by the Third Generation Partnership Project (3GPP), is based on Long-Term Evolution (LTE) and is designed for LPWAN.

For the specific case of NB-IoT, we assume that the attacker aims to disrupt the downlink channel with jamming interference. The attacker might monitor communication between User Equipment (UE) and Evolved Node B (eNodeB). The attacker can perform a targeted attack by submitting a jamming signal to disrupt parts of a specific physical channel, e.g. Narrowband Broadcast Channel (NPBCH) [1], Narrowband Primary Synchronisation Signal (NPSS) and Narrowband Secondary Synchronisation Signal (NSSS) [15], System Information Block (SIB) [16]. The attacker may only target NPBCH to perform a battery depletion attack or Narrowband Physical Downlink Shared Channel (NPDSCH) to disrupt data transmission between a UE and eNodeB. The attacker prevents the successful reception of specific subframes at the UE.

In this study we consider two selective jamming interference attacks (referred to as *Constant MIB* and *Selective MIB* attack) on the downstream NB-IoT channel based on attacks described in [16]. Both attacks target the NPBCH transporting the Master Information Block (MIB) in subframe 0. The MIB is transported in eight Code Sub-blocks (CSB) which are each repeated eight times. The *Constant MIB* attack assumes that the interferer targets subframe 0 (of 10 subframes present in the downlink channel). The *Selective MIB* attack assumes that only 16 consecutive CSB are subject to interference as this is the minimum level of interference required to prevent successful decoding of the MIB. Figure 1 illustrates the Transmission of the NPBCH-CSB with a Transmission Time

Interval (TTI) of 640ms and the Selective MIB targeted attack to only 16 consecutive subframes.

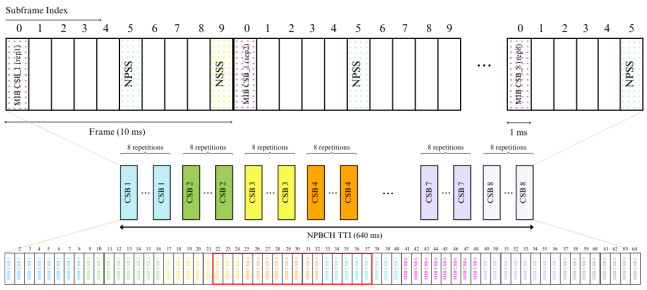


Fig. 1: NB-IoT Selective MIB Attack.

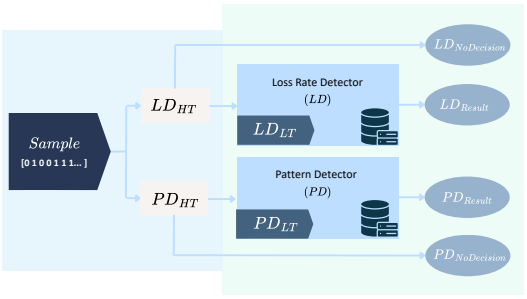


Fig. 3: Detector Design: The sequence S_n is received every T seconds describing PDU loss/reception. The detector decides if a targeted attack is present.

B. IEEE 802.15.4

The IEEE 802.15.4 MAC [18] defines two fundamental MAC modes: (i) non-beacon enabled and (ii) beacon-enabled.

The beacon-enabled approach adopts a slotted Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). Here, the network coordinator periodically broadcasts beacon frames to announce the presence of the network and provide timing and control information. These beacons divide the channel into superframes of the same direction. The superframe structure is depicted in Figure 2.

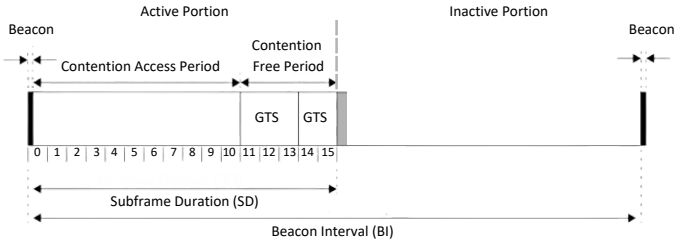


Fig. 2: IEEE 802.15.4 Superframe Structure.

A superframe is divided into two periods: a mandatory active period and an optional inactive period. The active period includes Contention Access Period (CAP) and Contention-Free Period (CFP). CAP is a part of the superframe during which devices can transmit data using a contention-based access method. In this period, devices that need to communicate attempt to transmit their data using a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism. CSMA/CA helps avoid collisions by checking the channel's availability before initiating data transmissions. The CFP period is a portion of the superframe reserved for contention-free communication. During this period, specific devices, typically associated with the network coordinator, can transmit data without the need for contention. The coordinator allocates time slots to individual devices, allowing them to transmit data in a coordinated and collision-free manner. The active period is equally divided into 16 slots.

In this study we consider an attack scenario using the IEEE 802.15.4 GTS in which an adversary device synchronises with the Personal Area Network (PAN) coordinator through the reception of the beacon messages. The beacon frame contains the GTS descriptor, which indicates the GTS starting slot, length, direction, and associated device address. The adversary node can learn the GTS times from the coordinator by extracting the GTS descriptor from the beacon. Once the attacker obtains the allocated GTS times, it can create jamming interference in any of these dedicated slots, causing a denial of service as these slots are assumed to provide collision-free communication [19], [2]. We refer to this attack as *Constant GTS*.

V. Technology Independent Targeted Interference Detection

We monitor on the device the reception success of Protocol Data Units (PDUs) over time. A PDU is a protocol element received periodically by the IoT device (e.g. a beacon, a data frame, a routing message). Not all possible PDUs need to be tracked for targeted interference monitoring. The assumption is that a device is able to tell if a PDU has been lost or has been received successfully.

For targeted interference detection, the sequence of reception success of the monitored PDUs is analysed. This task can be performed independent of the specifics of a wireless protocol and device once PDU monitoring specifics are defined.

The assumption is that an interferer will target a specific PDU to achieve his goal. Such an attack can be distinguished from naturally occurring interference. Targeted interference will cause an increased loss rate for one specific PDU. In addition, loss patterns for a PDU will exhibit periodicity in case of a targeted attack while natural interference will cause a random loss patterns.

The targeted interference detector is split into a small adapter interface layer and a detector. The adapter is specific to the wireless protocol and device considered, while the detector functions are protocol and device independent, and is therefore reusable in different IoT settings.

A. Adapter

The adapter requires the ability to monitor the reception success of η PDUs over a time period T. The adaptor may use a configuration syntax to specify what constitutes each PDU (e.g. mapping of message type or slot number to PDU). During T the adapter collects η sequences S_n of 0s and 1s denoting PDU_η reception failure or success. In a TDMA protocol where a PDU is mapped to a specific slot, a sequence S_n has implicit timing information; it is known when a reception succeeded or failed. In a contention-based protocol, it is not possible to reconstruct from the sequence S_n when each reception failure or success occurred within T. Every period T, the sequences S_n are passed to the detector.

B. Detector

The detector (shown in Figure 3) receives the sequences S_n every T seconds and must make a decision if the device is exposed to a targeted interference attack or not. The detector consists of a *Loss Rate Detector* and a *Loss Pattern Detector*. Each detector may be used on their own or the outputs are fused to obtain an overall decision.

A High Threshold (HT) (named LD_{HT} and PD_{HT} for the two detectors) is implemented to filter the sequences S_n to exclude sequences that exhibit a very high packet loss. In such cases it is impossible for any detector to make a decision; it cannot be decided if a targeted attack exists but it is obvious that the entire communication link is non-functional.

A Low Threshold (LT) (named LD_{LT} and PD_{LT} for the two detectors) is implemented within the detectors decision logic. If the observed losses are below this threshold an attack on the specific PDU will be inefficient and we count an ineffective attack as a non-attack case.

It has to be noted that multiple instances of *Loss Rate Detector* and a *Loss Pattern Detector* may be executed for each PDU to be monitored on an IoT device.

1) Loss Rate Detector (LD): The Loss Rate Detector evaluates the packet loss rate of one PDU in relation to the loss rates of all other PDUs. For each monitored PDU_η ($\forall \eta \in 0, 1, 2, 3, 4, ..., \eta$), we calculate the loss rate L_η. As a result, a set of loss rates L containing η values is collected. L indicates the interference environment to which the device is exposed, assuming that it is caused by either a targeted attack or natural interference. Under a targeted attack, the investigated PDU will exhibit a higher loss rate than other PDUs. As PDUs may differ in packet size or may use different coding schemes, it may be required to adjust loss rates accordingly, resulting in the adjusted loss rate set L'. A targeted attack is present if the adjusted loss rate in the investigated PDU L'_η is above the average loss rate $\widetilde{L}'$. To mitigate false alarms, we define φ as a safety margin to ensure that an attack is detected only if L'_η exhibits a significant deviation from the average $\widetilde{L}'$.

2) Loss Pattern Detector (PD): The Loss Pattern Detector analyses the loss pattern of a single PDU. For the specific monitored PDU_η ($\forall \eta \in 0, 1, 2, 3, 4, ..., \eta$), the sequence S_n is evaluated. Thus, a set of loss sequences S_n is collected per a specific PDU. The loss sequence S_n is processed using the autocorrelation with autocovariance coefficients ($corrcoef$) and then passing the result through an FFT function. If clear peaks are visible in the result a pattern exists. Under a targeted attack, the pattern detector will be able to determine if periodicity of losses is present; for natural interference no clear pattern will be visible. A targeted attack is present if a significant peak in the autocovariance function is identified. To mitigate false alarms, we define γ as a safety margin determining the significance of a present peak.

The Fused Detector integrates the outputs of the LD and the PD detectors. This integration involves processing the individual outcomes through two logical operations considering the union and the intersection of the results. The final detection decision is flagged by the Fused Detector. This dual-decision approach facilitates a comprehensive analysis of the detection effectiveness particularly when the nature of the network attack is unknown.

VI. Evaluation Setup

A. Simulation Environment and Scenarios

All simulations were carried out using MATLAB; we used the LTE Toolbox for NB-IoT and the Communications Toolbox Library for IEEE 802.15.4.

1) Narrowband-Internet of Things (NB-IoT): We simulate downlink communication between a base station (eNodeB) and a device (UE). The UE should determine if it is under attack using the detector. The complete transmitter and receiver signal processing chains of an NB-IoT system is simulated. The receiver chain completes the synchronisation, demodulation, and decoding of the NB-IoT downlink signal transmitted by the eNodeB. In our detector configuration, the adaption layer defines the NB-IoT downstream subframes as the PDUs of interest. The downlink signal may be subjected to interference, and it is possible to determine the accurate reception of subframes; 1 if decoding was successful, and 0 if unsuccessful, creating our sequences S_n.

2) IEEE 802.15.4: We simulate a PAN coordinator and a node receiving transmissions from the PAN coordinator via a GTS slot. The node should determine if it is under attack using the detector. We simulate an IEEE 802.15.4 beacon-enable MAC with GTS transmission. In our model, the node synchronises with the PAN Coordinator through the successful reception of a beacon. After that, the coordinator can communicate with the device using a GTS slot.

In our detector setup, the adaption layer defines the beacon messages and the specific GTS slot as two PDUs of interest. Receiving a message (either beacon or GTS transmission) is recorded a 1, a loss is recorded as 0 creating our sequences S_n for the two PDUs.

B. Noise and Attacks

The communication between the transmitter device (eNodeB or PAN coordinator) may be susceptible to interference. Three types of interference are considered.

- *Background Noise (BN):* refers to the presence of Additive White Gaussian Noise (AWGN) that remains constant while the transmission channel is active.
- *Background Traffic (BT):* refers to the presence of noise (interference) limited to specific time intervals. The duration of these noise periods and the intervals between them adhere to a Poisson process. This noise interference replicates the impact of interference caused by an additional communication network that is implemented in the currently analysed network space.
- *Targeted Attack (TA):* Interference signal that is present only in specific Protocol Data Unit (PDU)s, simulating a targeted attack on particular protocol elements.

Each interference type can be present at different Signal to Noise Radio (SNR) levels. The coexistence of different types of interference is possible.

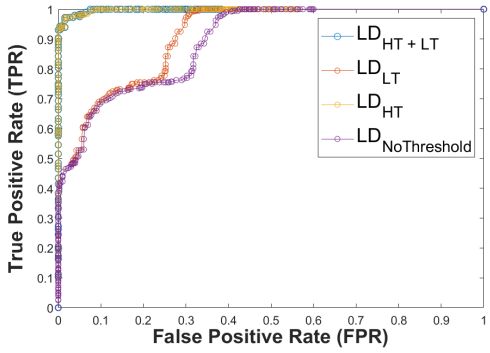
VII. Results

In these experiments, we create several scenarios with and without targeted interference attacks; we test the constant jamming attack for NB-IoT and IEEE 802.15.4 GTS, as well as the selective jamming attack and a third scenario that combines constant and selective approaches to simulate a more realistic attack when the node does not know the nature of the attack for NB-IoT. Then, we evaluate our loss rate and pattern detector's detection capacity by determining whether it can detect an attack is happening (True Positive) or if it incorrectly asserts an attack is present (False Positive).

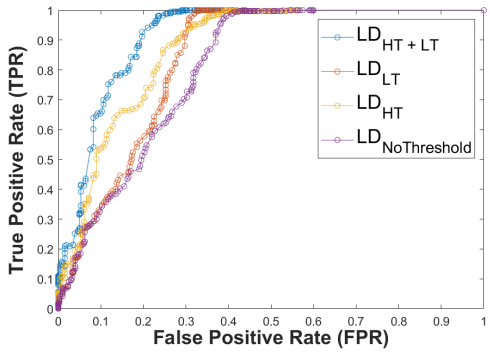
Considering that in a realistic scenario, even when a targeted attack is present, some background noise may also be present, we combine background noise (BN) and targeted attack (TA) noise to accomplish this. We create interference signals combining BN and TA at different SNR levels.

For NB-IoT, we generate 3 experiment sets with 800 transmission sequences each. i) 400 sequences in each set are attacks (400 Constant MIB, 400 Selective MIB and 200 Constant MIB and 200 Selective MIB in the third set) with TA-based SNR ranging from -9.75dB to 0B (steps of 0.25dB) and BN-based SNR ranging from -16dB to -10dB (steps of 2dB). This SNR setup guarantees that we are considering successful attacks. ii) 400 sequences in each set have no attack. We use here background noise (BN) based SNR ranging from 0dB to -4.75dB (steps of 0.25dB) and background traffic (BT) based SNR ranging from -5dB to -9.75dB(steps of 0.25dB).

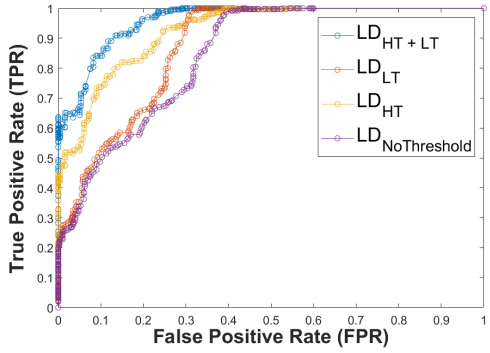
For IEEE 802.15.4, we generate one set with 560 sequences, distributed as follows: i) 280 attacks with TA-based SNR ranging from 0.0125dB to 0.510dB (steps of 0.0125dB), and BN-based SNR ranging from 0.26dB to 0.45dB (steps of 0.01dB). ii) 280 sequences in which no attack is present. We use background noise (BN) based SNR ranging from 0.275dB to 0.525dB (steps of 0.0125dB) and background traffic (BT) based SNR ranging from 0.0125dB to 0.250dB(steps of 0.0125dB).



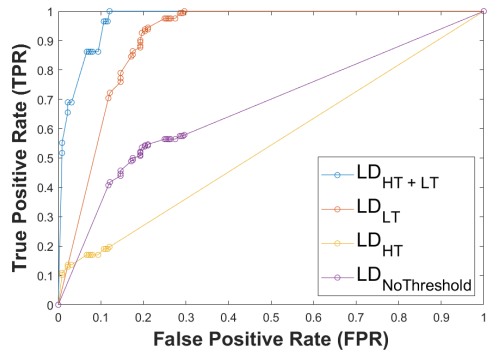
(a) Constant MIB



(b) Selective MIB

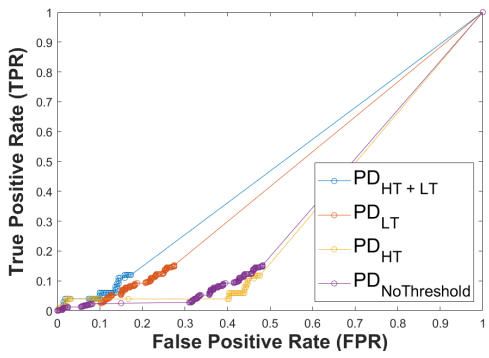


(c) Combined MIB

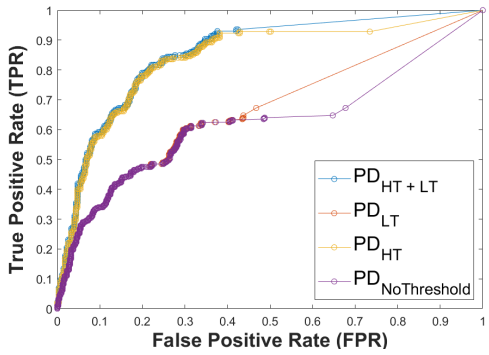


(d) Constant GTS

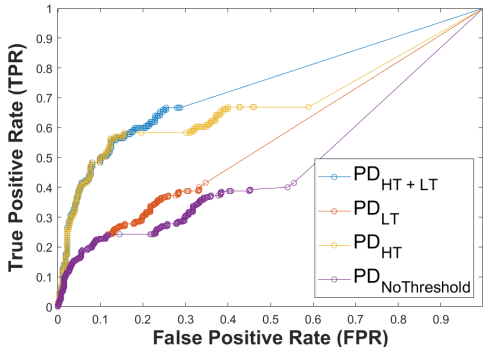
Fig. 4: Receiver Operating Characteristic (ROC) curves for NB-IoT and IEEE 802.15.4 GTS Jamming Attacks using Loss Rate Detector with different HT and LT.



(a) Constant MIB



(b) Selective MIB



(c) Combined

Fig. 5: Receiver Operating Characteristic (ROC) curves for NB-IoT Jamming Attacks using Pattern Detector with different HT and LT.

A. Loss Rate Detector

Figure 4 shows the result of our experiments in the form of a Receiver Operating Characteristic (ROC) curves for Narrowband-Internet of Things (NB-IoT) and IEEE 802.15.4 GTS. We use the Area Under the Curve (AUC) as one of the metrics to evaluate the detector design capabilities.

Figures 4a and 4b illustrate the ROC curves for the Loss Rate Detector (LD) applied to NB-IoT MIB Constant and Selective attack while Figure 4c shows the result if both attacks are present. Using our experimental data, we determined the parameters $LT = 12.5$ and $HT = 24$ provide a good fit, and we used these settings for our LD Detector in NB-IoT. For the Constant jamming attack, it can be observed that the LD_{HT+LT} setup performs better than the other setups. The detector can identify 93.04% of targeted interference cases with no false positives (an ideal setting as in a practical deployment false alarms should be avoided). For this scenario we obtain AUC $= 99.79\%$).

For the Selective and Combined jamming attack, it is shown that the LD_{HT+LT} detector performs better than the ones that consider only one of the thresholds (LD_{HT} or LD_{LT}) or no threshold at all. The detector identifies 11.11% and 63.64% of targeted interference cases with no false positives, respectively. We obtain an AUC $= 91.32\%$ for the Selective attack and the AUC $= 96.01\%$ for the Combined attack.

Figure 4d shows the Loss Rate Detector (LD) performance detection for IEEE 802.15.4 GTS attack. Using our experimental data, we determined the parameters $LT = 7.5$ and $HT = 55$ provide a good fit. It can be seen that the LD_{HT+LT} detector performs better than the ones that consider one of the thresholds or does not threshold at all. For this optimal scenario, the detector is able to identify 55.17% of targeted interference cases with no false positives, with an AUC=97.44%.

B. Pattern Detector

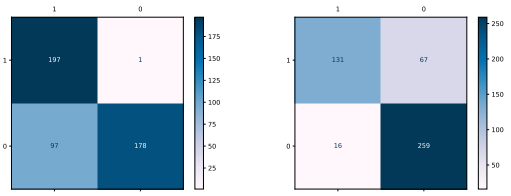
Figures 5a, 5b and 5c show the ROC curves for the Pattern Detector (PD) for NB-IoT MIB Constant, Selective and Combined targeted interference attack, respectively. Using our experimental data, we determined the parameters $LT = 7.5$ and $HT = 75$ provide a good fit. As expected, the detector underperforms against constant jamming attacks as it is unable to identify any pattern in the attack sequences. In this attack all MIB frames are destroyed resulting in a sequence S_n of (mostly) 1. A good performance is observed for the Selective attack. In this scenario, when the detector applies both thresholds (PD_{HT+LT}), it results in an $AUC = 90.18\%$. For the combined attack, an acceptable performance is also observed with an $AUC = 95.41\%$.

C. Fused Detector

Figure 6 shows the Confusion Matrix obtained for the optimal LD_ϕ and PD_γ for the Combined MIB targeted jamming attack. The Equal Error Rate values for the LD and PD are EER=11.72% and EER=28.l78%, respectively. As can be seen, the confusion matrix for the union of the results of the two detectors presents a high rate of true positives and false negatives and a low rate of true negatives, which will translate into reducing false alarms. Meanwhile, a higher level of true negatives is observed for the intersection operation.

VIII. Conclusion

We have demonstrated that it is possible to construct a technology-independent statistical anomaly detector capable of detecting targeted interference on IoT devices by utilising the available data in the node as an indirect measure of interference. We have shown that our detectors have excellent performance in detecting known-characteristic attacks $AUC = 99.79\%$ and $AUC = 97.44\%$ for LR detector with Constant jamming attack in NB-IoT and IEEE 802.15.4



(a) OR: $\phi \cup \gamma$ (b) AND: $\phi \cap \gamma$

Fig. 6: Confusion Matrix.

GTS, respectively; and $AUC = 90.18\%$ for PD detector with Selective jamming attack in NB-IoT. Also, we have shown that our detectors have adequate performance in detecting interference when the UE is unaware of the nature of the jamming attack; the Loss Rate detector has a AUC $= 96.01\%$ and can detect 63.64% of attacks without false positives, while the Pattern Detector has an AUC $= 95.41\%$. Finally, we have demonstrated that the threshold HT has more impact than LT.

ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant numbers 18/CRT/6222 and 13/RC/2077_P2 and the U.S. NSF NRT Grant 1829004. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

REFERENCES

[1] G. Morillo, U. Roedig, and D. Pesch, "Detecting targeted interference in NB-IoT," in *2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*. Institute of Electrical and Electronics Engineers IEEE, 2023, pp. 475–482.

[2] R. Sokullu, O. Dagdeviren, and I. Korkmaz, "On the IEEE 802.15.4 MAC Layer Attacks: GTS Attack," in *2008 Second International Conference on Sensor Technologies and Applications*, 2008, pp. 673–678.

[3] M. Achour, M. Mana, and A. Rachedi, "On the issues of selective jamming in IEEE 802.15.4-based wireless body area networks," *Peer-to-Peer Networking and Applications*, vol. 14, no. 1, pp. 135–150, 2021.

[4] Y. M. Amin and A. T. Abdel-Hamid, "Classification and analysis of IEEE 802.15.4 MAC layer attacks," in *2015 11th International Conference on Innovations in Information Technology (IIT)*, 2015, pp. 74–79.

[5] S. S. Jung, M. Valero, A. Bourgeois, and R. Beyah, "Attacking Beacon-Enabled 802.15.4 Networks," in *Security and Privacy in Communication Networks*, S. Jajodia and J. Zhou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 253–271.

[6] S. M. Sajjad and M. Yousaf, "Security analysis of IEEE 802.15.4 MAC in the context of Internet of Things (IoT)," in *2014 Conference on Information Assurance and Cyber Security (CIACS)*, 2014, pp. 9–14.

[7] H. Pirayesh, P. Kheirkhah Sangdeh, and H. Zeng, "Securing ZigBee Communications Against Constant Jamming Attack Using Neural Network," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4957–4968, 2021.

[8] D. Han, A. Li, L. Zhang, Y. Zhang, J. Li, T. Li, T. Zhu, and Y. Zhang, "Deep learning-guided jamming for cross-technology wireless networks: Attack and defense," *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 1922–1932, 2021.

[9] R. Daidone, G. Dini, and M. Tiloca, "A solution to the GTS-based selective jamming attack on IEEE 802.15.4 networks)," in *2014 Conference on Information Assurance and Cyber Security (CIACS)*, 2014.

[10] A. D. Wood, J. A. Stankovic, and G. Zhou, "DEEJAM: Defeating Energy-Efficient Jamming in IEEE 802.15.4-based Wireless Networks," in *2007 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2007, pp. 60–69.

[11] G. D. O'Mahony, P. J. Harris, and C. C. Murphy, "Detecting Interference in Wireless Sensor Network Received Samples: A Machine Learning Approach," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, 2020, pp. 1–6.

[12] C.-Y. Huang, C.-W. Lin, R.-G. Cheng, S. J. Yang, and S.-T. Sheu, "Experimental Evaluation of Jamming Threat in LoRaWAN," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019)*, 2019, pp. 1–6.

[13] I. Martinez, P. Tanguy, and F. Nouvel, "On the performance evaluation of LoRaWAN under Jamming," in *2019 12th IFIP Wireless and Mobile Networking Conference (WMNC)*, 2019, pp. 141–145.

[14] Y. Wang, S. Jere, S. Banerjee, L. Liu, S. Shetty, and S. Dayekh, "Anonymous Jamming Detection in 5G with Bayesian Network Model Based Inference Analysis," in *2022 IEEE 23rd International Conference on High Performance Switching and Routing*, 2022, pp. 151–156.

[15] G. Morillo and U. Roedig, "Jamming of nb-iot synchronisation signals," in *Computer Security. 26th European Symposium on Research in Computer Security ESORICS 2021 International, Darmstadt, Germany, October 4–8, 2021*. Springer Int. Publishing, 2021, pp. 759–763.

[16] V. Ionescu and U. Roedig, "Battery depletion attacks on nb-iot devices using interference," in *ADIoT ESORICS 2021*, 2021, p. 276–295.

[17] V. Ionesscu and U. Roedig, "NB-IoT Battery Depletion via Malicious Interference," in *Proceedings of the 2022 International Conference on Embedded Wireless Systems and Networks EWSN*, 2022, p. 244–249.

[18] IEEE, "IEEE Standard for Low-Rate Wireless Networks," *IEEE Std 802.15.4-2020*, pp. 1–800, 2020.

[19] R. Sokullu, I. Korkmaz, O. Dagdeviren, A. Mitseva, and N. R. Prasad, "An investigation on IEEE 802.15. 4 MAC layer attacks," in *Proc. of WPMC*, vol. 41, 2007, pp. 42–92.