Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Pocketpen: A Raspberry Pi Based Multi-Functional Security Testing and Iot Interaction Device Inspired By Flipper Zero

Abhinav K, Anfas Panginikadan, Anurag Ashok, Viswanth T, Roshmin Rizuvan

Department of Computer Science and Engineering (Cyber Security), Royal College of Engineering & Technology, Akkikavu, Kerala, India

ARTICLE INFO

Keywords:
Wireless security testing
Internet of Things
Flipper Zero
Raspberry Pi
RF fingerprinting
NFC/RFID security
Bluetooth Low Energy
TinyML
Edge anomaly detection

ABSTRACT

Smart locks, access badges, remote controls, wearables and Internet of Things (IoT) sensors rely on a mix of radio and optical links, namely Sub-GHz radio, infrared, Near Field Communication (NFC), Radio Frequency Identification (RFID), Wi-Fi and Bluetooth Low Energy (BLE), and every link type has documented weaknesses. Portable multi-protocol tools such as Flipper Zero have made interaction with these interfaces practical, but they leave the analysis of captured activity to the operator, while machine learning (ML) based detection is usually studied for one protocol or one dataset and on powerful hardware. This review examines the research needed for portable, analysis-capable security testing and IoT interaction devices. Fifty-one studies, seeded by fifteen recent works and verified against their bibliographic records, are organised into five streams: portable multi-protocol platforms, protocol-specific security, RF fingerprinting and physical-layer authentication, IoT intrusion detection datasets and ML-based detection, and lightweight edge ML. The comparison shows that each stream is maturing: RF fingerprinting identifies counterfeit NFC and RFID tags and forged transmitters, new datasets combine IP, Z-Wave and MQTT data, and quantised detectors reach F1-scores above 0.97 within tens of kilobytes on microcontrollers. However, no reviewed study combines multi-protocol acquisition with on-device ML analysis, public datasets do not cover the radio protocols handled by portable devices, and edge evaluations are limited to network-flow features. The paper consolidates these findings into research gaps, open challenges and future directions, including a common observation schema, authorised multi-protocol datasets, hybrid physical- and protocol-layer detection, end-to-end edge benchmarking and responsible open platforms.

1. INTRODUCTION

The number of devices that communicate wirelessly in homes, offices, vehicles and industrial sites has grown faster than the tools available to examine them. A single smart building may contain sub-gigahertz (Sub-GHz) alarm sensors and garage remotes, infrared (IR) controlled appliances, Near Field Communication (NFC) and Radio Frequency Identification (RFID) access badges, Wi-Fi cameras and Bluetooth Low Energy (BLE) wearables and trackers. Each of these technologies has its own physical layer, frame structure and trust model, and each has a documented history of weaknesses, ranging from replayable fixed codes and broken proprietary ciphers to unauthenticated management frames and privacy leakage through advertising packets [1–4]. Large-scale incidents such as the Mirai botnet have shown that weaknesses in individually inexpensive devices can be aggregated into damage of national significance [5], which explains the sustained research interest in Internet of Things (IoT) security and in machine learning (ML) based detection of IoT attacks [6].

Despite this interest, the day-to-day practice of wireless security assessment remains fragmented. A practitioner who must evaluate a Sub-GHz remote, an NFC badge and a BLE device typically uses three different instruments, each with its own capture format, and then interprets the captured activity largely by hand. Portable multi-protocol devices have reduced the hardware side of this burden considerably. Flipper Zero, introduced as a pocket-sized tool combining a Sub-GHz transceiver, 13.56 MHz NFC and 125 kHz RFID front-ends, an IR transceiver, BLE and general-purpose input/output (GPIO) pins on a single microcontroller platform [7], has attracted both enthusiasm from security researchers and concern from defenders [8]. Recent academic evaluations confirm that such a device can reproduce realistic attacks on everyday systems, including credential emulation, remote-control interference and Wi-Fi deauthentication [9–11]. However, these devices are designed as interactive tools: the analysis of what they capture, and the decision about whether an observation is benign or suspicious, is still left to the operator.

In parallel, three largely independent bodies of research have matured. The first studies the security of individual wireless protocols, for example BLE trackers [12,13], NFC-enabled IoT [14] or chipless RFID tags [15]. The second applies ML and deep learning (DL) to IoT intrusion detection,

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

supported by an expanding family of benchmark datasets [16–20]. The third, often referred to as tiny machine learning (TinyML) or edge ML, investigates how such detectors can be compressed and executed on resource-constrained hardware [21–24]. A separate stream of work on radio frequency (RF) fingerprinting uses hardware-induced signal imperfections to identify or authenticate individual transmitters [25–27]. Each stream contributes a component that a portable, analysis-capable security testing device would require, yet the streams are seldom combined, and the trade-offs between them are rarely discussed in one place.

This survey addresses that gap. Unlike earlier reviews that focus on a single protocol [28,29], a single technique [25] or ML-based IoT security in general [6], it examines the literature from the perspective of a portable, open, multi-protocol security testing and IoT interaction device that performs lightweight anomaly detection at the edge. The specific contributions are as follows:

- A structured review of portable multi-protocol security testing platforms, protocol-specific security studies, RF fingerprinting, IoT intrusion detection datasets and lightweight edge ML, with comparative tables for each stream.
- A unified technical background covering the attack surface of Sub-GHz, IR, NFC/RFID, Wi-Fi and BLE interfaces, the ML model families used for wireless anomaly detection, and the metrics used to evaluate both detection quality and resource consumption.
- A critical cross-stream comparison that identifies which capabilities, namely multi-protocol acquisition, signal-level features, on-device inference, model compression and open extensibility, are addressed by existing work and which remain isolated.
- A consolidated set of research gaps, open challenges and future research directions for analysis-capable portable wireless security platforms.

The remainder of the paper is organised as follows. Section 2 describes the review methodology. Section 3 introduces the background concepts. Section 4 surveys the existing research in five thematic streams. Section 5 presents the comparative analysis. Section 6 discusses research gaps, challenges and future directions, and Section 7 concludes the paper. Table 1 lists the abbreviations used throughout the manuscript.

Table 1
Abbreviations used in this review.

Abbreviation	Description	Abbreviation	Description
AWID	Aegean Wi-Fi Intrusion Dataset	MCU	Microcontroller Unit
BLE	Bluetooth Low Energy	MIMO	Multiple-Input Multiple-Output
CNN	Convolutional Neural Network	ML	Machine Learning
CR	Compression Ratio	MLP	Multi-Layer Perceptron
DL	Deep Learning	MQTT	Message Queuing Telemetry Transport
DOI	Digital Object Identifier	NFC	Near Field Communication
DT	Decision Tree	PLA	Physical-Layer Authentication
FN / FP	False Negative / False Positive	PMF	Protected Management Frames (IEEE 802.11w)
FNN	Feed-forward Neural Network	PTQ / QAT	Post-Training Quantisation / Quantisation-Aware Training
FPR	False Positive Rate	RF	Radio Frequency; Random Forest (context)
GPIO	General-Purpose Input/Output	RFF	Radio Frequency Fingerprinting
GRU	Gated Recurrent Unit	RFID	Radio Frequency Identification
HPO	Hyperparameter Optimisation	RKE	Remote Keyless Entry
IDS	Intrusion Detection System	RNN	Recurrent Neural Network
IoT	Internet of Things	SBC	Single-Board Computer
IQ	In-phase and Quadrature (samples)	SDR	Software-Defined Radio
IR	Infrared	SNR	Signal-to-Noise Ratio
KNN	k-Nearest Neighbours	TCN	Temporal Convolutional Network
LightGBM	Light Gradient Boosting Machine	TFLM	TensorFlow Lite Micro
LSTM	Long Short-Term Memory	TinyML	Tiny Machine Learning
MACC	Multiply–Accumulate Operation	XGBoost	Extreme Gradient Boosting

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

2. METHODOLOGY

2.1. Research questions

The review was designed to answer five research questions (RQs) that follow the functional layers of an analysis-capable portable wireless security device:

- RQ1: Which portable platforms are used for multi-protocol wireless security testing, and to what extent do they support automated analysis of the activity they capture?
- RQ2: Which weaknesses and threat models characterise the Sub-GHz, IR, NFC/RFID, Wi-Fi and BLE interfaces that such platforms interact with?
- RQ3: Which ML approaches, feature representations and datasets have been used to detect anomalous activity or to identify devices in these settings, and how have they been evaluated?
- RQ4: What evidence exists that such models can operate within the memory, latency and energy budgets of portable hardware?
- RQ5: Which gaps prevent these components from being integrated, and which research directions could close them?

RQ1 is addressed in Sections 3.2 and 4.1, RQ2 in Sections 3.1 and 4.2, RQ3 in Sections 3.3, 4.3 and 4.4, RQ4 in Section 4.5, and RQ5 in Sections 5 and 6.

2.2. Search and selection protocol

The review followed a structured identification, screening and synthesis protocol, summarised in Fig. 1. The starting point was a seed set of fifteen recent studies (2025–2026) that address the core technical concerns of the target device class: multi-source IoT intrusion datasets, BLE tracker security, NFC and RFID authentication, RF fingerprinting and lightweight ML for constrained hardware. From this seed set, the literature was expanded in two ways. First, keyword searches were run against the Crossref and OpenAlex scholarly indexes and the Semantic Scholar database using combinations of device-level terms (“Flipper Zero”, “portable”, “Raspberry Pi”, “penetration testing”), protocol terms (“Sub-GHz”, “remote keyless entry”, “NFC”, “RFID”, “BLE”, “802.11”, “infrared”) and analysis terms (“RF fingerprinting”, “physical-layer authentication”, “intrusion detection”, “anomaly detection”, “TinyML”, “quantization”). Second, backward citation chasing was performed on the seed studies and on recent surveys in order to include foundational works on which the newer studies build, such as the original datasets, attack papers and quantisation methods.

Studies were retained when they satisfied three criteria: (i) they were published in a peer-reviewed journal or conference, or are widely used preprints of systems that the community relies on (for example, embedded inference runtimes and benchmarks); (ii) they report a concrete method, dataset, attack, measurement or deployment rather than an opinion; and (iii) they are directly relevant to at least one of the functional layers of an analysis-capable portable wireless security device. Studies whose bibliographic record could not be confirmed against its digital object identifier (DOI) or an equivalent authoritative listing were excluded. Two vendor documentation sources were retained solely to report hardware specifications and are not counted as studies. The retained studies were then grouped into the five thematic clusters shown in Fig. 1, and each cluster was synthesised with respect to its contribution, methodology, evaluation setting and limitations. Particular attention was paid to whether reported results were obtained on real radio captures or on network-level traffic datasets, and whether models were evaluated on the hardware on which they would actually run, since both factors strongly affect how far a result transfers to a portable device.

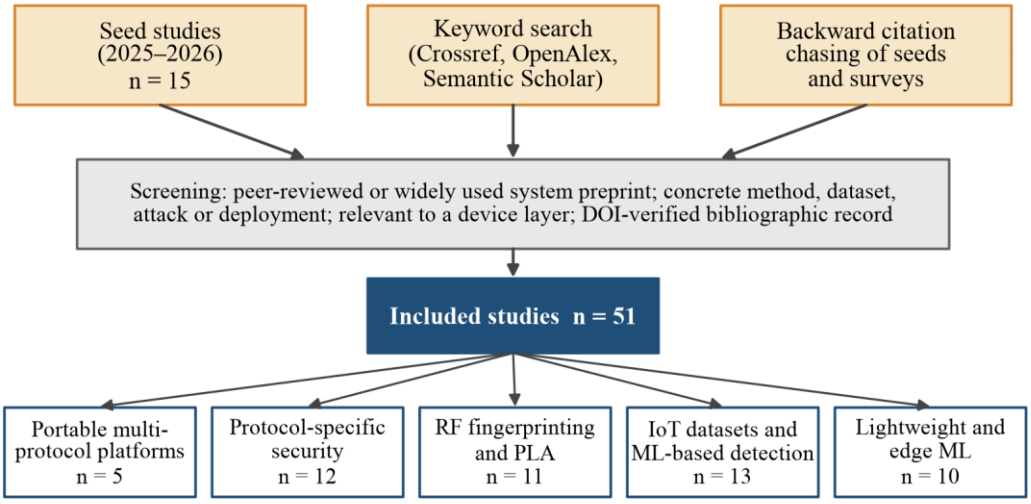


Fig. 1. Review protocol: identification, screening and thematic clustering of the included studies.

3. BACKGROUND

This section introduces the concepts that the surveyed works build on. It first describes the wireless interfaces that make up the attack surface of a typical IoT environment, then the portable platforms used to interact with them, and finally the ML methods and evaluation measures used for automated

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

analysis. Fig. 2 organises these concepts into a taxonomy that is used throughout the remainder of the review.

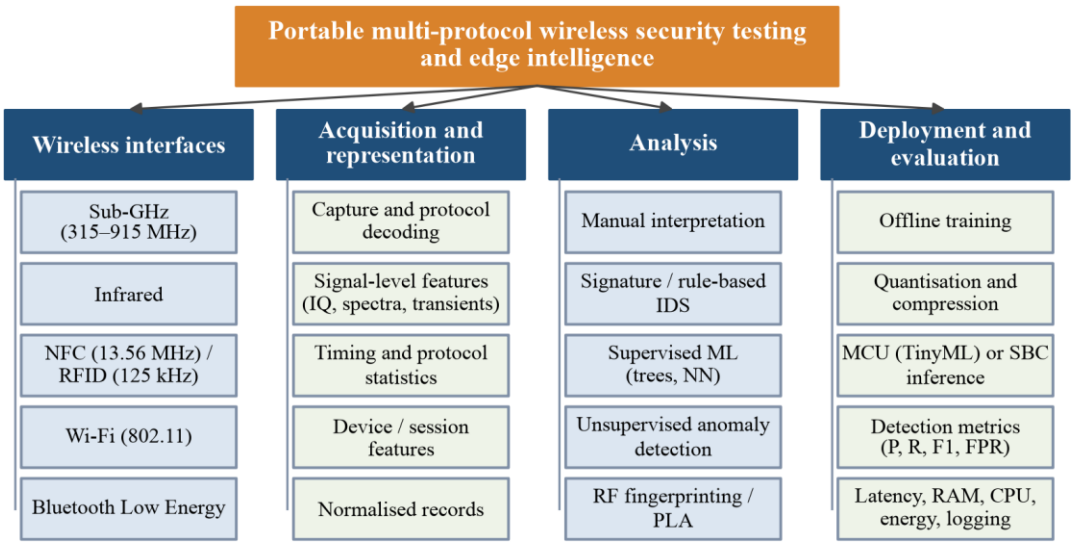


Fig. 2. Taxonomy of research relevant to portable multi-protocol wireless security testing with edge intelligence.

3.1. Wireless interfaces of the IoT attack surface

Table 2 summarises the five interface families considered in this survey, together with representative weaknesses reported in the literature. Although the families differ widely in range, data rate and protocol complexity, they share two properties that are relevant to automated analysis: their security frequently depends on implementation details that are invisible at higher layers, and their traffic can be captured passively by inexpensive hardware.

Table 2
Wireless interfaces considered in this survey and representative weaknesses reported in the literature.

Interface	Typical band	Example devices	Representative weaknesses	Refs.
Sub-GHz	315/433/868/915 MHz	Garage and gate openers, car key fobs, alarm sensors	Fixed-code replay; weak rolling-code ciphers and key management; jamming	[1,30,31]
Infrared	≈940 nm, ≈38 kHz carrier	Remote controls, appliances, surveillance cameras	No authentication or encryption; trivial replay; covert channels	[7,32]
NFC / RFID	13.56 MHz HF; 125 kHz LF	Access badges, payment and transit cards, product tags	Cloning, skimming, eavesdropping, relay; broken proprietary ciphers; counterfeit tags	[2,14,15,29]
Wi-Fi	2.4/5 GHz	Cameras, hubs, smart appliances, sensors	Handshake flaws; unprotected management frames enabling deauthentication	[3,10,33]
BLE	2.4 GHz	Wearables, item trackers, smart locks	Tracking via advertising data; spoofing; firmware tampering; location forgery	[4,12,13,28]

3.1.1. Sub-GHz radio

Licence-free bands around 315, 433, 868 and 915 MHz are used by garage and gate openers, car key fobs, wireless doorbells, weather stations and alarm sensors. Many low-cost devices transmit fixed codes with simple amplitude- or frequency-shift keying, which makes them vulnerable to capture-and-replay. Rolling-code schemes mitigate replay, but their security depends on the underlying cipher and key management. Garcia et al. showed that remote keyless entry systems used in a large number of vehicles relied on a small number of global master keys or on the weak Hitag2 cipher, allowing an attacker to clone a key after eavesdropping on very few transmissions [1]. Wouters et al. demonstrated that a passive keyless entry and start system protected by a 40-bit proprietary cipher could be cloned within seconds [30]. On the defensive side, Joo et al. showed that physical-layer characteristics of key-fob transmissions can be used to fingerprint legitimate keys and detect relayed or forged signals [31], which anticipates the RF fingerprinting approaches discussed in Section 4.3.

3.1.2. Infrared

Consumer IR remote controls modulate a near-infrared carrier, typically around 38 kHz, using protocols such as NEC, RC5/RC6 or SIRC [7]. These protocols provide neither authentication nor confidentiality, so any recorded command can be replayed by a device in line of sight. IR is usually regarded as a low-risk channel, but its ubiquity in building infrastructure gives it security relevance: Guri and Bykhovsky

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

showed that the IR illumination of surveillance cameras can be modulated to create covert exfiltration and infiltration channels for air-gapped networks [32]. IR is therefore relevant to a multi-protocol testing device both as a target and as an unexpected side channel.

3.1.3. NFC and RFID

Low-frequency (125 kHz) RFID and high-frequency (13.56 MHz) NFC are widely used for access control, payments, ticketing and product authentication. Kumar et al. classify the threats to RFID systems into eavesdropping, skimming, cloning, tracking, relay, replay and denial-of-service attacks, and note that many deployed tags lack the computational capacity for strong cryptography [29]. The cryptanalysis of the MIFARE Classic Crypto-1 cipher by Garcia et al. remains the canonical example of a widely deployed proprietary scheme failing under scrutiny [2]. Because cryptographic protection is often weak or absent, recent work has turned to physical-layer properties of tags, discussed in Section 4.3, and to learning-based threat detection for NFC-enabled IoT [14].

3.1.4. Wi-Fi

IEEE 802.11 networks are the dominant local connectivity for IoT devices that need moderate bandwidth. Two classes of weakness are particularly relevant here. Protocol-level flaws, exemplified by the key reinstallation attacks (KRACK) against the WPA2 four-way handshake, can undermine confidentiality even when strong ciphers are used [3]. Management-frame attacks, especially deauthentication flooding, exploit the fact that management frames are unprotected unless Protected Management Frames (IEEE 802.11w) are enabled; Drăghici et al. confirmed experimentally that such attacks interrupt the sensor-to-cloud connectivity of an ESP32-based IoT sensor [10]. The AWID corpus by Kolias et al. was one of the first public datasets to support ML-based intrusion detection for 802.11 networks [33].

3.1.5. Bluetooth Low Energy

BLE is used by wearables, medical devices, smart locks and item trackers. Barua et al. survey the BLE threat landscape, covering pairing weaknesses, eavesdropping, man-in-the-middle attacks, denial of service and privacy threats [28]. Celosia and Cunche showed that information carried in BLE advertisements can defeat the address randomisation intended to prevent tracking [4], and Antonioli et al. demonstrated impersonation attacks against the authentication procedure of Bluetooth Classic secure connections [34]. Item trackers combine BLE with crowd-sourced location networks and cloud services, which multiplies the attack surface [12,13].

3.2. Portable security testing platforms

Two families of hardware platform are used for portable wireless experimentation. The first comprises integrated microcontroller-based devices. Flipper Zero is built around an STM32WB55 microcontroller with an Arm Cortex-M4 application core and a Cortex-M0+ radio core sharing 1 MB of flash and 256 KB of SRAM; it integrates a CC1101 Sub-GHz transceiver, an ST25R3916 NFC front-end, a 125 kHz RFID circuit, an IR transceiver, BLE and 13 user GPIO pins [7]. Such devices are compact, power-efficient and immediately usable, but their computational budget is in the range targeted by TinyML rather than by general-purpose analytics, and their hardware architecture is fixed by the vendor. Single-purpose microcontroller tools, such as ESP8266-based deauthentication devices, fall into the same family [10].

The second family comprises single-board computers (SBCs), of which the Raspberry Pi is the most widely used. An SBC runs a general-purpose Linux distribution on a multi-core application processor, exposes a 40-pin GPIO header for attaching radio modules [35], and can host standard software stacks for packet capture, signal processing and ML. Sforzin et al. showed as early as 2016 that a Raspberry Pi could act as a portable intrusion detection system for IoT environments using a signature-based engine [36]. SBC-based platforms trade battery life and form factor for flexibility, which makes them attractive as open research platforms on which acquisition, feature extraction and inference can all be modified. Fig. 3 contrasts the processing chain of a tool-centric device with that of an analysis-capable edge platform; the distinction between the two is central to the comparison in Section 5.

Between these two families lie microcontroller boards that are used as edge-ML targets rather than as security tools, such as the Arduino, ESP32 and STM32 boards on which recent intrusion detectors have been benchmarked [23,24], and laboratory set-ups built around software-defined radios (SDRs), which offer the greatest signal-level flexibility but depend on a host computer [37,38]. Table 3 compares these platform classes with respect to the properties that matter for portable, analysis-capable operation. The comparison shows a clear trade-off: the classes that offer the richest radio access are either closed and resource-limited or not portable, while the class that offers an open software stack and ample computing resources, the SBC, requires external radio modules for most non-Wi-Fi and non-Bluetooth interfaces.

Table 3
Comparison of hardware platform classes used for portable wireless experimentation and edge ML.

Platform class	Example	Compute and software	Radio interfaces	Extensibility	On-device ML suitability	Refs.
Integrated MCU multi-tool	Flipper Zero	Cortex-M4 at 64 MHz with 256 KB SRAM; embedded firmware	Sub-GHz, 13.56 MHz NFC, 125 kHz RFID, IR, BLE, GPIO	Applications and GPIO modules; fixed vendor hardware	TinyML-scale models only	[7,9]
Single-purpose MCU tool	ESP8266 Deauther	Wi-Fi SoC; embedded firmware	Wi-Fi	Low	Very limited	[10]

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Table 3
Comparison of hardware platform classes used for portable wireless experimentation and edge ML.

Platform class	Example	Compute and software	Radio interfaces	Extensibility	On-device ML suitability	Refs.
MCU edge-ML target	Arduino Mega 2560, ESP32-C3, STM32H7 board	8–400 KB SRAM in the cited evaluations; bare-metal or RTOS	Board-dependent	Moderate	Quantised trees and MLPs verified on hardware	[23,24]
Single-board computer	Raspberry Pi	Multi-core application processor; Linux	On-board Wi-Fi and Bluetooth on recent models; 40-pin GPIO for add-on radios	High; open software stack	Standard ML libraries; larger models	[35,36]
Laboratory SDR set-up	SDR with host computer	Host PC or workstation	Wideband, software-defined	Very high	High on the host; not portable	[37,38]

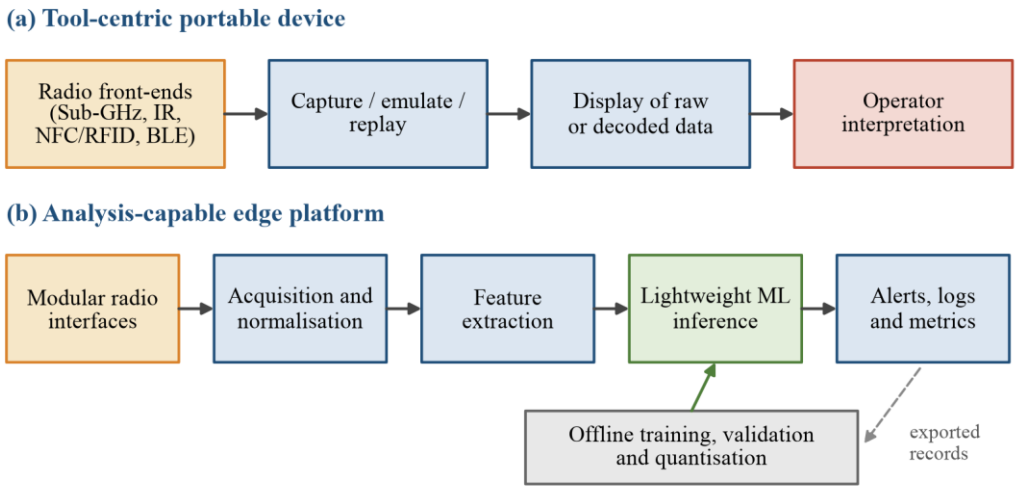


Fig. 3. Processing chain of (a) a tool-centric portable device and (b) an analysis-capable edge platform with offline-trained, on-device ML.

3.3. Machine learning methods for wireless and IoT anomaly detection

Chandola et al. define anomaly detection as the problem of finding patterns in data that do not conform to expected behaviour, and distinguish supervised, semi-supervised and unsupervised settings according to the availability of labels [39]. All three settings appear in the surveyed literature.

3.3.1. Tree-based models

Decision trees (DT), random forests (RF) and gradient-boosted trees such as XGBoost and LightGBM dominate the lightweight end of IoT intrusion detection. They handle heterogeneous tabular features without normalisation, their inference cost grows with tree depth rather than input dimensionality, and individual trees can be compiled directly to branching code for microcontrollers. De Bernardis et al. report a depth-5 decision tree executing in 5.6 µs on an ESP32-C3 microcontroller [24], which illustrates why tree models are a natural baseline for constrained devices. Their main limitation is that they require labelled examples of each attack class and do not learn representations from raw signals.

3.3.2. Neural models

Feed-forward networks or multi-layer perceptrons (FNN/MLP), convolutional neural networks (CNN), recurrent networks (RNN, LSTM, GRU) and temporal convolutional networks (TCN) are used when features must be learned from sequences or raw signals. CNNs are the standard choice for RF fingerprinting from in-phase/quadrature (IQ) samples [25,37], whereas recurrent and temporal models are used for flow sequences [23]. Neural models can be compressed by quantisation, pruning or architectural simplification, which is the focus of TinyML research [21]. Their weaknesses are larger training data requirements, sensitivity to distribution shift and limited interpretability.

3.3.3. Unsupervised and one-class models

Autoencoders learn to reconstruct benign behaviour and flag observations with high reconstruction error; N-BaIoT trained one autoencoder per device to detect botnet traffic [16], and Kitsune used an ensemble of small autoencoders for online, unsupervised intrusion detection [40]. Isolation forest isolates anomalies by random partitioning and has low memory and time requirements [41]. In the radio domain, Rajendran et al. proposed SAIFE, an adversarial-autoencoder approach to

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

unsupervised spectrum anomaly detection with interpretable features [42]. Unsupervised methods are attractive for a field device because labelled attack data are scarce, but Sommer and Paxson caution that anomaly detection in security settings suffers from high costs of false alarms, diverse benign traffic and a semantic gap between “unusual” and “malicious” [43].

3.4. Performance evaluation

Detection quality is commonly evaluated with the confusion-matrix counts of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), treating the anomalous class as positive:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$Precision = TP / (TP + FP) \quad (2)$$

$$Recall = TP / (TP + FN) \quad (3)$$

$$F1 = 2 \cdot Precision \cdot Recall / (Precision + Recall) \quad (4)$$

$$FPR = FP / (FP + TN) \quad (5)$$

For edge deployment, these measures must be complemented by resource measures: inference latency per observation, peak RAM, flash or storage footprint, the number of multiply–accumulate operations (MACCs) and, for battery-powered devices, energy per inference [23,27,44]. When a 32-bit floating-point model is quantised to 8-bit integers, real values r are represented by integers q through an affine mapping with scale S and zero-point Z [45]:

$$r = S (q - Z) \quad (6)$$

and the resulting compression ratio is CR = M(FP32) / M(INT8), where M denotes the stored model size. Because accuracy, false positive rate and latency interact, several authors combine them into a single deployability score [23]. In the context of a portable testing device, the false positive rate deserves particular emphasis: a detector that raises frequent false alarms during an assessment quickly loses the operator's trust [43].

Beyond the choice of metrics, the reliability of reported results depends on the evaluation protocol. Three pitfalls recur in the surveyed literature. First, intrusion datasets are typically imbalanced, so overall accuracy can remain high even when minority attack classes are missed; per-class or macro-averaged F1-scores and confusion matrices are more informative. Second, random train–test splits of flow records allow near-duplicate records from the same session to appear on both sides of the split, inflating performance; splits by time, device or capture session are more realistic. Third, results obtained on one dataset rarely transfer to another: De Bernardis et al. measured normalised distribution shifts of up to 5.95 in a packet-asymmetry feature between TON_IoT and CICIoT2023 [24], and Sommer and Paxson argued that closed-world evaluation systematically overestimates field performance [43]. For a portable device that is moved between environments, cross-environment testing should therefore be regarded as a primary rather than an optional evaluation.

3.5. Strengths and limitations of ML methods

Table 4 summarises the strengths and limitations of the ML model families with respect to the requirements of an analysis-capable portable device. No single family dominates: tree ensembles offer the best accuracy-to-cost ratio on engineered tabular features, convolutional models are required when the input is a raw radio signal, and unsupervised models are the only option when labelled attack data are unavailable.

Table 4
Strengths and limitations of ML model families for an analysis-capable portable device.

Family	Typical models	Strengths	Limitations	Suitability (evidence)
Tree-based	DT, RF, XGBoost, LightGBM	Strong on tabular features; small; microsecond inference; compilable to C	Require labelled attacks; no representation learning from raw signals	High for engineered features [22,24]
Feed-forward NN	FNN, MLP	Compact; INT8-quantisable; good accuracy–cost trade-off	Need normalised inputs; limited interpretability	High after quantisation [22,23]
Convolutional / temporal	CNN, TCN	Learn features from IQ samples or sequences	More MACCs and memory; sensitive to channel and device drift	Medium; needed for RFF [27,37]
Recurrent	RNN, LSTM, GRU	Model long temporal dependencies	Higher latency and RAM on MCUs	Low–medium [23]
Unsupervised / one-class	Autoencoders, isolation forest, adversarial autoencoder	No attack labels required; can flag novel behaviour	Higher false alarms; threshold tuning; semantic gap	Medium–high in the field [40–43]

4. SURVEY OF EXISTING RESEARCH

This section reviews the literature in five thematic streams, following the taxonomy of Fig. 2. Each subsection closes with a comparative table that records the year, approach, main contribution and principal limitation of the reviewed works.

4.1. Portable multi-protocol security testing devices

Academic literature on portable multi-protocol devices is recent and still small, which is itself a finding of this survey. Cass provided an early hands-on assessment of Flipper Zero and framed the central tension that subsequent studies return to: the same device is a debugging aid for researchers and hardware developers and a concern for those responsible for physical and network security [8]. Muñoz-Ruiz et al. systematically exercised the infrared, RFID/NFC, Sub-GHz, USB and Bluetooth functions of the device against everyday systems and identified practical weaknesses such as access-credential emulation and interference with remote-control

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

signals [9]. Their study is valuable as a capability inventory, but its analysis is qualitative: the device performs the interaction and the researcher interprets the outcome.

Drăghici et al. compared Flipper Zero with an ESP8266-based deauther in a controlled laboratory study of IEEE 802.11 deauthentication against an ESP32 temperature sensor, describing the command sets of both tools and evaluating defences such as Protected Management Frames and adaptive channel hopping [10]. Garrido Briones et al. used Flipper Zero to measure the distance dependence of replay attacks on authorised RFID and Sub-GHz access-control testbeds and proposed a reproducible red-teaming protocol for education [11]. These two studies are notable for their emphasis on controlled, authorised experimentation and reproducibility, which are prerequisites for using a portable device as a research instrument rather than as an attack tool.

On the analysis side, the closest precedent is the Raspberry Pi based intrusion detection system of Sforzin et al., which demonstrated that an inexpensive SBC can monitor IoT traffic in real time with a signature-based engine [36]. However, it addressed IP traffic only, used rules rather than learned models, and did not interact with non-IP radio interfaces. Table 5 compares the reviewed platform studies. Two observations follow. First, the existing devices either interact with many protocols but leave analysis to the human, or analyse traffic automatically but on a single protocol family. Second, none of the reviewed studies reports quantitative detection metrics for an on-device ML model operating on data captured by the device itself.

Critical assessment. The platform literature answers RQ1 only partially. It establishes what portable devices can do but not how well their captures can support automated decisions, because none of the studies treats the device as a data source for a learning pipeline. Reproducibility also varies: magazine-style and capability-inventory studies document what was attempted, whereas the laboratory protocols of [10] and [11] specify set-ups and distances precisely enough to be repeated. Future platform studies would benefit from reporting capture formats, sampling parameters and the proportion of interactions that succeed under controlled conditions, which would make their results usable as ground truth for ML studies.

Table 5
Comparison of studies on portable multi-protocol security testing platforms.

Study	Year	Platform	Interfaces	Analysis	Main contribution	Principal limitation
Cass [8]	2023	Flipper Zero	Sub-GHz, NFC/RFID, IR, BLE, GPIO	Manual	Hands-on assessment; frames the dual-use tension	Magazine review; no measurements
Muñoz-Ruiz et al. [9]	2026	Flipper Zero	IR, RFID/NFC, Sub-GHz, USB, Bluetooth	Manual	Capability inventory; credential emulation and remote-control interference	Qualitative; no automated analysis
Drăghici et al. [10]	2025	Flipper Zero; ESP8266 Deauther	Wi-Fi	Manual, lab-measured	Deauthentication of an ESP32 sensor; PMF and channel-hopping defences	Single attack class; no detector
Garrido Briones et al. [11]	2026	Flipper Zero	RFID, Sub-GHz	Manual, measured	Distance-dependent replay; reproducible educational red-teaming protocol	Pilot scale; replay only
Sforzin et al. [36]	2016	Raspberry Pi	IP traffic	Signature-based IDS	Feasibility of a portable SBC-based IDS for IoT	No non-IP radio interfaces; no ML

4.2. Protocol-specific security analysis

Protocol-specific studies provide the threat models and ground truth that any automated analysis must be built on. For BLE, Alamleh et al. analysed Apple AirTags and Samsung Galaxy SmartTags across five layers, namely hardware, firmware, radio signals, companion applications and cloud services, and identified weaknesses including spoofing, firmware tampering, jamming and location forgery [12]. They report that privacy-oriented design on one platform leaves authentication gaps while cloud-centred design on the other increases data-exposure risk, and that neither design uses secure boot. The methodological lesson is that a vulnerability visible at one layer can only be explained by evidence from another, which argues for acquisition pipelines that record radio-level and protocol-level observations together. Earlier work by Heinrich et al. on Apple's Find My network [13] and by Celosia and Cunche on BLE advertising privacy [4] supports the same conclusion from the privacy perspective.

For NFC, Rehman et al. proposed DC-NFC, a deep learning framework composed of a component for capturing temporal and spatial patterns, a privacy-preserving module enforcing end-to-end privacy constraints, and an adaptive threat-detection module for real-time detection and model adaptation [14]. The framework was evaluated on UNSW-NB15, Bot-IoT, TON_IoT telemetry and Edge-IIoTset, reaching accuracy of up to 95% on UNSW-NB15 and F1-scores above 0.90 across datasets. The study is important because it explicitly targets NFC-enabled IoT; at the same time, it illustrates a recurring limitation of the field, namely that the evaluation relies on network-level IoT datasets rather than on captures of NFC exchanges, so the link between the detected patterns and NFC-specific attacks such as relay or eavesdropping is indirect.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

For Sub-GHz and RFID, the attacks of Garcia et al. on remote keyless entry [1] and on MIFARE Classic [2], and the keyless-entry cloning of Wouters et al. [30], show that cryptographic weaknesses are often discovered only after mass deployment, by which time hardware cannot be patched. This motivates detection-oriented and physical-layer defences that do not require changes to deployed devices. For Wi-Fi, KRACK [3] and management-frame attacks [10] play the same role, and AWID [33] shows how such attacks can be converted into labelled datasets. Table 6 compares the protocol-specific studies.

Critical assessment. Protocol studies provide the most reliable evidence about what can go wrong, because they are grounded in real devices. Their limitation, from the perspective of this survey, is that their output is knowledge rather than reusable artefacts: few release captures, labelled traces or detection rules that could be used to train or validate a detector. Moreover, the most severe vulnerabilities, such as broken proprietary ciphers in keyless entry and access-control cards [1,2,30], are design flaws that cannot be fixed in the field. In such cases, the realistic defensive contribution of a portable device is detection and auditing rather than prevention, which strengthens the case for embedding analysis capabilities in testing tools.

Table 6
Comparison of protocol-specific security studies.

Study	Year	Protocol / target	Approach	Key finding	Principal limitation
Alamleh et al. [12]	2025	BLE trackers	Five-layer analysis: hardware, firmware, radio, app, cloud	Spoofing, firmware tampering, jamming, location forgery; no secure boot	Two products; no automated detection
Heinrich et al. [13]	2021	BLE (Find My)	Reverse engineering; security and privacy analysis	Weaknesses in crowd-sourced location tracking	Single ecosystem
Celosia and Cunche [4]	2019	BLE advertising	Analysis of advertised data	Address randomisation undermined by advertising content	Privacy only
Antonioli et al. [34]	2020	Bluetooth Classic	Protocol attack	Impersonation during secure-connection authentication	BR/EDR rather than BLE
Rehman et al. [14]	2025	NFC-enabled IoT	DL framework with privacy and adaptive threat modules	Accuracy up to 95% (UNSW-NB15); F1 > 0.90	Evaluated on network datasets, not NFC captures
Garcia et al. [2]	2008	MIFARE Classic	Cryptanalysis	Crypto-1 weaknesses enable key recovery	One card family
Kumar et al. [29]	2021	RFID	Taxonomy	Systematic classification of threats and countermeasures	No experiments
Garcia et al. [1]	2016	Sub-GHz RKE	Cryptanalysis and RF capture	Global master keys and Hitag2 weaknesses enable key cloning	Specific vehicle families
Wouters et al. [30]	2019	Passive keyless entry	Reverse engineering	40-bit proprietary cipher allows key-fob cloning	One system
Guri and Bykhovsky [32]	2019	Infrared	Covert channel	Camera IR illumination used for air-gap exfiltration	Not remote-control protocols
Vanhoef and Piessens [3]	2017	Wi-Fi (WPA2)	Protocol attack	Key reinstallation in the four-way handshake	Mitigated in patched clients

4.3. RF fingerprinting and physical-layer authentication

RF fingerprinting exploits small, device-specific imperfections in transmitter hardware, such as oscillator offsets, power-amplifier non-linearities and transient behaviour, to identify devices from their signals rather than from the identifiers they claim. Jagannath et al. survey traditional feature-engineered approaches and deep learning approaches, and highlight open challenges including sensitivity to channel conditions, limited dataset diversity and robustness against adversaries [25]. Sankhe et al. demonstrated with ORACLE that a CNN operating on raw IQ samples can distinguish nominally identical software-defined radio transmitters [37].

Several recent studies apply this idea directly to the interfaces of interest here. Adigopula and Subramanyam proposed DCNN-RFF-NFC, a deep convolutional network that identifies NFC tags from their RF characteristics, enabling tags that carry identical logical data to be distinguished [26]. Lee and Baek likewise used deep neural networks to fingerprint NFC tags [46]. For chipless RFID, Khan et al. generated fingerprints from the backscattered electromagnetic responses of tags captured by a portable multiple-input multiple-output (MIMO) reader; using one authentic and twenty-one counterfeit tags, they reported clone-detection accuracy of 98.41% with frequency-domain fingerprints under 35 dB noise and 92.07% with time-domain fingerprints under 50 dB noise, using similarity measures rather than a trained classifier [15]. Wu et al. addressed the degradation of physical-layer RFID classification at low signal-to-noise ratio by a cognitive risk control strategy that adjusts tag distance, together with an

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

enriched set of 104 time- and frequency-domain features; using software-defined radio and seven tag types from three manufacturers, they report an average accuracy gain of about 11% from risk control and further gains of 4.3–5.3% over smaller feature sets [38].

Resource consumption is addressed explicitly in three studies. Owusu et al. designed RFNet, a full-capacity CNN achieving 97.48% identification accuracy on 30 commercial LoRa devices, and TinyRFNet, which keeps 93.19% accuracy with more than 34 times fewer parameters; an energy-aware policy that switches between them according to battery level, prediction confidence and context achieved 94.54% overall accuracy with 17% lower energy consumption than the large model alone [27]. Shiltagh Al-Jamali et al. combined RF fingerprinting with a spiking convolutional neural network and reported better accuracy than one-dimensional CNN and spiking baselines with lower memory use and execution time than CNN-based alternatives on a public dataset from Northeastern University [47]. Yao et al. studied lightweight physical-layer authentication for IoT device access under jamming attacks, targeting authentication from radio characteristics at a computational cost compatible with constrained IoT devices and identifying devices whose signals do not match enrolled profiles [48]. Finally, Joo et al. showed that the same principle can protect vehicles by fingerprinting key-fob transmissions [31].

Table 7 compares these studies. The stream as a whole demonstrates convincingly that signal-level features carry security-relevant information that protocol-level inspection cannot see. Its limitations are equally consistent: each study addresses one protocol and one hardware setup, most rely on laboratory-grade receivers or software-defined radios, and with the exception of [27] the energy and latency of inference on portable hardware are not measured.

Critical assessment. Three methodological differences complicate comparison within this stream. The studies use different receivers, ranging from commercial readers to portable MIMO readers and SDRs, and fingerprints are known to depend on the receiver as well as the transmitter [25]. Population sizes differ by an order of magnitude, from 22 tags [15] to 30 LoRa devices [27], and small populations make high accuracy easier to achieve. Finally, authentication accuracy at a fixed distance and noise level may not hold when a handheld device is moved, as the distance sensitivity reported by Wu et al. illustrates [38]. Reporting performance as a function of distance, orientation and receiver, and testing on enrolment sessions separated in time, would make results more transferable to portable use.

Table 7
Comparison of RF fingerprinting and physical-layer authentication studies.

Study	Year	Target	Method	Reported result	Principal limitation
Sankhe et al. [37]	2019	SDR transmitters	CNN on raw IQ samples (ORACLE)	Classifies nominally identical radios	Laboratory SDRs; channel sensitivity
Joo et al. [31]	2020	Car key fobs	Physical-layer fingerprinting	Detects forged or relayed key signals	Single application
Jagannath et al. [25]	2022	General	Survey	Traditional vs. DL RFF; open challenges	No new experiments
Adigopula and Subramanyam [26]	2025	NFC tags	Deep CNN RF fingerprinting	Distinguishes tags carrying identical data	NFC only; laboratory setup
Lee and Baek [46]	2026	NFC tags	DNN-aided RF fingerprinting	DNN-based identification of NFC tags	NFC only
Khan et al. [15]	2025	Chipless RFID	Portable MIMO reader; MSE/SSI similarity	98.41% (FD, 35 dB); 92.07% (TD, 50 dB)	22 tags; one authentic tag
Wu et al. [38]	2025	RFID tags (7 types)	SDR; cognitive risk control; 104 features	≈11% accuracy gain at low SNR	Requires control of tag distance
Owusu et al. [27]	2025	LoRa devices (30)	RFNet / TinyRFNet; energy-aware switching	94.54% accuracy; 17% less energy	LoRa only
Shiltagh Al-Jamali et al. [47]	2025	IoT transmitters	RFF with spiking CNN (RFFI-SCNN)	Higher accuracy than 1CNN/1SNN; less memory and time than DCNN	Public dataset only
Yao et al. [48]	2025	IoT access under jamming	Lightweight PLA	Low-complexity authentication under jamming	Single threat scenario
Rajendran et al. [42]	2018	Spectrum	Adversarial autoencoder (SAIFE)	Unsupervised spectrum anomaly detection with interpretable features	Spectrum-level, not device-level

4.4. IoT intrusion detection datasets and ML-based detection

Datasets determine what an ML-based detector can learn, and their evolution reveals how the field's view of the IoT has changed. N-BaIoT captured traffic from commercial IoT devices infected with real botnets [16]; Bot-IoT [17] and TON_IoT [18] added larger testbeds, telemetry and operating-system data; Edge-IIoTset targeted industrial IoT and federated learning [49]; and CICIoT2023 provided a large real-time benchmark of attacks against a broad device topology [19]. Shafi and Lashkari analysed thirty public smart-home datasets, identified twenty-two shortcomings including

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

limited device diversity, absence of multimodal data and lack of multi-protocol support, and introduced BCCC-IoT-IDS-Zwave-2025 [20]. Their dataset was collected over five months from more than fifty devices and more than eighty attack scenarios and, unlike its predecessors, combines IP traffic, Z-Wave radio communication, device activity and MQTT traffic and logs; they also released IoT-ZwaveNetLyzer, a traffic analyser for Z-Wave. State-of-the-art classifiers exceeded 95% average detection accuracy on the dataset.

On the detection side, Al-Garadi et al. survey ML and DL methods for IoT security and conclude that DL methods are promising but that their computational demands are a barrier for constrained devices [6]. Kitsune [40] and N-BaIoT [16] established autoencoder-based detection as a strong unsupervised baseline. Table 8 compares the datasets with respect to the properties that matter for a multi-protocol portable device. Two gaps stand out. Apart from AWID for 802.11 [33] and the Z-Wave component of BCCC-IoT-IDS-Zwave-2025 [20], public datasets describe IP-level traffic; none of the reviewed datasets contains labelled Sub-GHz, IR, NFC/RFID and BLE observations captured with a common schema. Moreover, the high accuracies reported on these datasets are known to transfer poorly across datasets: De Bernardis et al. quantified large normalised distribution shifts between TON_IoT and CICIoT2023 [24], echoing the long-standing warning of Sommer and Paxson about closed-world evaluation [43].

Critical assessment. The dataset literature has responded to criticism of earlier corpora by increasing scale, device diversity and data modalities, and the multi-source design of BCCC-IoT-IDS-Zwave-2025 is the closest existing match to the needs of a multi-protocol device [20]. Nevertheless, datasets remain tied to the testbeds on which they were captured, and the features they provide are computed by specific flow analysers. A detector trained on such features presupposes the same feature extractor at deployment time, which a portable device capturing raw radio activity does not have. This mismatch between the representation used in training and the representation available in the field is, in the authors' assessment, the single largest obstacle to reusing existing IoT detection results on portable hardware.

Table 8
Comparison of IoT and wireless intrusion detection datasets.

Dataset	Year	Data sources	Protocol coverage	Relevance and limitation for portable multi-protocol devices
AWID [33]	2016	IEEE 802.11 frames	Wi-Fi	Radio-adjacent and labelled; Wi-Fi only
N-BaIoT [16]	2018	Traffic of commercial IoT devices infected by botnets	IP	Per-device benign baselines; botnet-specific
Bot-IoT [17]	2019	Realistic testbed network traffic	IP	Large and labelled; IP only
TON_IoT [18]	2021	Network traffic, IoT telemetry, operating-system logs	IP + telemetry	Multi-source; no radio layer
Edge-IIoTset [49]	2022	IoT/IIoT network traffic	IP (IIoT)	Supports centralised and federated learning; no radio layer
CICIoT2023 [19]	2023	Network traffic from a large IoT topology	IP	Real-time benchmark of many attacks; no radio layer
BCCC-IoT-IDS-Zwave-2025 [20]	2025	IP traffic, Z-Wave signals, device activity, MQTT	IP + Z-Wave	Multi-source and multi-protocol; no Sub-GHz remotes, IR, NFC/RFID or BLE

4.5. Lightweight and edge machine learning

Edge computing moves computation towards the data source to reduce latency, bandwidth and privacy exposure [50]. TinyML pushes this idea to microcontrollers with kilobytes of memory, relying on embedded runtimes such as TensorFlow Lite Micro [51] and on benchmarks such as MLPerf Tiny for comparable measurement [44]; Ray reviews the state of the art and its prospects [21]. Integer-only quantisation, which represents weights and activations as 8-bit integers using the affine mapping of Eq. (6), is the most widely adopted compression technique [45]. Fig. 4 shows the offline-training and edge-deployment workflow shared by the studies reviewed in this subsection.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

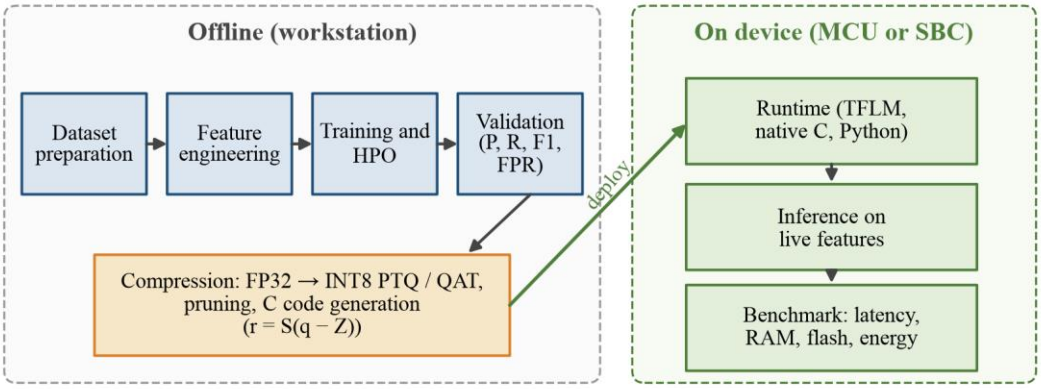


Fig. 4. Offline-training and edge-deployment workflow common to the reviewed lightweight ML studies.

Lundqvist et al. compared a decision tree, LightGBM, an FNN and a TinyML-optimised FNN for multiclass intrusion detection on TON_IoT under realistic resource limits [22]. The TinyML FNN gave the best balance between accuracy and deployability, with an F1-score of 0.976, a throughput of approximately 120,000 packets per second and a model size of only 31 KB. Mazinani et al. evaluated LSTM, GRU, RNN, CNN, TCN and MLP models on CICIoT2023 for 2-, 8- and 34-class detection, tuned with Bayesian optimisation and profiled in terms of MACCs, RAM and inference time on an STM32H7S78-DK board [23]. Using a trade-off score that balances accuracy and computational cost, MLP and TCN performed best, and post-training INT8 quantisation reduced model size by more than 90% with negligible loss of performance.

De Bernardis et al. extended the analysis to physical execution on two microcontrollers, an 8-bit Arduino Mega 2560 with 8 KB of SRAM and a RISC-V ESP32-C3 with 400 KB of SRAM [24]. Logistic regression, a depth-5 decision tree, random forest, XGBoost, LightGBM and an MLP were trained on TON_IoT; LightGBM reached an F1-score of 0.9993 in five-fold cross-validation. Three quantisation pipelines were compared: generated C code for the lightest models, TensorFlow Lite Micro full-integer INT8 for the MLP (a 9.34-fold size reduction to 13.03 KB), and a custom INT8 format for the boosted ensembles (an 18.91-fold reduction to 73.85 KB for LightGBM, which retained an F1-score of 0.9992). Firmware predictions were verified to match the floating-point models, and the RISC-V device was 5.8–7.8 times faster than the 8-bit device for identical model code. Hamdouchi and Idri compared singular and ensemble TinyML classifiers for IoT intrusion detection [52], and Fusco et al. addressed the scarcity of labelled data and the communication cost of collaborative training by combining Siamese networks with reinforcement-learning-driven gradient sparsification for few-shot federated TinyML intrusion detection [53].

Table 9 compares the edge-ML studies. Collectively, they establish that accurate intrusion detection is feasible within tens of kilobytes and microseconds to milliseconds of latency, and they provide a methodology, consisting of offline training, quantisation, firmware verification and on-device benchmarking, that transfers directly to portable devices. Their common limitation is the input: all of them classify features derived from IP network flows, so their results do not yet show whether the same budgets suffice for heterogeneous radio observations.

Critical assessment. The edge-ML stream is methodologically the most mature of the five: recent studies report hardware, memory, latency and post-quantisation accuracy, verify firmware outputs against reference models [24] and introduce explicit accuracy–cost trade-off scores [23]. Its open questions concern scope rather than rigour. The reviewed studies classify one input stream with one model, whereas a multi-protocol device would run several feature pipelines and possibly several models concurrently; energy is measured only in the RF fingerprinting work of Owusu et al. [27]; and continual adaptation after deployment is addressed only through federated few-shot learning [53].

Table 9
Comparison of lightweight and edge ML studies for intrusion detection.

Study	Year	Models	Dataset	Hardware / target	Optimisation	Reported result
Lundqvist et al. [22]	2025	DT, LightGBM, FNN, TinyML FNN	TON_IoT	Constrained IoT profile	TinyML-optimised FNN	F1 0.976; 31 KB; ≈120,000 packets/s
Mazinani et al. [23]	2026	LSTM, GRU, RNN, CNN, TCN, MLP	CICIoT2023	STM32H7S78-DK	Bayesian HPO; INT8 PTQ	MLP/TCN best trade-off; >90% size reduction
De Bernardis et al. [24]	2026	LR, DT, RF, XGBoost, LightGBM, MLP	TON_IoT (+ CICIoT2023 shift)	Arduino Mega 2560; ESP32-C3	C code generation; TFLM INT8; custom INT8	LightGBM INT8 F1 0.9992 at 73.85 KB; DT 5.6 µs
Hamdouchi and Idri [52]	2025	Singular and ensemble classifiers	IoT IDS data	TinyML	TinyML deployment	Singular vs. ensemble comparison

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Table 9
Comparison of lightweight and edge ML studies for intrusion detection.

Study	Year	Models	Dataset	Hardware / target	Optimisation	Reported result
Fusco et al. [53]	2025	Siamese neural network	CSE-CIC-IDS2018	Federated TinyML devices	ε-greedy gradient sparsification	Few-shot federated IDS with reduced communication

5. COMPARATIVE ANALYSIS

The preceding section reviewed each stream on its own terms. This section compares them across the functional layers of an analysis-capable portable wireless security device, namely acquisition, representation, analysis and deployment, in order to determine which capabilities exist in combination and which exist only in isolation.

Table 10 maps representative studies against seven capabilities: multi-protocol acquisition, use of signal-level (physical-layer) features, use of protocol- or network-level features, automated ML-based analysis, on-device or edge inference, explicit model compression, and an open or reproducible platform. The pattern is clear. Platform studies provide multi-protocol acquisition and openness but no automated analysis [9,11]. Protocol-security studies provide deep, often cross-layer, insight into one protocol but rarely produce reusable detectors [1,12]. RF fingerprinting studies combine signal-level features with learned analysis, and occasionally with lightweight models [27,47], but always for a single protocol. Dataset and IDS studies provide rich labelled data and strong detectors, but only for IP-level and, in one case, Z-Wave traffic [19,20]. Edge-ML studies provide compression and on-device inference with verified hardware measurements [23,24], but on network-flow features. No reviewed study combines more than four of the seven capabilities, and none combines multi-protocol acquisition with on-device ML.

Table 10
Capability matrix of representative studies (● addressed, ◐ partially addressed, ○ not addressed). C1: multi-protocol acquisition; C2: signal-level features; C3: protocol/network-level features; C4: automated ML analysis; C5: on-device or edge inference; C6: model compression; C7: open or reproducible platform.

Representative studies	Stream	C1	C2	C3	C4	C5	C6	C7
Flipper Zero evaluations [9,11]	Platform	●	○	◐	○	○	○	◐
Deauthentication assessment [10]	Platform	○	○	●	○	○	○	◐
RPiDS [36]	Platform	○	○	●	○	●	○	●
BLE tracker analysis [12]	Protocol	○	◐	●	○	○	○	○
DC-NFC [14]	Protocol	○	○	◐	●	◐	○	○
RFID fingerprinting [15,38]	RFF	○	●	○	◐	◐	○	○
NFC fingerprinting [26,46]	RFF	○	●	○	●	○	○	○
Energy-aware RFF [27]	RFF	○	●	○	●	◐	●	○
BCCC-IoT-IDS-Zwave-2025 [20]	Dataset	◐	◐	●	●	○	○	●
TinyML IDS [22–24]	Edge ML	○	○	●	●	●	●	◐

Fig. 5 shows the distribution of the reviewed studies across the five thematic clusters and three publication periods. It confirms that research on portable multi-protocol platforms is both the smallest cluster and almost entirely recent, whereas protocol security and IoT detection have a longer history. The growth of RF fingerprinting and edge-ML work since 2024 indicates that the building blocks needed for analysis-capable portable devices are maturing at the same time as the devices themselves become popular.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

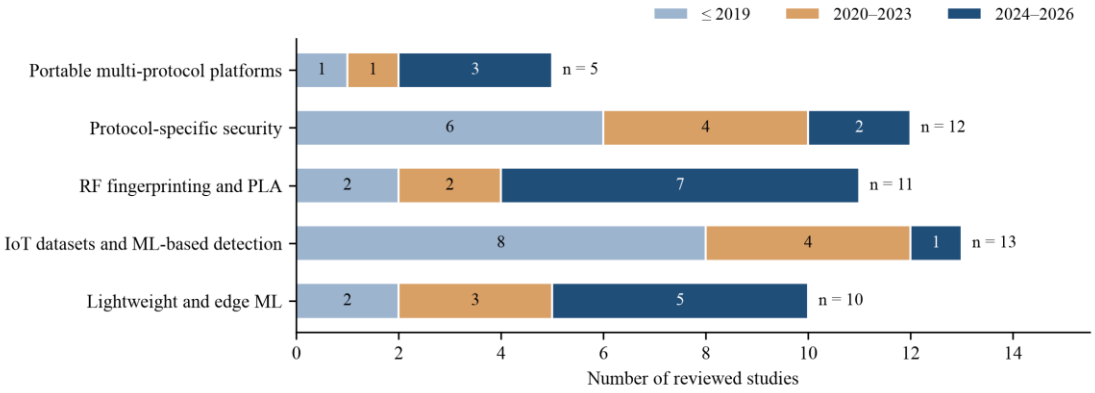


Fig. 5. Distribution of the reviewed studies across thematic clusters and publication periods.

Several further comparisons emerge from the tables.

Evaluation setting. Studies in the platform and protocol streams evaluate on real devices but report qualitative outcomes, whereas studies in the dataset and edge-ML streams report precise metrics but on curated datasets. RF fingerprinting studies sit between the two, using real radio captures with quantitative metrics, which is why they are the most directly transferable to a field device. A recurring weakness, noted for DC-NFC [14], is that protocol-specific claims are validated on traffic that does not originate from that protocol.

Feature representation. Signal-level features, such as IQ samples, backscatter spectra and transient shapes, enable device identification and counterfeit detection [15,26,38], while protocol-level and flow-level features enable attack classification [20,22]. These representations are complementary: a replayed Sub-GHz code is protocol-valid but may be physically distinguishable [31], whereas a protocol-level anomaly such as a deauthentication flood is invisible in the fingerprint of the transmitting radio. A portable analytics pipeline therefore benefits from combining both, but no reviewed work defines a common record format that carries both kinds of features across protocols.

Model choice. For tabular features, tree ensembles achieve the highest reported accuracy at small size and are straightforward to quantise or compile [22,24]; compact MLPs and TCNs are competitive when sequence information matters [23]. For raw signals, CNNs remain necessary, and the work of Owusu et al. shows that adaptive switching between a small and a large model can recover much of the accuracy lost by compression [27]. Unsupervised models are underused in the recent literature despite being better aligned with field conditions in which attacks are unlabelled [40,42].

Hardware target. The edge-ML literature targets microcontrollers with 8–400 KB of SRAM [23,24], while SBC-class devices offer orders of magnitude more memory and a full operating system [35,36]. This difference matters: on an SBC, the constraint is less the size of a single model than the end-to-end latency and power of acquisition, decoding, feature extraction and inference running concurrently, a combination that none of the reviewed studies measures.

Data and ground truth. Table 11 collects the quantitative results reported by the surveyed studies. The numbers are not directly comparable, because tasks, datasets, class definitions and hardware differ, and the table is intended to show the range of reported performance and the caveats that accompany each figure rather than to rank the methods. Two observations nevertheless stand out. Reported accuracies and F1-scores cluster above 0.9 across very different tasks, which suggests that the limiting factor for portable deployment is not raw classification performance on curated data but representation, generalisation and resource cost. In addition, the studies with the smallest evaluation populations report some of the highest accuracies, which reinforces the need for larger, multi-environment evaluations.

Table 11
Quantitative results reported by the surveyed studies and caveats for comparison (values as reported by the original authors).

Study	Task	Evaluation data	Metric	Reported value	Comparability caveat
Lundqvist et al. [22]	Multiclass IDS	TON_IoT	F1; model size	0.976; 31 KB	Flow features; single dataset
De Bernardis et al. [24]	Binary and 10-class IDS	TON_IoT	F1 (LightGBM INT8); size	0.9992; 73.85 KB	Large shift to CICIoT2023 reported
Mazinani et al. [23]	2-, 8-, 34-class IDS	CICIoT2023	Size reduction after INT8 PTQ	> 90%	Study-specific trade-off score
Shafi and Lashkari [20]	IoT intrusion detection	BCCC-IoT-IDS-Zwave-2025	Average detection accuracy	> 95%	Accuracy on a new, multi-source dataset
Rehman et al. [14]	Threat detection for NFC-enabled IoT	UNSW-NB15, Bot-IoT, TON_IoT, Edge-IIoTset	Accuracy; F1	Up to 95%; > 0.90	No NFC radio captures

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Table 11
Quantitative results reported by the surveyed studies and caveats for comparison (values as reported by the original authors).

Study	Task	Evaluation data	Metric	Reported value	Comparability caveat
Khan et al. [15]	Chipless RFID clone detection	22 tags (1 authentic)	Accuracy	98.41% (FD); 92.07% (TD)	Small population; similarity matching
Wu et al. [38]	RFID tag classification	7 tag types, 3 manufacturers	Accuracy gain	≈ 11% (risk control)	Requires distance adjustment
Owusu et al. [27]	LoRa device identification	30 commercial devices	Accuracy; energy	94.54%; 17% lower energy	Single protocol

Reproducibility. Public datasets and analysis tools, as released for BCCC-IoT-IDS-Zwave-2025 [20], CICIoT2023 [19] and TON_IoT [18], make IDS results reproducible in principle, and hardware-in-the-loop verification [24] makes edge results reproducible on commodity boards. In contrast, platform and RF fingerprinting studies depend on physical devices and radio conditions that are difficult to replicate, and few publish raw captures. Standardised capture formats and shared radio datasets would therefore benefit these two streams most.

6. RESEARCH GAPS, CHALLENGES AND FUTURE DIRECTIONS

The comparative analysis shows that the field has produced the components of an analysis-capable portable wireless security platform, but not their integration. Fig. 6 summarises the situation as three research silos, namely protocol-specific security, ML-based IoT detection and edge deployment, whose intersection is largely unpopulated. This section consolidates the gaps and challenges that follow from the review and outlines directions for future research. Table 12 maps each gap to the corresponding challenge and direction. It should be emphasised that the directions below are proposals derived from the reviewed evidence, not established results.

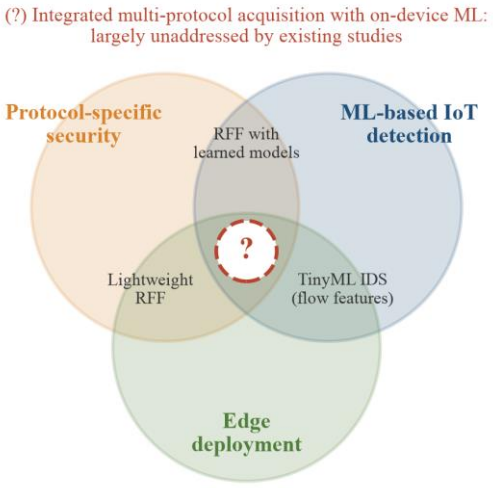


Fig. 6. Research gap: the three research silos and their largely unaddressed intersection.

6.1. Research gaps

Gap 1: Absence of integrated acquisition-to-analysis pipelines. Existing portable devices interact with many protocols but leave interpretation to the operator [8,9], while automated detectors operate on a single protocol family [20,36]. No reviewed study reports on-device ML detection over observations captured by a multi-protocol device.

Gap 2: Lack of multi-protocol radio datasets. Public datasets are overwhelmingly IP-centric; BCCC-IoT-IDS-Zwave-2025 is a notable step towards multi-source data [20], and AWID covers 802.11 [33], but no dataset provides labelled Sub-GHz, IR, NFC/RFID and BLE observations under a common schema with benign and adversarial activity.

Gap 3: Mismatch between protocol claims and evaluation data. Protocol-specific detectors are sometimes validated on network datasets that do not contain the protocol in question [14], and laboratory fingerprinting results are rarely tested under the channel variability of handheld operation [25].

Gap 4: Edge evaluation limited to network flows. Edge-ML studies provide rigorous hardware measurements, but only for flow features [22–24]; the cost of signal-level feature extraction and of running several protocol pipelines concurrently on portable hardware has not been characterised.

Gap 5: Generalisation and drift. Cross-dataset shifts are large [24] and closed-world evaluation overstates performance [43]; for a device carried between environments, distribution shift is the normal operating condition rather than an exception.

6.2. Open challenges

Heterogeneous representation. Sub-GHz pulse trains, IR command frames, NFC anticollision exchanges, 802.11 management frames and BLE advertisements differ in time scale, structure and semantics. Designing a normalised record that preserves protocol-specific information while exposing comparable signal, timing, protocol-statistics and device or session features is a non-trivial representation problem.

Label scarcity and ethics of data collection. Labelled attack data for radio protocols can only be generated by performing attacks, which must be confined to devices that researchers own or are authorised to test [11]. This limits dataset scale and diversity and favours unsupervised and few-shot methods [40,53].

Resource contention. On portable hardware, acquisition, decoding, feature extraction, inference, display and logging

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

compete for the same processor, memory and battery. Energy-aware model selection [27] and integer-only inference [45] mitigate the cost of inference, but system-level scheduling remains unexplored.

False alarms and interpretability. A field detector must justify its alarms to an operator. Trees offer partial interpretability, and interpretable spectrum features have been proposed [42], but explanation mechanisms for multi-protocol detectors have not been studied.

Dual-use risk. The same capabilities that support assessment can support misuse, as the reception of Flipper Zero illustrates [8]. Research platforms therefore need built-in safeguards such as authorised-use workflows, audit logging of every observation and event, and restriction of active transmission features.

Table 12
Mapping of research gaps to open challenges and future research directions.

Research gap	Evidence	Open challenge	Future research direction
G1. No integrated acquisition-to-analysis pipeline	[8,9,36]	Heterogeneous representation; resource contention	Open, modular SBC platforms with a unified pipeline and on-device ML
G2. No multi-protocol radio datasets	[20,33]	Label scarcity; authorised data collection	Common observation schema; open datasets collected under reproducible, authorised protocols
G3. Protocol claims validated on unrelated data	[14,25]	Ground truth; channel variability of handheld use	Protocol-native captures; field evaluation of fingerprinting
G4. Edge evaluation limited to flow features	[22–24]	Concurrent pipelines on battery-powered hardware	End-to-end hardware benchmarking of complete pipelines
G5. Poor generalisation and drift	[24,43]	Distribution shift; false alarms	Unsupervised, few-shot, federated and energy-aware adaptive models
G6. Dual-use risk of capable devices	[8,11]	Preventing misuse without hindering research	Authorised-use workflows, audit logging and restricted transmission by design

6.3. Future research directions

Unified wireless observation schema and open multi-protocol datasets. A first direction is the definition of a protocol-agnostic observation record, containing a timestamp, interface, decoded protocol, normalised content and a feature vector of signal, timing, protocol-statistics and device or session features, together with community datasets collected under authorised conditions following reproducible protocols such as that of [11]. Multi-source collection as in [20] could be extended from Z-Wave and MQTT to Sub-GHz, IR, NFC/RFID and BLE.

Hybrid physical-layer and protocol-layer detection. Combining RF fingerprints [15,26,27] with protocol-level features would allow a device to detect both forged transmitters and protocol abuse. Late fusion of per-protocol detectors and shared representation learning across protocols are both open options.

End-to-end edge benchmarking. The hardware-in-the-loop methodology of [24] and [23] should be extended from single models to complete pipelines on SBC-class and microcontroller-class portable devices, reporting accuracy, precision, recall, F1-score, false positive rate, end-to-end latency, RAM and CPU utilisation and energy, ideally with standardised procedures in the spirit of MLPerf Tiny [44]. Such benchmarks should separate the cost of acquisition and decoding, feature extraction and inference, since on SBC-class hardware the first two stages may dominate, and should report results per protocol as well as in aggregate.

Adaptive, label-efficient and collaborative learning. Unsupervised and one-class models [40,41], few-shot learning [53] and federated updates across devices can reduce dependence on labelled attacks and support adaptation to new environments, while energy-aware model switching [27] can balance accuracy against battery life.

Hierarchical and explainable detection. A practical portable detector could combine inexpensive rule-based checks, such as known replay patterns or deauthentication rates, with lightweight learned models that are invoked only when rules are inconclusive, reducing average latency and energy. Each alarm should carry the features that triggered it and a confidence value, so that the operator can judge it, and every observation and decision should be logged with a timestamp to support later evaluation and audit [42,43].

Open, modular and responsible platforms. Open SBC-based designs, in which interface modules, feature extractors and models can be replaced independently, would turn portable testing devices into research instruments. One example of this direction is the authors' ongoing work on a Raspberry Pi based device inspired by Flipper Zero, which combines modular RF/Sub-GHz, IR, NFC/RFID and Wi-Fi/BLE interfaces with a lightweight on-device anomaly detector; such designs must embed authorised-use controls and complete audit logging from the outset.

7. CONCLUSION

This survey has reviewed the research relevant to portable, multi-protocol security testing and IoT interaction devices that perform automated analysis at the edge. The literature was organised into five streams: portable multi-protocol platforms, protocol-specific security, RF fingerprinting and physical-layer authentication, IoT intrusion detection datasets and ML-based detection, and lightweight edge ML. For each stream,

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

the contributions, methods and limitations of representative studies were compared, and the streams were then compared with one another across acquisition, representation, analysis and deployment.

The review finds that each stream is individually mature. Portable devices such as Flipper Zero make multi-protocol interaction practical; protocol studies continue to uncover weaknesses in BLE, NFC, RFID, Sub-GHz, IR and Wi-Fi systems; RF fingerprinting shows that signal-level features reveal counterfeit and forged transmitters; new datasets broaden the scope of IoT detection towards multi-source data; and TinyML studies demonstrate that accurate detectors can run within kilobytes of memory and microseconds to milliseconds of latency. However, the streams remain separate. No reviewed study combines multi-protocol acquisition with on-device ML analysis, public datasets do not cover the radio protocols handled by portable devices, and edge evaluations have been limited to network-flow features.

Bridging these silos requires a common observation schema, authorised multi-protocol datasets, hybrid physical- and protocol-layer detection, end-to-end edge benchmarking, label-efficient and adaptive learning, and open platforms with built-in safeguards against misuse. Progress on these directions would transform portable wireless tools from interactive instruments that depend on the expertise of their operators into research platforms that produce measurable, reproducible security evidence.

ACKNOWLEDGEMENT

The authors thank Ms. Roshmin Rizuvan, Assistant Professor, Department of Computer Science and Engineering, Royal College of Engineering & Technology, Akkikavu, for her guidance, and the Department of Computer Science and Engineering (Cyber Security) for its support.

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.

Data availability

No data was used for the research described in the article.

REFERENCES

[1] Garcia FD, Oswald D, Kasper T, Pavlidès P. Lock it and still lose it – On the (in)security of automotive remote keyless entry systems. In: 25th USENIX security symposium. 2016, p. 929–44.

[2] Garcia FD, de Koning Gans G, Muijrers R, van Rossum P, Verdult R, Schreur RW, et al. Dismantling MIFARE classic. In: Computer security – ESORICS 2008. Lecture notes in computer science, Berlin: Springer; 2008, p. 97–114. http://dx.doi.org/10.1007/978-3-540-88313-5_7.

[3] Vanhoef M, Piessens F. Key reinstallation attacks: Forcing nonce reuse in WPA2. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security. 2017, p. 1313–28. http://dx.doi.org/10.1145/3133956.3134027.

[4] Celosia G, Cunche M. Saving private addresses: An analysis of privacy issues in the Bluetooth-Low-Energy advertising mechanism. In: Proceedings of the 16th EAI international conference on mobile and ubiquitous systems: computing, networking and services. 2019, p. 444–53. http://dx.doi.org/10.1145/3360774.3360777.

[5] Kolias C, Kambourakis G, Stavrou A, Voas J. DDoS in the IoT: Mirai and other botnets. Computer 2017;50(7):80–4. http://dx.doi.org/10.1109/MC.2017.201.

[6] Al-Garadi MA, Mohamed A, Al-Ali AK, Du X, Ali I, Guizani M. A survey of machine and deep learning methods for Internet of Things (IoT) security. IEEE Commun Surv Tutor 2020;22(3):1646–85. http://dx.doi.org/10.1109/COMST.2020.2988293.

[7] Flipper Devices Inc. Flipper Zero technical specifications. 2026, https://docs.flipper.net/zero/development/hardware/tech-specs [Accessed 11 November 2026].

[8] Cass S. A hacker's delight: You'll either love or hate the Flipper Zero. IEEE Spectr 2023;60(5):18–20. http://dx.doi.org/10.1109/MSPEC.2023.10120663.

[9] Muñoz-Ruiz FJ, Di-Bartolo AJ, Broncano-Morgado F, Ramírez-Gabardino BM, Ávila M. Exploring the real capabilities of the Flipper Zero. Eng Proc 2026;123:6. http://dx.doi.org/10.3390/engproc2026123006.

[10] Drăghici BG, Dobre AE, Jibotean M-A, Ionică D, Stan OP, Miclea LC. Assessment of deauthentication threats in IoT: An empirical analysis of ESP8266 Deauther and Flipper Zero. In: 2025 29th international conference on system theory, control and computing. ICSTCC, 2025, p. 661–6. http://dx.doi.org/10.1109/ICSTCC66753.2025.11240388.

[11] Garrido Briones C, Bravo Lara O, Rehl Muñoz J. Distance-dependent replay results on authorized RFID and sub-GHz wireless access-control testbeds: A reproducible educational red-teaming protocol and pilot study with Flipper Zero. In: Communications in computer and information science. Cham: Springer; 2026, p. 39–60. http://dx.doi.org/10.1007/978-3-032-42785-4_3.

[12] Alamleh H, AlQahtani AAS, Gogarty M, Ruddel D. Securing the invisible thread: a comprehensive analysis of BLE tracker security in Apple AirTags and Samsung SmartTags. J Umm Al-Qura Univ Eng Archit 2025;17(3):540–58. http://dx.doi.org/10.1007/s43995-025-00248-4.

[13] Heinrich A, Stute M, Kornhuber T, Hollick M. Who can Find My devices? Security and privacy of Apple's crowd-sourced Bluetooth location tracking system. Proc Priv Enhancing Technol 2021;2021(3):227–45. http://dx.doi.org/10.2478/popets-2021-0045.

[14] Rehman A, Alharbi O, Qasaymeh Y, Aljaedi A. DC-NFC: A custom deep learning framework for security and privacy in NFC-enabled IoT. Sensors 2025;25(5):1381. http://dx.doi.org/10.3390/s25051381.

[15] Khan S, Ray B, Karmakar N. Fingerprinting chipless RFID with a MIMO system for tag authentication in Internet of Things. Internet Things 2025;31:101542. http://dx.doi.org/10.1016/j.iot.2025.101542.

[16] Meidan Y, Bohadana M, Mathov Y, Mirsky Y, Shabtai A, Breitenbacher D, et al. N-BaIoT—Network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Comput 2018;17(3):12–22. http://dx.doi.org/10.1109/MPRV.2018.03367731.

[17] Koroniotis N, Moustafa N, Sitnikova E, Turnbull B. Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. Future Gener Comput Syst 2019;100:779–96. http://dx.doi.org/10.1016/j.future.2019.05.041.

[18] Moustafa N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. Sustainable Cities Soc 2021;72:102994. http://dx.doi.org/10.1016/j.scs.2021.102994.

[19] Neto ECP, Dadkhah S, Ferreira R, Zohourian A, Lu R, Ghorbani AA. CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. Sensors 2023;23(13):5941. http://dx.doi.org/10.3390/s23135941.

[20] Shafi M, Lashkari AH. Toward generating a large-scale IoT-Zwave intrusion detection dataset: Smart device profiling, intruders behavior, and traffic characterization. Internet Things 2025;34:101747. http://dx.doi.org/10.1016/j.iot.2025.101747.

[21] Ray PP. A review on TinyML: State-of-the-art and prospects. J King Saud Univ Comput Inf Sci 2022;34(4):1595–623. http://dx.doi.org/10.1016/j.jksuci.2021.11.019.

[22] Lundqvist J, Hadzic A, Kirkeluten TM, Pedersen H, Holth J, Johansson MH, et al. Lightweight machine learning models for intrusion detection on IoT devices. In: Norsk IKT-konferanse for forskning og utdanning (NIKT), vol. 37, no. 3. 2025. http://dx.doi.org/10.5324/jrxdjb92.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

[23] Mazinani A, Antonucci D, Davoli L, Ferrari G. Performance assessment of DL for network intrusion detection on a constrained IoT device. Future Internet 2026;18(1):34. http://dx.doi.org/10.3390/fi18010034.

[24] De Bernardis EP, Kuznetsov O, Arnesano M, Zhansaya P, Sydykova M. Lightweight machine learning intrusion detection for IoT/IIoT networks: Quantisation strategies and physical deployment on resource-constrained microcontrollers. Electronics 2026;15(13):2869. http://dx.doi.org/10.3390/electronics15132869.

[25] Jagannath A, Jagannath J, Kumar PSPV. A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges. Comput Netw 2022;219:109455. http://dx.doi.org/10.1016/j.comnet.2022.109455.

[26] Adigopula S, Subramanyam MV. DCNN-RFF-NFC: a novel design of NFC security using deep convolution neural network-based RF fingerprinting. Neural Comput Appl 2025;37(6):4439–53. http://dx.doi.org/10.1007/s00521-024-10861-4.

[27] Owusu EO, Iddrisu D, Klogo GS, Boateng KO, Akowuah EK. Energy-aware RF fingerprinting for device identification in ultra-low-power IoT systems. Eng Rep 2025;7(7):e70293. http://dx.doi.org/10.1002/eng2.70293.

[28] Barua A, Al Alamin MA, Hossain MS, Hossain E. Security and privacy threats for Bluetooth low energy in IoT and wearable devices: A comprehensive survey. IEEE Open J Commun Soc 2022;3:251–81. http://dx.doi.org/10.1109/OJCOMS.2022.3149732.

[29] Kumar A, Jain AK, Dua M. A comprehensive taxonomy of security and privacy issues in RFID. Complex Intell Syst 2021;7(3):1327–47. http://dx.doi.org/10.1007/s40747-021-00280-6.

[30] Wouters L, Marin E, Ashur T, Gierlichs B, Preneel B. Fast, furious and insecure: Passive keyless entry and start systems in modern supercars. IACR Trans Cryptogr Hardw Embed Syst 2019;2019(3):66–85. http://dx.doi.org/10.46586/tches.v2019.i3.66-85.

[31] Joo K, Choi W, Lee DH. Hold the door! Fingerprinting your car key to prevent keyless entry car theft. In: Proceedings 2020 network and distributed system security symposium. NDSS, 2020. http://dx.doi.org/10.14722/ndss.2020.23107.

[32] Guri M, Bykhovsky D. aIR-Jumper: Covert air-gap exfiltration/infiltration via security cameras & infrared (IR). Comput Secur 2019;82:15–29. http://dx.doi.org/10.1016/j.cose.2018.11.004.

[33] Kolias C, Kambourakis G, Stavrou A, Gritzalis S. Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. IEEE Commun Surv Tutor 2016;18(1):184–208. http://dx.doi.org/10.1109/COMST.2015.2402161.

[34] Antonioli D, Tippenhauer NO, Rasmussen K. BIAS: Bluetooth impersonation attacks. In: 2020 IEEE symposium on security and privacy. 2020, p. 549–62. http://dx.doi.org/10.1109/SP40000.2020.00093.

[35] Raspberry Pi Ltd. Raspberry Pi hardware documentation. 2026, https://www.raspberrypi.com/documentation/computers/raspberry-pi.html [Accessed 11 November 2026].

[36] Sforzin A, Mármol FG, Conti M, Bohli J-M. RPiDS: Raspberry Pi IDS — A fruitful intrusion detection system for IoT. In: 2016 Intl IEEE conferences on ubiquitous intelligence & computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people, and smart world congress. 2016, p. 440–8. http://dx.doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0080.

[37] Sankhe K, Belgiovine M, Zhou F, Riyaz S, Ioannidis S, Chowdhury K. ORACLE: Optimized radio classification through convolutional neural networks. In: IEEE INFOCOM 2019 - IEEE conference on computer communications. 2019, p. 370–8. http://dx.doi.org/10.1109/INFOCOM.2019.8737463.

[38] Wu H, Wang S, Pu C, Ma J, Zeng Y. Enhancing counterfeit RFID tag classification through distance based cognitive risk control. Sci Rep 2025;15:4150. http://dx.doi.org/10.1038/s41598-025-87809-8.

[39] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM Comput Surv 2009;41(3):1–58. http://dx.doi.org/10.1145/1541880.1541882.

[40] Mirsky Y, Doitshman T, Elovici Y, Shabtai A. Kitsune: An ensemble of autoencoders for online network intrusion detection. In: Proceedings 2018 network and distributed system security symposium. NDSS, 2018. http://dx.doi.org/10.14722/ndss.2018.23204.

[41] Liu FT, Ting KM, Zhou Z-H. Isolation forest. In: 2008 eighth IEEE international conference on data mining. 2008, p. 413–22. http://dx.doi.org/10.1109/ICDM.2008.17.

[42] Rajendran S, Meert W, Lenders V, Pollin S. SAIFE: Unsupervised wireless spectrum anomaly detection with interpretable features. In: 2018 IEEE international symposium on dynamic spectrum access networks. DySPAN, 2018, p. 1–9. http://dx.doi.org/10.1109/DySPAN.2018.8610471.

[43] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection. In: 2010 IEEE symposium on security and privacy. 2010, p. 305–16. http://dx.doi.org/10.1109/SP.2010.25.

[44] Banbury C, Janapa Reddi V, Torelli P, Holleman J, Jeffries N, Király C, et al. MLPerf tiny benchmark. 2021, arXiv preprint arXiv:2106.07597. http://dx.doi.org/10.48550/arXiv.2106.07597.

[45] Jacob B, Kligys S, Chen B, Zhu M, Tang M, Howard A, et al. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In: 2018 IEEE/CVF conference on computer vision and pattern recognition. 2018, p. 2704–13. http://dx.doi.org/10.1109/CVPR.2018.00286.

[46] Lee W, Baek SY. Deep neural network-aided radio frequency fingerprinting for identification of near field communication tags. Expert Syst Appl 2026;296:129016. http://dx.doi.org/10.1016/j.eswa.2025.129016.

[47] Shiltagh Al-Jamali NA, Zarzoor AR, Al-Raweshidy HS, Jawad Abbas TM. Hybrid model-based RF fingerprinting and spiking neural networks for IoT device classification. IET Wirel Sens Syst 2025;15(1):e70018. http://dx.doi.org/10.1049/wss2.70018.

[48] Yao X, Yang H, Zeng W. Lightweight physical-layer authentication for IoT devices access against jamming attacks. Phys Commun 2025;72:102787. http://dx.doi.org/10.1016/j.phycom.2025.102787.

[49] Ferrag MA, Friha O, Hamouda D, Maglaras L, Janicke H. Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning. IEEE Access 2022;10:40281–306. http://dx.doi.org/10.1109/ACCESS.2022.3165809.

[50] Shi W, Cao J, Zhang Q, Li Y, Xu L. Edge computing: Vision and challenges. IEEE Internet Things J 2016;3(5):637–46. http://dx.doi.org/10.1109/JIOT.2016.2579198.

[51] David R, Duke J, Jain A, Janapa Reddi V, Jeffries N, Li J, et al. TensorFlow lite micro: Embedded machine learning on TinyML systems. 2020, arXiv preprint arXiv:2010.08678. http://dx.doi.org/10.48550/arXiv.2010.08678.

[52] Hamdouchi A, Idri A. Evaluating the performance of TinyML singular and ensemble techniques for intrusion detection in IoT networks. Microprocess Microsyst 2025;117:105172. http://dx.doi.org/10.1016/j.micpro.2025.105172.

[53] Fusco P, Palmieri F, Ficco M. Combining epsilon-greedy reinforcement learning based gradient sparsification and siamese neural networks for few-shot federated tinyML intrusion detection in IoT. Internet Things 2025;34:101820. http://dx.doi.org/10.1016/j.iot.2025.101820.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)