sensors

MDPI

Article

Generating Datasets for Anomaly-Based Intrusion Detection Systems in IoT and Industrial IoT Networks

Ismael Essop [1], José C. Ribeiro [2,*], Maria Papaioannou [1,2], Georgios Zachos [1,2], Georgios Mantas [1,2] and Jonathan Rodriguez [2,3]

[1] Faculty of Engineering and Science, University of Greenwich, Chatham Maritime ME4 4TB, UK; i.a.essop@gre.ac.uk (I.E.); m.papaioannou@av.it.pt (M.P.); g.zachos@gre.ac.uk (G.Z.); gimantas@av.it.pt (G.M.)
[2] Instituto de Telecomunicações, Aveiro 3810-193, Portugal; jonathan@av.it.pt
[3] Faculty of Computing, Engineering and Science, University of South Wales, Pontypridd CF37 1DL, UK
* Correspondence: jcarlosvgr@av.it.pt

Abstract: Over the past few years, we have witnessed the emergence of Internet of Things (IoT) and Industrial IoT networks that bring significant benefits to citizens, society, and industry. However, their heterogeneous and resource-constrained nature makes them vulnerable to a wide range of threats. Therefore, there is an urgent need for novel security mechanisms such as accurate and efficient anomaly-based intrusion detection systems (AIDSs) to be developed before these networks reach their full potential. Nevertheless, there is a lack of up-to-date, representative, and well-structured IoT/IIoT-specific datasets which are publicly available and constitute benchmark datasets for training and evaluating machine learning models used in AIDSs for IoT/IIoT networks. Contribution to filling this research gap is the main target of our recent research work and thus, we focus on the generation of new labelled IoT/IIoT-specific datasets by utilising the Cooja simulator. To the best of our knowledge, this is the first time that the Cooja simulator is used, in a systematic way, to generate comprehensive IoT/IIoT datasets. In this paper, we present the approach that we followed to generate an initial set of benign and malicious IoT/IIoT datasets. The generated IIoT-specific information was captured from the Contiki plugin "powertrace" and the Cooja tool "Radio messages".

Keywords: IoT; Industrial IoT; benign datasets generation; malicious datasets generation; Cooja simulator; Contiki OS; anomaly-based intrusion detection

check for updates

Citation: Essop, I.; Ribeiro, J.C.; Papaioannou, M.; Zachos, G.; Mantas, G.; Rodriguez, J. Generating Datasets for Anomaly-Based Intrusion Detection Systems in IoT and Industrial IoT Networks. *Sensors* **2021**, *21*, 1528. https://doi.org/10.3390/s21041528

Academic Editor: David Plets

Received: 8 January 2021
Accepted: 18 February 2021
Published: 23 February 2021

1. Introduction

Despite the significant benefits that IoT and Industrial IoT (IIoT) networks bring to citizens, society, and industry, the fact that these networks incorporate a wide range of different communication technologies (e.g., WLANs, Bluetooth, and Zigbee) and types of nodes/devices (e.g., temperature/humidity sensors), which are vulnerable to various types of security threats, raises many security and privacy challenges in IoT/IIoT-based systems. For instance, attackers may compromise IoT/IIoT networks in order to manipulate sensing data (e.g., by injecting fake data) and cause malfunction to the IoT/IIoT-based systems that rely on the compromised IoT/IIoT networks. It is worthwhile to mention that IoT/IIoT networks can become an attractive target of attackers with a wide spectrum of motivations ranging from criminal intents aimed at financial gain to industrial espionage and cyber-sabotage. Therefore, security solutions protecting IoT/IIoT networks from attackers are critical for the acceptance and wide adoption of such networks in the coming next years. Nevertheless, the high resource requirements of complex and heavyweight conventional security mechanisms cannot be afforded by (i) the resource-constrained IoT/IIoT nodes (e.g., sensors) with limited processing power, storage capacity, and battery life; and/or (ii) the constrained environment in which the nodes are deployed and interconnected using lightweight communication protocols. Consequently, there is an urgent need for novel security mechanisms, such as accurate and efficient anomaly-based intrusion detection systems

(AIDSs) tailored to the resource-constrained characteristics of IoT/IIoT networks, to be developed in order to address the pressing security challenges of IoT/IIoT networks with reasonable cost, in terms of processing and energy, before IoT/IIoT networks gain the trust of all involved stakeholders and reach their full potential in the market [1–3]. However, there is a lack of up-to-date, representative and well-structured IoT/IIoT-specific datasets that are publicly available to the research community and constitute benchmark datasets for training and evaluating machine learning (ML) models used in AIDSs for IoT/IIoT networks [4,5]. This lack of benchmark IoT/IIoT datasets constitutes a significant research gap that should be addressed in order to develop more accurate and efficient IoT/IIoT-specific AIDS whose effectiveness is evaluated based on their performance to detect IoT/IIoT attacks which is a process reliant on comprehensive IoT/IIoT-specific datasets.

In fact, although several datasets, such as KDDCUP99 [6], NSL-KDD [7], UNSW-NB15 [8], and CICD2017 [9] have been created over the past two decades for evaluation purposes of network-based intrusion detection systems (IDSs), they do not include any specific characteristics of IoT/IIoT networks as these datasets do not contain sensors' reading data or IoT/IIoT network traffic [4,5]. To respond to this major issue, few efforts focused on the generation of IoT-specific datasets have also been seen in the literature recently. However, they are characterised by some limitations in terms of the IoT-specific information they include. For instance, the datasets proposed in [10,11] are IoT-specific datasets but they lack of events reflecting attack scenarios. To address this limitation, the IoT-specific and network-related datasets proposed in [12,13] contain events reflecting attack scenarios; however, they do not cover a diverse set of attack scenarios and do not include sensors' reading data or information related to the behaviour of the IoT/IIoT devices (e.g., sensors/actuators) within the network. Therefore, these IoT datasets can mainly be used for detecting only a limited number of network-based attacks against IoT/IIoT networks as they do not contain adequate information for detecting a wide range of network-based attacks and/or attacks that manipulate sensor measurement data or compromise IoT/IIoT devices within the IoT/IIoT network.

Consequently, there is an urgent need for comprehensive IoT/IIoT-specific datasets containing not only network-related information (e.g., packet-level information and flow-level information) but also events reflecting multiple benign and attack scenarios from current IoT/IIoT network environments, sensor measurement data, and information related to the behaviour of the IoT/IIoT devices deployed within the IoT/IIoT network for efficient and effective training and evaluation of AIDSs suitable for IoT/IIoT networks. Towards this direction, the recent work of [4] has proposed, for the first time, to the best of our knowledge, a new dataset that includes events of a variety of IoT-related attacks and legitimate scenarios, IoT telemetry data collected from heterogeneous IoT/IIoT data sources, network traffic of IoT/IIoT network, and audit traces of operating systems [4]. Therefore, it is clear that more comprehensive IoT/IIoT-specific datasets including events reflecting multiple benign and attack scenarios, sensor measurement data, network-related information, and information related to the behaviour of the IoT/IIoT devices are required to be generated and become publicly available to the research community so as to fill this significant research gap of lack of benchmark IoT/IIoT datasets and more accurate and efficient IoT/IIoT-specific AIDS to be developed.

Contribution to filling this research gap is the main target of our recent research work. In particular, our focus is on the generation of new labelled IoT/IIoT datasets that will be publicly available to the research community and include: (a) events reflecting multiple benign and attack scenarios from current IoT/IIoT network environments, (b) sensor measurement data, (c) network-related information (e.g., packet-level information and flow-level information) from the IoT/IIoT network, and (d) information related to the behaviour of the IoT/IIoT devices deployed within the IoT/IIoT network. It is worthwhile to mention that the new labelled IoT/IIoT datasets are generated by implementing various benign IoT/IIoT network scenarios and IoT/IIoT network attack scenarios in the Cooja simulator which is the companion network simulator of the open source Contiki Operating

System (OS) that is one of the most popular OSs for resource constrained IoT devices [14]. To the best of our knowledge, this is the first time that the Cooja simulator is going to be used, in a systematic way, to generate comprehensive IoT/IIoT datasets. In this paper, we present the approach that we followed to generate an initial set of benign IoT/IIoT datasets (i.e., including only normal events) and malicious IoT/IIoT datasets (i.e., including attack and normal events) by utilising the Cooja simulator that was the simulation environment where the corresponding benign and attack scenarios were implemented.

The rest of this paper is organised as follows. In Section 2, the main threats against the IoT/IIoT network (i.e., perception domain) are presented and in Section 3, examples of anomaly-based intrusion detection systems for IoT/IIoT networks are discussed. In Section 4, a detailed description of the approach followed to generate a set of benign datasets by implementing a benign IIoT network scenario in the Cooja simulator is provided. In Section 5, a detailed description of the approach followed to generate a set of malicious datasets by implementing a User Datagram Protocol (UDP) flooding attack scenario in the Cooja is provided as well. In Section 6, a discussion on the generated datasets is given. Finally, Section 7 concludes this paper.

2. Threat Analysis of the IoT/IIoT Network (Perception Domain)

The perception domain, as shown in Figure 1, can be perceived as the device layer in the ITU-T reference model [15]. As the main purpose of the perception domain is to gather data, the security challenges in this domain target to forge collected IoT/IIoT data and damage perception devices, as presented below.

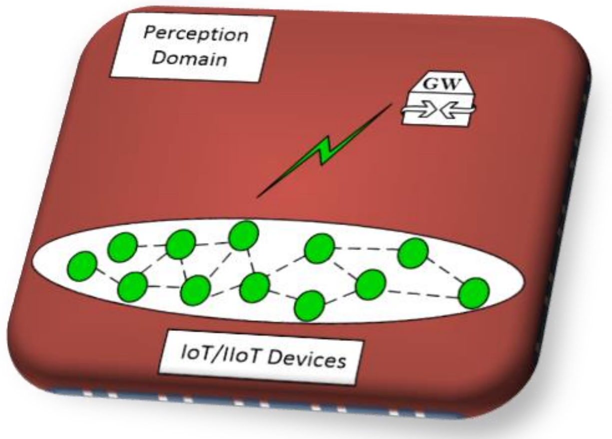

Figure 1. IoT/IIoT Network (Perception Domain).

2.1. Sinkhole Attacks

In this type of attacks, a compromised IoT/IIoT node (i.e., IoT/IIoT gateway [16]) in the perception domain proclaims very appealing capabilities of power, computation and communication [17] so that nearby nodes (i.e., IoT/IIoT sensors) will choose it as the forwarding node in the routing process due to its very attractive capabilities. As a consequence, the compromised IoT/IIoT node can increase the amount of data obtained before it is delivered to the cloud domain of the IoT-based monitoring system. Therefore, a sinkhole attack can not only compromise the confidentiality of the manufacturing data but also can comprise an initial step to launch additional attacks such as DoS/DDoS attacks [17,18].

2.2. Node Capture Attacks

In this type of attack, the adversary is able to extract important information about the captured node, such as the group communication key, radio key, etc. [17]. Additionally, the adversary can copy the important information related to the captured node to a malicious node, and afterwards fake the malicious node as a legitimate node to connect to the

IoT/IIoT network (i.e., perception domain). This type of attack is also known as node cloning/replication attack [17,19]. This attack may lead to compromising the security of the complete IoT/IIoT-based monitoring system.

2.3. Malicious Code Injection Attacks

An attacker can take control of an IoT/IIoT node or device in the perception domain by exploiting its security vulnerabilities in software and hardware and injecting malicious code into its memory. Afterwards, using the malicious code, the attacker can force the node or device to perform unintended operations. For example, the infected IoT/IIoT node(s) or device(s) can be used as a bot(s) to launch further attacks (e.g., DoS and DDoS) against other devices or nodes within the perception domain or even against the other domains (i.e., Network domain and Cloud domain). In addition, the attacker can use the injected malicious code in the infected device or node to get access into the IoT/IIoT-based system and/or get full control of the system [19].

2.4. False Data Injection Attacks

After capturing an IoT/IIoT node or device in the perception domain, the adversary can inject false data in place of benign data measured by the captured IoT/IIoT node or device and transmit the false data to the Cloud domain [17]. Thereafter, receiving the false data, the IoT/IIoT-based system may provide wrong services, which further negatively impacts the effectiveness of system itself.

2.5. Replay Attacks

In the perception domain, the attacker can use a malicious IoT/IIoT node or device to transmit to the destination host (i.e., IoT/IIoT gateway) with legitimate identification information, already received by the destination host, so that the malicious node or device can become a trusted node/device to the destination host [17]. Replay attacks are commonly launched in authentication process to destroy the validity of certification.

2.6. Eavesdropping

As the IoT/IIoT nodes and devices in perception domain communicate via wireless networks, an attacker (i.e., eavesdropper) can retrieve sensitive manufacturing data by overhearing the wireless transmission. For instance, an adversary within the perception domain can eavesdrop exchanged information by tracking wireless communications and reading the contents of the transmitted packages [17]. The eavesdropper can passively intercept the wireless communication between a sensor (e.g., environment industrial sensors or sensors on the machine resources) and the IoT/IIoT gateway, and extract confidential data (e.g., through traffic analysis) in order to maliciously use them.

2.7. Sleep Deprivation Attacks or Denial of Sleep Attacks

These attacks target to drain the battery of the resource constrained IoT/IIoT devices of the perception domain. In principle, the IoT/IIoT devices in the perception domain are usually programmed to follow a sleep routine when they are inactive in order to reduce the power consumption and extend their life cycle. However, an adversary may break the programmed sleep routines and keep the IoT/IIoT devices of the perception domain continuously active until they are shut down due to a drained battery. Attackers can achieve this by running infinite loops in these devices using malicious code or by artificially increasing their power consumption [20].

2.8. Sybil Attacks

In a sybil attack, a malicious or sybil node or device can illegitimately claim multiple identities, allowing it to impersonate them within the perception domain. For instance, the malicious node can achieve to connect with several other devices in order to maximise its influence and even deceive the complete system to draw incorrect conclusions [21].

2.9. Denial of Service (DoS) Attacks

The main target of these attacks is to deplete resources of the perception domain in order to make the whole IoT/IIoT network or specific nodes (e.g., machine or/and environment resources) or devices (e.g., IoT/IIoT gateway) unavailable. For instance, jamming attacks are a type of DoS attacks where an attacker transmits a high-range signal to overload the communication channel between two communicating entities and disrupt their communication. Within the perception domain of the IoT/IIoT-based system, jamming attacks can disrupt the communication between the IoT/IIoT sensors and the Gateway in order to prevent data from being transmitted to the Gateway, leading to malfunctions in the provided services to the authorised users. Jamming attacks can be performed by passively listening to the wireless medium so as to broadcast on the same frequency band as the legitimate transmitting signal. Finally, distributed denial of service (DDoS) attacks are a large-scale variant of DoS attacks and in the case of the perception domain an example of DDoS attack is when a large number of nodes (e.g., IoT/IIoT sensors) are compromised so as to flood the Gateway with a lot of transmitted data/requests and render it unavailable or disrupt its normal operations [22,23].

3. Anomaly-Based Intrusion Detection Systems for IoT/IIoT Networks

In this Section, two examples of anomaly-based intrusion detection systems for IoT/IIoT networks are discussed. Moustafa et al. in [24] proposed an ensemble network intrusion detection technique which utilises established statistical flow features. The goal is to mitigate malicious events, and more specifically botnet attacks against DNS, HTTP and MQTT protocols that are employed in IoT networks. The first step of their work revolves around the deep analysis of the TCP/IP model and the subsequent extraction of a set of features from the network traffic protocols MQTT, HTTP, and DNS protocols. The Bro-IDS tool is used by the authors for basic features while they also employ, in parallel, their own extractor module to generate additional statistical features of the transactional flows. Consequently, features are filtered and only the most important ones are selected in order to simplify the NIDS and decrease its computational cost. In this step, the authors utilise the correlation coefficient on result features as a means of features selection. Lastly, an AdaBoost ensemble learning method is developed to detect the attacks. The method is based on the combination of three different Machine Learning (ML) algorithms; decision tree (DT), Naive Bayes (NB), and artificial neural network (ANN) algorithms. These classification techniques were chosen mainly due to the core entropy measure that was calculated from the feature vectors. The AdaBoost (Adaptive Boosting) method improves the performance of the detection in comparison to using each machine learning algorithm separately. In case of small differences of the feature vectors, an error function is employed. The importance of the error function lies in computing the error value for each instance of the distributed input data. Based on this error value, it is possible to understand and evaluate which learners are best suited to classify each instance. The experiments results show that the ensemble technique achieved a high detection rate (95.25%–99.86%) and a low false positive rate (between 0.01% and 0.72%) compared to existing state-of-the-art techniques. The authors employed the UNSWNB15 and NIMS botnet datasets with simulated IoT sensor data to support their findings.

Furthermore, a multi-layer perceptron (MLP), which is a type of supervised artificial neural network [25]), is used in an offline IDS for IoT networks [26]. The ANN consists of 3 layers and each of the hidden and output layers' neurons use a unipolar sigmoid transfer function to transform their input values to a specific output value. The network was trained using a stochastic learning algorithm with mean square error function. The training process included both feed-forward and backward training algorithms. To perform its task, the ANN analyses the Internet packet traces and attempts to detect DoS and DDoS attacks in IoT network. In order to evaluate the IoT IDS, an experimental architecture was created with four client nodes and a server relay node. The server node was subjected not only to DOS attacks from a single host with more than 10 million UDP packets sent but also to DDoS attacks from three hosts each sending over 10 million UDP packets at wire speed. The results of their simulations showed a detection accuracy of 99.4% and 0.6% false positive rate. The authors used a training dataset consisting of a total of 2313 samples, 496 of them deployed for validation and 496 of them for testing [5].

4. Generation of Benign IoT/IIoT Datasets

In this Section, we provide a detailed description of the approach followed to generate a set of benign datasets by implementing a benign IoT/IIoT network scenario in the Cooja simulator, as shown in Figure 2. The generated IoT/IIoT-specific information from the simulated scenario was captured from the Contiki plugin "powertrace" (i.e., features such as CPU consumption) and the Cooja tool "Radio messages" (i.e., network traffic features) in order to generate the "powertrace" dataset and the network traffic dataset for the simulated benign IoT/IIoT network scenario.

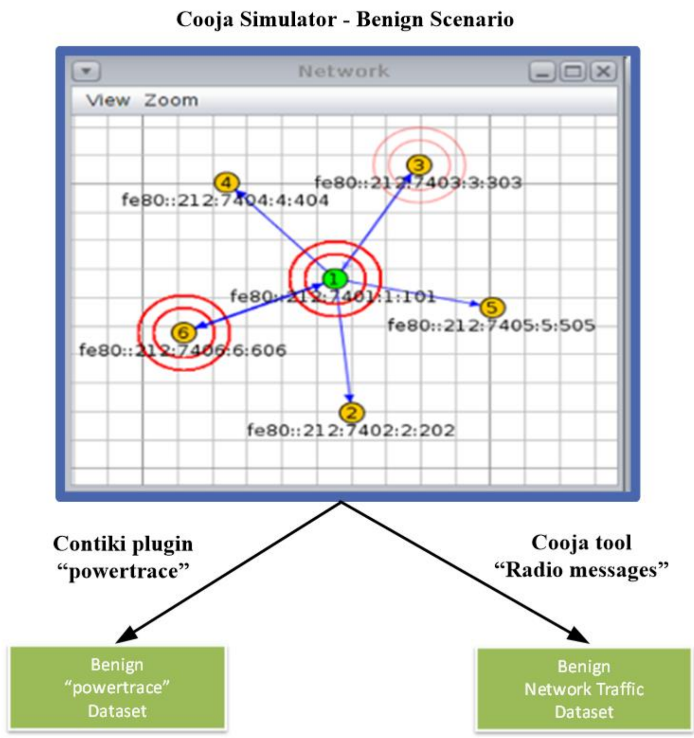

Figure 2. Benign datasets generation by utilizing the Cooja simulator.

The network topology of the simulated benign IoT/IIoT network scenario in the Cooja simulator environment consists of 5 yellow UDP-client motes (i.e., motes 2, 3, 4, 5 and 6) and the green UDP-server mote (i.e., mote 1), as depicted in Figure 2. The simulation duration was set to 60 min and the motes' outputs were printed out in the respective window (e.g., Mote output) while simulations run, as shown in Figure 3. In addition, the yellow UDP-client motes were configured to send text messages every 10 s, approximately, to the green UDP-sever mote that was configured to provide a corresponding response. The UDP protocol was used at the Transport Layer and the IPv6 at the network layer. Moreover, the type of motes used in this scenario was the Tmote Sky that is an ultra-low power wireless module for use in sensor networks, monitoring applications, and rapid application prototyping. In addition, Tmote Sky motes leverage industry standards such as USB and IEEE 802.15.4 to interoperate seamlessly with other devices. By using industry standards, integrating humidity, temperature, and light sensors, and providing flexible interconnection with peripherals, Tmote Sky motes enable several mesh network applications [27].

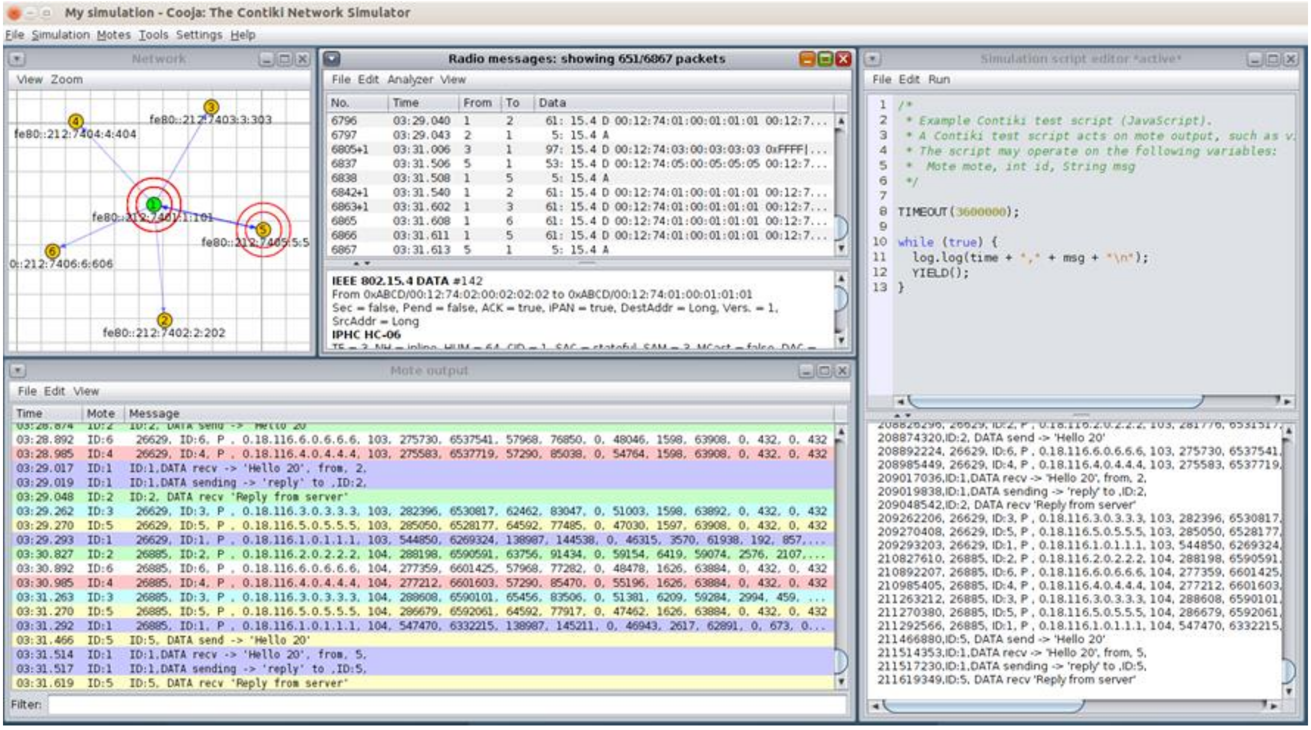

Figure 3. Cooja Simulator—motes' outputs.

4.1. Benign "Powertrace" Dataset Generation

4.1.1. Benign "Powertrace" Dataset Generation

The "powertrace" dataset includes information about features such as total CPU energy consumption and low power mode (LPM) energy consumption. In fact, it is the dataset of the simulated benign IIoT network scenario that includes records about information related to the energy consumption of the IIoT devices (i.e., motes) deployed within the simulated IIoT network. To enable the "powertrace" plugin and generate the "powertrace" dataset, we programmed the motes of the benign IIoT network to make use of the "powertrace" plugin for collecting "powertrace" related features every 2 s. In particular, we included the "powertrace.h" library into the code of each mote (i.e., #include "powertrace.h"), as shown in Figure 4, and defined to start powertracing, once every 2 s, in the code of each mote as shown in Figure 5.

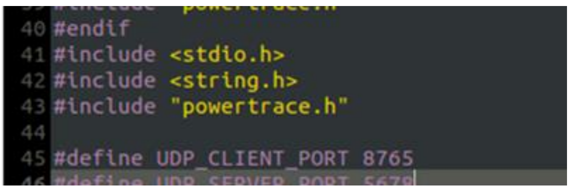

Figure 4. "powertrace.h" library in the mote code.

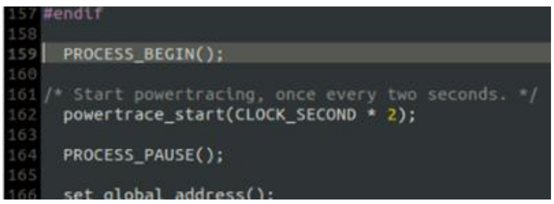

Figure 5. Powertracing Begin.

More precisely, the "powertrace" plugin captured raw information, every 2 s, about the set of features summarised in Table 1. In particular, the "powertrace" plugin tracks the duration (i.e., number of cpu ticks) of activities of a mote being in each power state. Particularly, the outputs demonstrate the fraction of time in which a mote remains for a given power state. There are the following six power states: (i) cpu; (ii) lpm; (iii) transmit; (iv) listen; (v) idle_transmit; and (vi) idle_listen, as shown in Table 1. These are measured with a hardware timer (i.e., clock frequency is defined in RTIMER_SECOND or 32,768 Hz for XM1000).

Table 1. "powertrace" plugin—Set of Captured Features.

Index	Feature	Description
1	sim time	simulation time
2	clock_time()	clock time (i.e.,by default, 128 ticks/second)
3	ID	Mote ID
4	P	label
5	rimeaddr	rime address
6	seqno	sequence number
7	all_cpu	accumulated CPU energy consumption
8	all_lpm	accumulated Low Power Mode energy consumption
9	all_transmit	accumulated transmission energy consumption
10	all_listen	accumulated listen energy consumption
11	all_idle_transmit	accumulated idle transmission energy consumption
12	all_idle_listen	accumulated idle listen energy consumption
13	cpu	CPU energy consumption for this cycle
14	lpm	LPM energy consumption for this cycle
15	transmit	transmission energy consumption for this cycle
16	listen	listen energy consumption for this cycle
17	idle_transmit	idle transmission energy consumption for this cycle
18	idle_listen	idle listen energy consumption for this cycle

In Figure 6, the depicted Mote output window displays the captured "powertrace" information every 2 s and also the messages sent and received by each mote (printouts/printf messages from each mote).

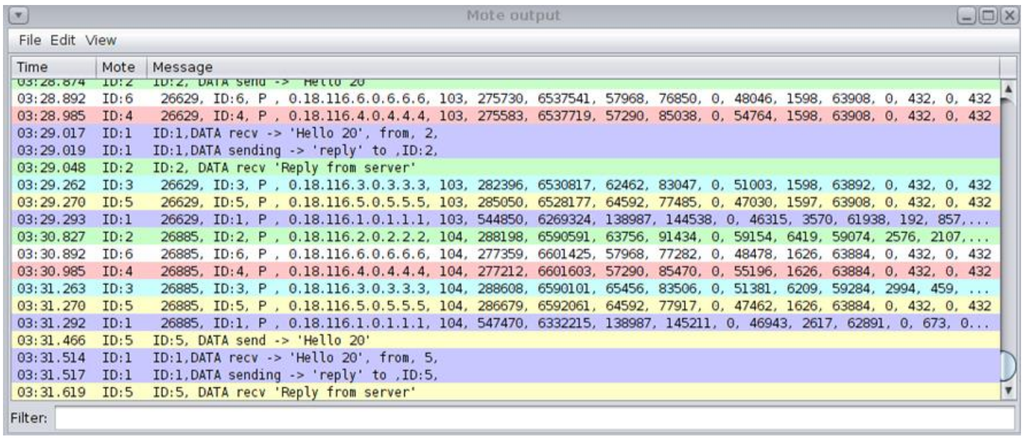

Figure 6. Cooja Simulator—Mote output window.

Furthermore, the Simulation script editor, shown in Figure 7, is a Cooja tool used to display messages and set a timer on the simulation. As shown in Figure 7, the upper part of the Simulation script editor was used to create scripts and the lower part to show the captured "powertrace" information and the printouts (i.e., printf messages) from the motes until the timeout occurs. In our implementation, we considered the simulation duration to be 60 min and thus, the timeout was set at 3,600,000 ms. When the timeout occurred, the simulation stopped, and all the captured information and prints were stored in the log file named "COOJA.testlog".

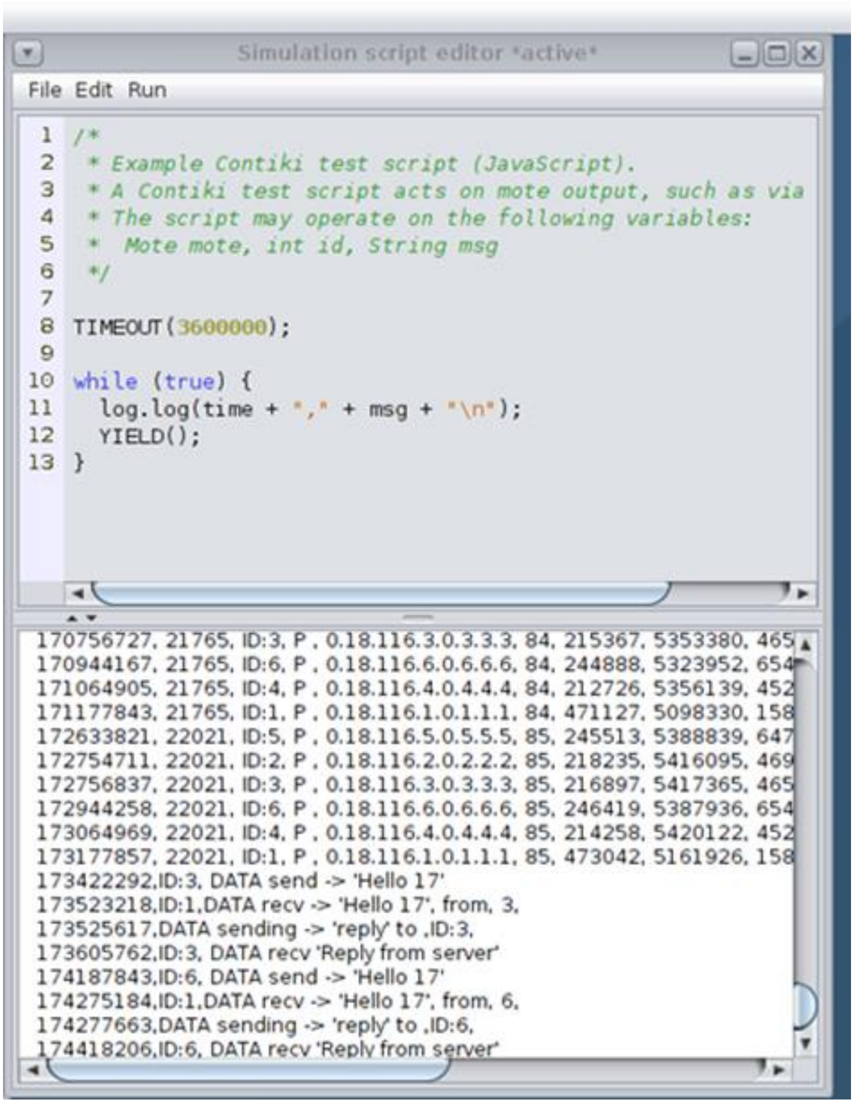

Figure 7. Simulation script editor.

Having collected all the captured raw information from the "powertrace" plugin in the "COOJA.testlog" file, the challenging task was to extract this information from the "COOJA.testlog" file to a csv file that would be the "powertrace" dataset of the simulated benign IIoT network scenario including records about the energy consumption of the motes. To address this challenge, we developed the "IoT_Simul.sh" bash file in order to extract all the required "powertrace" information from the "COOJA.testlog" file to the "pwrtrace.csv" file. An extract of the "IoT_Simul.sh" bash file is shown in Figure 8.

```
mkdir -p "$dir/$DATE/log"
mkdir -p "$dir/$DATE/dataset"
mkdir -p "$dir/$DATE/motedata"
mkdir -p "$dir/$DATE/nettraffic"

cd $cooja_home; ant run_nogui -Dargs=$dirtest/$TEST.csc |tee $dir/$DATE/log/COOJA.log

if [ -f "build/COOJA.testlog" ]; then
    mv build/COOJA.testlog $dir/$DATE/log/COOJA.testlog
fi

if [ -f "build/radiolog.pcap" ]; then
    mv build/radiolog.pcap $dir/$DATE/nettraffic/radiolog.pcap
fi

cd $dir/$DATE;

`echo ",,,,,,,Total measurements from the begining of the simulation,,,
,,Measurements for each of the 2-sec monitoring period" > dataset/pwrtrace.csv`
`echo "Real time,Simul time,Mote ID,,Rime Address,seq n,total_CPU_usage,
total_LPM_usage,total_transmitting_usage- transmission mode,total_listenig_usage-reception mode,
total_idle_transmitting,total_idle_reception,CPU_usage,LPM_usage,transmitting_usage - transmission mode,
listenig_usage-reception mode,idle_transmission,idle_recepction" >> dataset/pwrtrace.csv`
`echo "[us],(in ticks),,,,,(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),
(in ticks),(in ticks),(in ticks),(in ticks),(in ticks)">> dataset/pwrtrace.csv`

`grep " P " log/COOJA.testlog >> dataset/pwrtrace.csv`
`grep "DATA recv" log/COOJA.testlog >> dataset/recv.csv`
`grep "DATA send" log/COOJA.testlog >> dataset/send.csv`

 for n in `seq 1 6`;do

`echo ",,,,,,,Total measurements from the begining of the simulation,,,
,,Measurements for each of the 2-sec monitoring period" > motedata/mote$n.csv`
`echo "Real time,Simul time,Mote ID,,Rime Address,seq n,total_CPU_usage,total_LPM_usage,
total_transmitting_usage- transmission mode,total_listenig_usage-reception mode,total_idle_transmitting,
total_idle_reception,CPU_usage,LPM_usage,transmitting_usage - transmission mode,
listenig_usage-reception mode,idle_transmission,idle_recepction" >> motedata/mote$n.csv`
`echo "[us],(in ticks),,,,,(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),
(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),(in ticks),(in ticks)" >> motedata/mote$n.csv`

    `grep "ID:"$n dataset/pwrtrace.csv >> motedata/mote$n.csv`
done
```

Figure 8. Extract of the "IoT_Simul.sh" bash file.

Initially, the "IoT_Simul.sh" file created the root folder which was named with the simulation date and time (i.e., "2020-11-19-17-45-22" folder), as shown below in the left part of Figure 9. Afterwards, the bash file created the "log" folder, inside the "2020-11-19-17-45-22" folder, where the "COOJA.testlog" file was copied from the " ... /cooja/build" folder located in the Cooja Simulator environment.

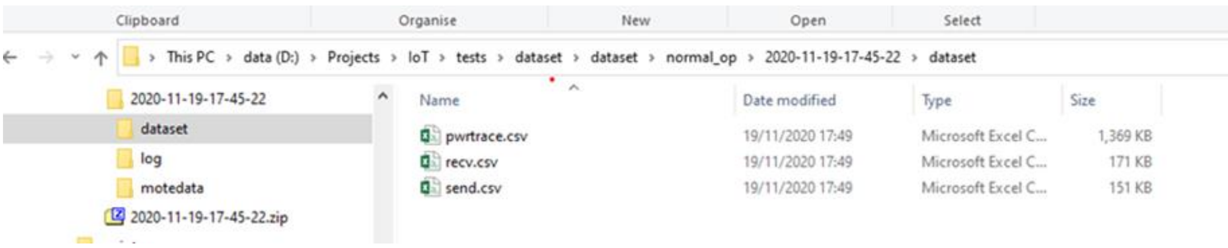

Figure 9. Location of the generated "pwrtrace.csv", "recv.csv", and "send.csv" files by the "IoT_Simul.sh" file.

In addition, in the "IoT_Simul.sh" file, we used the Linux tool "grep" in order to extract the required "powertrace" information by selecting the label "P" in each powertrace row (i.e., grep "P" log/COOJA.testlog >> dataset/pwrtrace.csv) from the "COOJA.testlog" file and save it in the "pwrtrace.csv" file in the "dataset" folder that was created by the batch file inside the "2020-11-19-17-45-22" folder, as shown in the left part of Figure 9. In the "dataset" folder, apart from the "pwrtrace.csv" file, the "IoT_Simul.sh" file generated two more files, based on the information included in the "COOJA.testlog" file, as shown in Figure 9; the "recv.csv" file and the "send.csv" file that include the "received" and "sent"messages printed by the motes, respectively.

Finally, the "IoT_Simul.sh" file extracted the information related to each mote, from the "pwrtrace.csv" file, and generated one csv file for each mote with the corresponding information from the "pwrtrace.csv" file. The generated 6 csv files (i.e., mote1.csv, mote2.csv, mote3.csv, mote4.csv, mote5.csv, mote6.csv) were stored in the "motedata" folder. The "motedata" folder was also created by the "IoT_Simul.sh" file inside the "2020-11-19-17-45-22" folder.

An overview of the above mentioned process followed to extract the required information from the "COOJA.testlog" file to the "pwrtrace.csv", "recv.csv", and "send.csv", "mote1.csv", "mote2.csv", "mote3.csv", "mote4.csv", "mote5.csv", and "mote6.csv" files are depicted in the Figure 10.

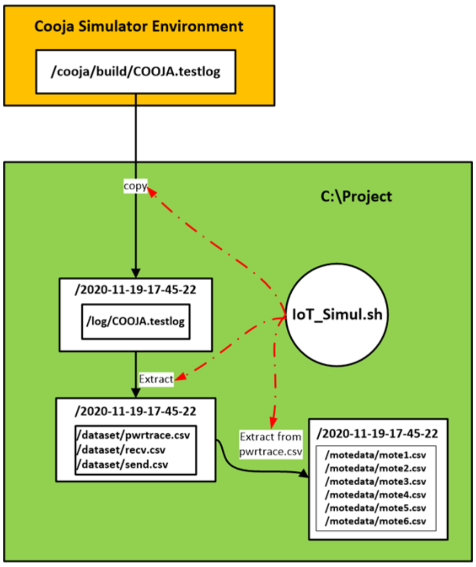

Figure 10. An overview of the process followed by the "IoT_Simul.sh" file to extract all the required "powertrace" information from the "COOJA.testlog" file.

4.1.2. Benign "Powertrace" Datasets—Results

Benign "pwrtrace.csv": The generated benign "pwrtrace.csv" file consists of 10,794 records and its first 38 records (i.e., 1–38) and its last 38 records (10,757–10,794) are depicted in Figures 11 and 12, respectively.

							Total measurements from the begining of the simulation						Measurements every 2 seconds (monitoring period)					
No	Real time [us]	clock_time (in ticks)	ID	P	rimeaddr	seqno	all_cpu (in ticks)	all_lpm (in ticks)	all_transmit (in ticks)	all_listen (in ticks)	all_idle_transmit (in ticks)	all_idle_listen (in ticks)	cpu (in ticks)	lpm (in ticks)	transmit (in ticks)	listen (in ticks)	idle_transmit (in ticks)	idle_listen (in ticks)
1	2587177	261	ID:6	P	0.18.116.6.0.6.6.6	0	6737	59719	2588	442	0	364	6737	59719	2588	442	0	364
2	2816245	261	ID:3	P	0.18.116.3.0.3.3.3	0	2184	64270	0	390	0	390	2184	64270	0	390	0	390
3	2907083	261	ID:1	P	0.18.116.1.0.1.1.1	0	2827	63628	0	1003	0	744	2827	63628	0	1003	0	744
4	3163478	261	ID:4	P	0.18.116.4.0.4.4.4	0	2184	64270	0	390	0	390	2184	64270	0	390	0	390
5	3183394	261	ID:5	P	0.18.116.5.0.5.5.5	0	6737	59719	2588	442	0	364	6737	59719	2588	442	0	364
6	3305496	261	ID:2	P	0.18.116.2.0.2.2.2	0	6737	59719	2588	442	0	364	6737	59719	2588	442	0	364
7	4586462	517	ID:6	P	0.18.116.6.0.6.6.6	1	7899	124068	2588	858	0	780	1159	64349	0	416	0	416
8	4821159	517	ID:3	P	0.18.116.3.0.3.3.3	1	3569	128521	0	1094	0	1043	1382	64251	0	704	0	653
9	4909515	517	ID:1	P	0.18.116.1.0.1.1.1	1	8583	123383	2980	1472	0	1134	5753	59755	2980	469	0	390
10	5164736	517	ID:4	P	0.18.116.4.0.4.4.4	1	3625	128346	0	1105	0	1056	1438	64076	0	715	0	666
11	5183527	517	ID:5	P	0.18.116.5.0.5.5.5	1	8248	123723	2588	1158	0	1030	1508	64004	0	716	0	666
12	5305285	517	ID:2	P	0.18.116.2.0.2.2.2	1	8250	123721	2588	1133	0	754	1510	64002	0	691	0	390
13	6587344	773	ID:6	P	0.18.116.6.0.6.6.6	2	9564	187918	2588	1525	0	1170	1662	63850	0	667	0	390
14	6817450	773	ID:3	P	0.18.116.3.0.3.3.3	2	4957	192526	0	1510	0	1459	1385	64005	0	416	0	416
15	6909795	773	ID:1	P	0.18.116.1.0.1.1.1	2	10071	187437	2980	2173	0	1537	1486	64054	0	701	0	403
16	7164686	773	ID:4	P	0.18.116.4.0.4.4.4	2	5014	192466	0	1521	0	1472	1386	64120	0	416	0	416
17	7184917	773	ID:5	P	0.18.116.5.0.5.5.5	2	14271	183210	5572	1630	0	1420	6020	59487	2984	472	0	390
18	7307013	773	ID:2	P	0.18.116.2.0.2.2.2	2	14255	183225	5565	1601	0	1144	6002	59504	2977	468	0	390
19	8588987	1029	ID:6	P	0.18.116.6.0.6.6.6	3	15557	247436	5573	1968	0	1534	5990	59518	2985	443	0	364
20	8820944	1029	ID:3	P	0.18.116.3.0.3.3.3	3	18623	244367	7942	4101	0	1823	13663	51841	7942	2591	0	364
21	8909413	1029	ID:1	P	0.18.116.1.0.1.1.1	3	13115	249855	2980	4355	0	2453	3041	62418	0	2182	0	916
22	9168183	1029	ID:4	P	0.18.116.4.0.4.4.4	3	18227	244754	7542	3922	0	1810	13210	52288	7542	2401	0	338
23	9185894	1029	ID:5	P	0.18.116.5.0.5.5.5	3	23353	239636	10452	4102	0	1784	9079	56426	4880	2472	0	364
24	9306227	1029	ID:2	P	0.18.116.2.0.2.2.2	3	15749	247241	5565	2017	0	1560	1491	64016	0	416	0	416
25	10656477	1293	ID:6	P	0.18.116.6.0.6.6.6	4	19726	310973	7091	3102	0	1950	4166	63537	1518	1134	0	416
26	10819122	1285	ID:3	P	0.18.116.3.0.3.3.3	4	20093	308390	7942	4517	0	2239	1468	64023	0	416	0	416
27	10909061	1285	ID:1	P	0.18.116.1.0.1.1.1	4	15371	313112	2980	5170	0	2817	2253	63257	0	815	0	364
28	11166334	1285	ID:4	P	0.18.116.4.0.4.4.4	4	19655	308818	7542	4338	0	2226	1426	64064	0	416	0	416
29	11184417	1285	ID:5	P	0.18.116.5.0.5.5.5	4	24780	303701	10452	4518	0	2200	1425	64065	0	416	0	416
30	11306888	1285	ID:2	P	0.18.116.2.0.2.2.2	4	17828	310652	5726	2610	0	1976	2076	63411	161	593	0	416
31	12588011	1541	ID:6	P	0.18.116.6.0.6.6.6	5	21486	372532	7091	3990	0	2543	1757	61559	0	888	0	593
32	12819256	1541	ID:3	P	0.18.116.3.0.3.3.3	5	21848	372149	7942	5306	0	2806	1753	63759	0	789	0	567
33	12910930	1541	ID:1	P	0.18.116.1.0.1.1.1	5	26285	367714	8402	7027	0	3142	10911	54602	5422	1857	0	325
34	13168185	1541	ID:4	P	0.18.116.4.0.4.4.4	5	26062	367921	10016	6692	0	2780	6404	59103	2474	2354	0	554
35	13184502	1541	ID:5	P	0.18.116.5.0.5.5.5	5	26537	367458	10452	5169	0	2590	1754	63757	0	651	0	390
36	13306333	1541	ID:2	P	0.18.116.2.0.2.2.2	5	19600	374395	5726	3329	0	2379	1769	63743	0	719	0	403
37	14589286	1797	ID:6	P	0.18.116.6.0.6.6.6	6	27550	431978	10073	4458	0	2933	6061	59446	2982	468	0	390
38	14819832	1797	ID:3	P	0.18.116.3.0.3.3.3	6	24234	435261	8052	6066	0	3209	2383	63112	110	760	0	403

Figure 11. Benign "pwrtrace.csv"—1 to 38 records.

No	Real time [us]	clock_time (in ticks)	ID	P	rimeaddr	seqno	Total measurements from the begining of the simulation						Measurements every 2 seconds (monitoring period)					
							all_cpu (in ticks)	all_lpm (in ticks)	all_transmit (in ticks)	all_listen (in ticks)	all_idle_transmit (in ticks)	all_idle_listen (in ticks)	cpu (in ticks)	lpm (in ticks)	transmit (in ticks)	listen (in ticks)	idle_transmit (in ticks)	idle_listen (in ticks)
10757	3587190301	459013	ID:5	P	0.18.116.5.0.5.5.5	1792	4227046	1.1E+08	696153	1413275	0	958376	1605	63887	0	763	0	763
10758	3587313763	459013	ID:2	P	0.18.116.2.0.2.2.2	1792	4226306	1.1E+08	696849	1221793	0	754382	6257	59244	2508	2059	0	364
10759	3588594047	459269	ID:6	P	0.18.116.6.0.6.6.6	1793	4274768	1.1E+08	722143	1356760	0	875849	1587	63923	0	416	0	416
10760	3588825278	459269	ID:3	P	0.18.116.3.0.3.3.3	1793	4117288	1.1E+08	624064	1372127	0	948031	1587	63923	0	416	0	416
10761	3588916319	459269	ID:1	P	0.18.116.1.0.1.1.1	1793	8613082	1.1E+08	2425789	2442763	0	720613	2961	62549	131	758	0	403
10762	3589172501	459269	ID:4	P	0.18.116.4.0.4.4.4	1793	4237391	1.1E+08	696699	1285112	0	825907	1586	63924	0	416	0	416
10763	3589191656	459269	ID:5	P	0.18.116.5.0.5.5.5	1793	4233517	1.1E+08	698773	1415424	0	958956	6468	59033	2620	2149	0	580
10764	3589312310	459269	ID:2	P	0.18.116.2.0.2.2.2	1793	4227941	1.1E+08	696849	1222209	0	754798	1632	63878	0	416	0	416
10765	3590594057	459525	ID:6	P	0.18.116.6.0.6.6.6	1794	4276346	1.1E+08	722143	1357176	0	876265	1575	63933	0	416	0	416
10766	3590825297	459525	ID:3	P	0.18.116.3.0.3.3.3	1794	4118865	1.1E+08	624064	1372543	0	948447	1574	63933	0	416	0	416
10767	3590915623	459525	ID:1	P	0.18.116.1.0.1.1.1	1794	8615070	1.1E+08	2425789	2443179	0	721029	1985	63523	0	416	0	416
10768	3591172527	459525	ID:4	P	0.18.116.4.0.4.4.4	1794	4238968	1.1E+08	696699	1285528	0	826323	1574	63934	0	416	0	416
10769	3591190298	459525	ID:5	P	0.18.116.5.0.5.5.5	1794	4235140	1.1E+08	698773	1415840	0	959372	1620	63887	0	416	0	416
10770	3591312382	459525	ID:2	P	0.18.116.2.0.2.2.2	1794	4229517	1.1E+08	696849	1222625	0	755214	1573	63935	0	416	0	416
10771	3592594079	459781	ID:6	P	0.18.116.6.0.6.6.6	1795	4277971	1.1E+08	722143	1357592	0	876681	1622	63870	0	416	0	416
10772	3592825311	459781	ID:3	P	0.18.116.3.0.3.3.3	1795	4120490	1.1E+08	624064	1372959	0	948863	1622	63870	0	416	0	416
10773	3592915644	459781	ID:1	P	0.18.116.1.0.1.1.1	1795	8617027	1.1E+08	2425789	2443595	0	721445	1954	63555	0	416	0	416
10774	3593172522	459781	ID:4	P	0.18.116.4.0.4.4.4	1795	4240593	1.1E+08	696699	1285944	0	826739	1622	63871	0	416	0	416
10775	3593190317	459781	ID:5	P	0.18.116.5.0.5.5.5	1795	4236766	1.1E+08	698773	1416256	0	959788	1623	63870	0	416	0	416
10776	3593312391	459781	ID:2	P	0.18.116.2.0.2.2.2	1795	4231142	1.1E+08	696849	1223041	0	755630	1622	63870	0	416	0	416
10777	3594594061	460037	ID:6	P	0.18.116.6.0.6.6.6	1796	4279589	1.1E+08	722143	1358008	0	877097	1615	63876	0	416	0	416
10778	3594825310	460037	ID:3	P	0.18.116.3.0.3.3.3	1796	4122130	1.1E+08	624064	1373552	0	949456	1637	63856	0	593	0	593
10779	3594934655	460039	ID:1	P	0.18.116.1.0.1.1.1	1796	8623860	1.1E+08	2428553	2445570	0	721809	6830	59263	2764	1975	0	364
10780	3595172515	460037	ID:4	P	0.18.116.4.0.4.4.4	1796	4242210	1.1E+08	696699	1286537	0	827332	1614	63877	0	593	0	593
10781	3595190283	460037	ID:5	P	0.18.116.5.0.5.5.5	1796	4238397	1.1E+08	698773	1416966	0	960498	1628	63865	0	710	0	710
10782	3595313074	460037	ID:2	P	0.18.116.2.0.2.2.2	1796	4234321	1.1E+08	697363	1223992	0	756020	3176	62322	514	951	0	390
10783	3596594058	460293	ID:6	P	0.18.116.6.0.6.6.6	1797	4281196	1.1E+08	722143	1358614	0	877703	1604	63887	0	606	0	606
10784	3596825303	460293	ID:3	P	0.18.116.3.0.3.3.3	1797	4123737	1.1E+08	624064	1373968	0	949872	1604	63887	0	416	0	416
10785	3596915641	460293	ID:1	P	0.18.116.1.0.1.1.1	1797	8625902	1.1E+08	2428553	2445986	0	722225	2039	62889	0	416	0	416
10786	3597172526	460293	ID:4	P	0.18.116.4.0.4.4.4	1797	4243816	1.1E+08	696699	1286953	0	827748	1603	63888	0	416	0	416
10787	3597190263	460293	ID:5	P	0.18.116.5.0.5.5.5	1797	4240004	1.1E+08	698773	1417382	0	960914	1604	63887	0	416	0	416
10788	3597312372	460293	ID:2	P	0.18.116.2.0.2.2.2	1797	4235922	1.1E+08	697363	1224408	0	756436	1598	63910	0	416	0	416
10789	3598594083	460549	ID:6	P	0.18.116.6.0.6.6.6	1798	4282793	1.1E+08	722143	1359030	0	878119	1594	63897	0	416	0	416
10790	3598826646	460549	ID:3	P	0.18.116.3.0.3.3.3	1798	4128484	1.1E+08	625601	1375503	0	950262	4744	60758	1537	1535	0	390
10791	3598916331	460549	ID:1	P	0.18.116.1.0.1.1.1	1798	8629132	1.1E+08	2428874	2446790	0	722615	3227	62283	321	804	0	390
10792	3599172530	460549	ID:4	P	0.18.116.4.0.4.4.4	1798	4245412	1.1E+08	696699	1287369	0	828164	1593	63897	0	416	0	416
10793	3599191309	460549	ID:5	P	0.18.116.5.0.5.5.5	1798	4243080	1.1E+08	699229	1418517	0	961494	3073	62419	456	1135	0	580
10794	3599312385	460549	ID:2	P	0.18.116.2.0.2.2.2	1798	4237491	1.1E+08	697363	1224824	0	756852	1566	63942	0	416	0	416

Figure 12. Benign "pwrtrace.csv"—10,757 to 10,794 records.

Benign "recv.csv": The generated benign "recv.csv" file consists of 3586 records and its first 25 records (i.e., 1–25) are depicted below in Figure 13.

No	Real time [us]	ID (Receiver)	Received Message		ID (Sender)
1	11635659	ID:1	DATA recv -> 'Hello 1'	from	4
2	11768650	ID:4	DATA recv 'Reply from server'		
3	14510081	ID:1	DATA recv -> 'Hello 1'	from	3
4	14545397	ID:3	DATA recv 'Reply from server'		
5	16259239	ID:1	DATA recv -> 'Hello 1'	from	2
6	16531142	ID:2	DATA recv 'Reply from server'		
7	18258289	ID:1	DATA recv -> 'Hello 1'	from	5
8	18283595	ID:5	DATA recv 'Reply from server'		
9	19884821	ID:1	DATA recv -> 'Hello 1'	from	6
10	19937444	ID:6	DATA recv 'Reply from server'		
11	23761798	ID:1	DATA recv -> 'Hello 2'	from	4
12	23891542	ID:4	DATA recv 'Reply from server'		
13	24385405	ID:1	DATA recv -> 'Hello 2'	from	6
14	24437891	ID:6	DATA recv 'Reply from server'		
15	28008873	ID:1	DATA recv -> 'Hello 2'	from	5
16	28034048	ID:5	DATA recv 'Reply from server'		
17	29634363	ID:1	DATA recv -> 'Hello 2'	from	3
18	29669812	ID:3	DATA recv 'Reply from server'		
19	30134905	ID:1	DATA recv -> 'Hello 2'	from	2
20	30281255	ID:2	DATA recv 'Reply from server'		
21	31258819	ID:1	DATA recv -> 'Hello 3'	from	3
22	31294158	ID:3	DATA recv 'Reply from server'		
23	35260414	ID:1	DATA recv -> 'Hello 3'	from	6
24	35312814	ID:6	DATA recv 'Reply from server'		
25	38883782	ID:1	DATA recv -> 'Hello 3'	from	2

Figure 13. Benign "recv.csv"—1 to 25 records.

4.2. Benign Network Traffic Dataset Generation

4.2.1. Benign Network Traffic Dataset Generation

The generated network traffic dataset constitutes the dataset of the simulated benign IIoT network scenario that includes records consisting of IIoT network traffic features such as source/destination IPv6 address, packet size, and communication protocol. The Cooja simulator provides the "Radio messages" tool that allowed the collection of data related to the corresponding network traffic features. In Figure 14, the "Radio messages" output window is depicted along with the three configuration options that are provided by the "Radio messages" tool:

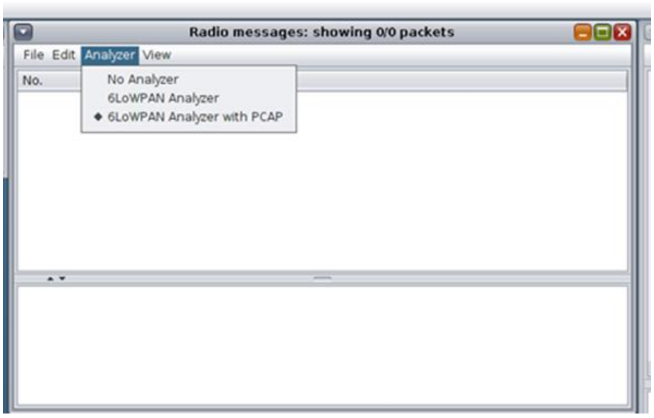

Figure 14. "Radio messages" tool—output window.

The "6LoWPAN Analyzer with PCAP" option was selected and the "Radio messages" tool saved the captured network traffic data from the simulated IIoT network into a pcap file whose file-naming format was as follows: "radiolog-" + System.currentTimeMillis() + "pcap".

During the simulation, the network traffic information about the transmitted data was also being shown in the top part of the "Radio messages" output window as depicted in the top part of Figure 15. When the simulation stopped, the generated pcap file was saved as "radiolog-1605811324302.pcap" within the " . . . /cooja/build" folder.

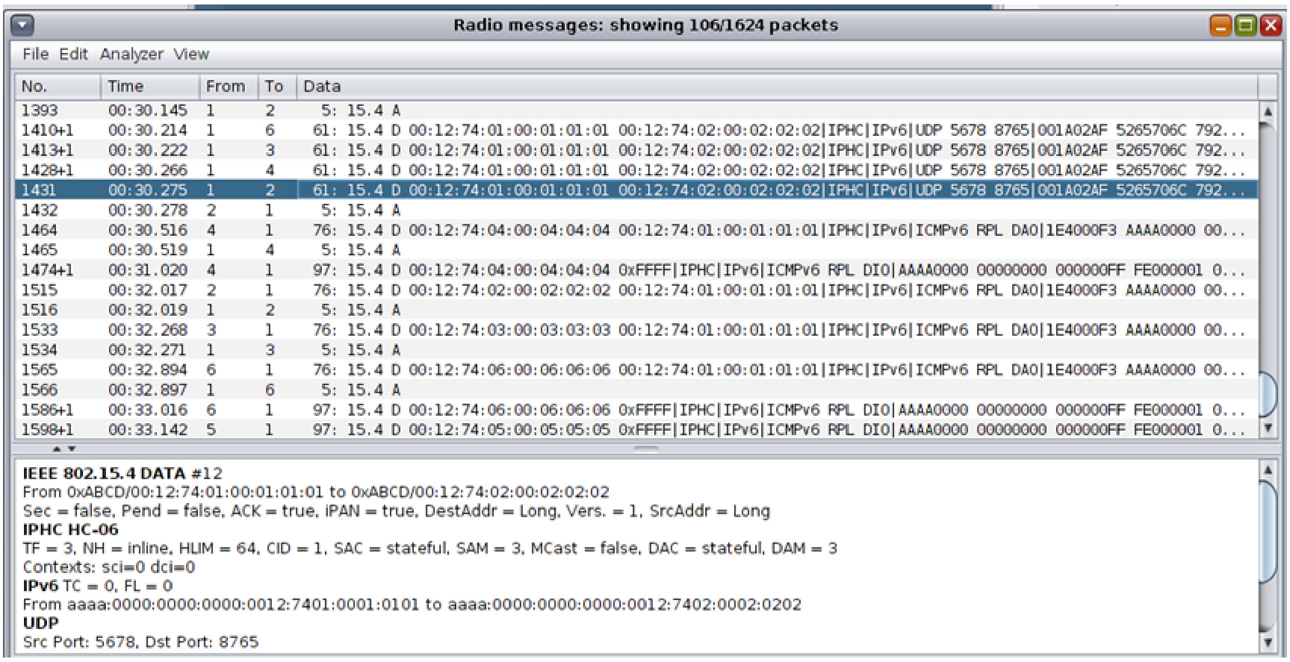

Figure 15. Network traffic information from the benign scenario in the "Radio messages" output window.

Having now saved all the captured raw network traffic information, through the "Radio messages" tool, into a pcap file, the challenging task was to extract this information from the pcap file to a csv file that would be the network traffic dataset of the simulated benign IIoT network scenario. This challenge was addressed by utilising the "IoT_Simul.sh" file that was also used in the "powertrace" dataset generation process, as described in Section 4.1, and the well-known network protocol analyser Wireshark [28].

In particular, the first step was the use of the "IoT_Simul.sh" file in order to copy the "radiolog-1605811324302.pcap" file from the " . . . /cooja/build" folder located in the Cooja Simulator environment to the "nettraffic" folder that was created by the "IoT_Simul.sh" file inside the root folder "2020-11-19-17-45-22" that was also created by the "IoT_Simul.sh" during the "powertrace" dataset generation process. The "nettraffic" folder inside the root folder "2020-11-19-17-45-22" and the copy of the "radiolog-1605811324302.pcap" file in the "nettraffic" folder is shown in Figure 16.

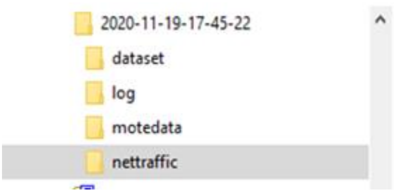

Figure 16. The "nettraffic" folder inside the root folder "2020-11-19-17-45-22" and the copy of the "radiolog-1605811324302.pcap" file.

After having the copy of the "radiolog-1605811324302.pcap" file in the "nettraffic" folder, the next step was the extraction of the stored network traffic information from the "radiolog-1605811324302.pcap" file to the "radiolog.csv" file. This was achieved through Wireshark as Wireshark allows opening a pcap file and exporting data to a csv file. In Figure 17, the upper panel of the Wireshark window shows the seventeen first packets included in the "radiolog-1605811324302.pcap" file that was opened via Wireshark. The middle panel shows the protocol details of the 10th packet selected in the upper panel and the bottom panel presents the protocol details of the selected 10th packet in both HEX and ASCII format.

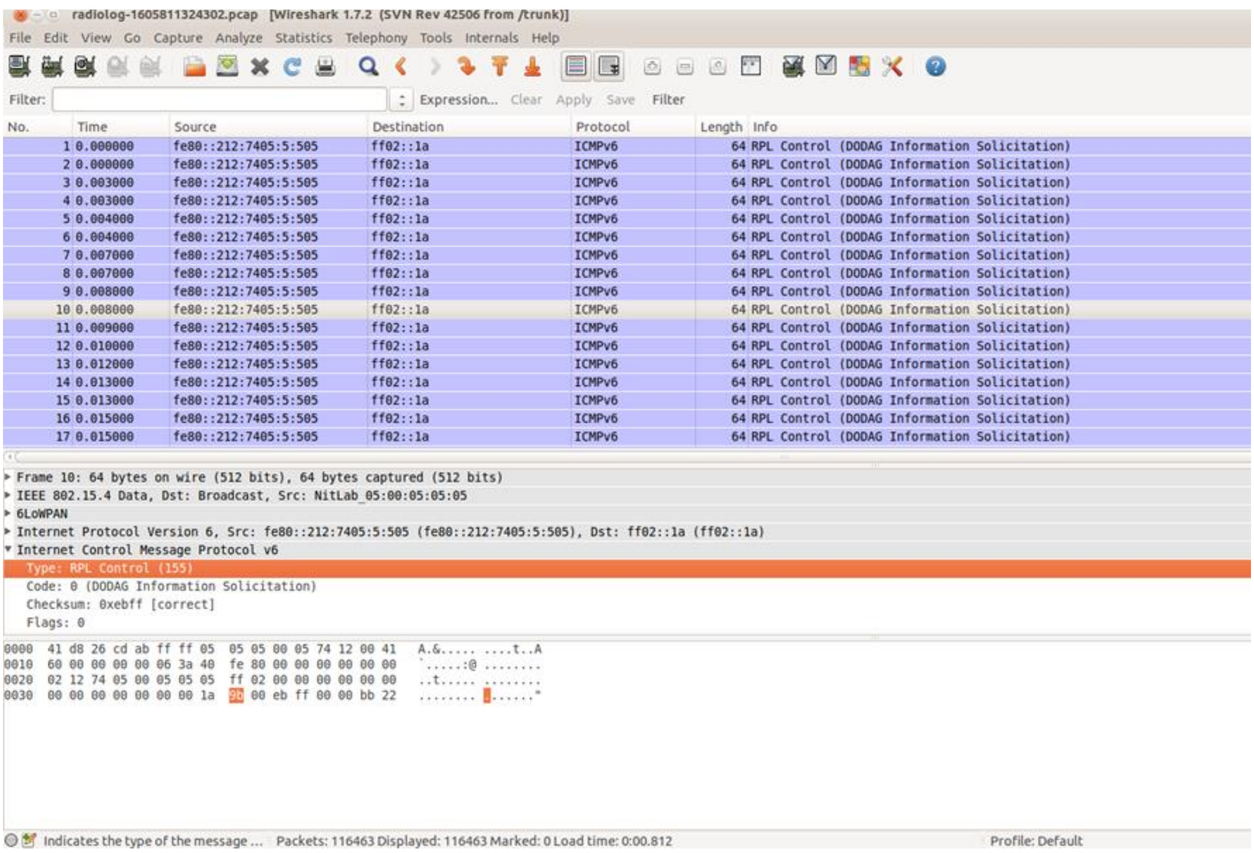

Figure 17. The first seventeenth packets in the "radiolog-1605811324302.pcap" file.

The data from the "radiolog-1605811324302.pcap" file were exported and saved, through Wireshark, into the "radiolog.csv" file in the "nettraffic" folder in the project environment, as shown in Figure 18. Furthermore, it is worthwhile to mention that we also used Wireshark to filter the "radiolog-1605811324302.pcap" file based on the ICMPv6 protocol and the UDP protocol and then exported and saved the filtered results, through Wireshark, in the "radiologICMPv6.csv" file and the "radiologUDP.csv" file, respectively, in the "nettraffic" folder in the project environment, as shown in Figure 19. The radiologICMPv6.csv" file and the "radiologUDP.csv" file facilitated the analysis of the capture traffic as shown in Section 6.

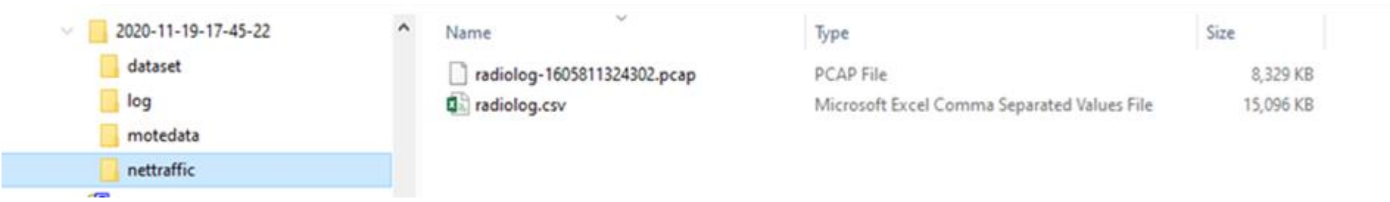

Figure 18. The "radiolog.csv" file in the "nettraffic" folder in the project environment.

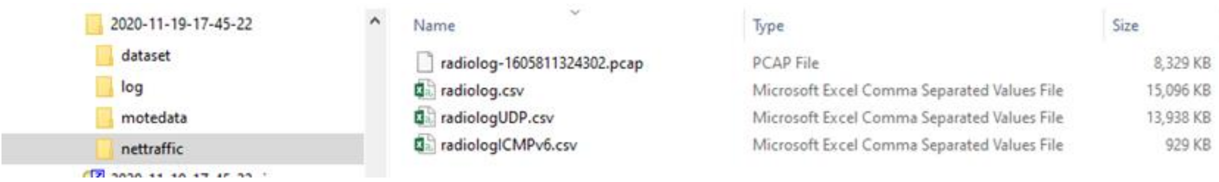

Figure 19. The "radiologICMPv6.csv" file and the "radiologUDP.csv" file in the "nettraffic" folder in the project environment.

Finally, an overview of the above mentioned process followed to extract the required information from the "radiolog-1605811324302.pcap" file to the "radiolog.csv", "radiolog-ICMPv6.csv" and "radiologUDP.csv" files is depicted in Figure 20.

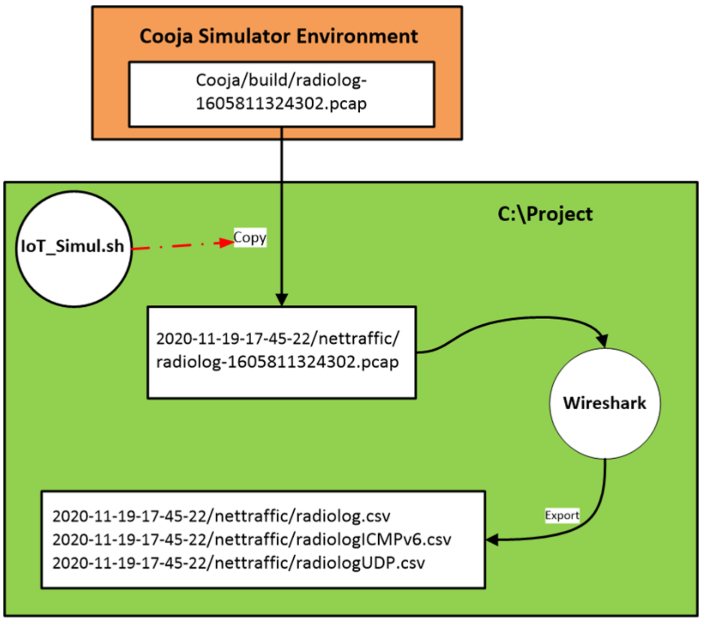

Figure 20. An overview of the process followed to extract all the required network traffic information from the "radiolog-1605811324302.pcap" file.

4.2.2. Benign Network Traffic Datasets—Results

"radiolog.csv": The generated benign "radiolog.csv" file consists of 116,463 records and its first 40 records (i.e., 1–40) are depicted below in Figure 21.

No	Time (sec)	Source Address (IPv6)	Destination Address (IPv6)	Protocol	Length (bytes)	Info
1	0	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
2	0	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
3	0.003	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
4	0.003	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
5	0.004	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
6	0.004	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
7	0.007	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
8	0.007	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
9	0.008	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
10	0.008	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
11	0.009	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
12	0.01	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
13	0.012	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
14	0.013	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
15	0.013	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
16	0.015	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
17	0.015	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
18	0.019	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
19	0.02	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
20	0.021	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
21	0.021	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
22	0.022	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
23	0.023	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
24	0.024	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
25	0.028	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
26	0.029	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
27	0.029	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
28	0.029	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
29	0.03	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
30	0.031	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
31	0.031	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
32	0.039	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
33	0.039	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
34	0.04	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
35	0.04	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
36	0.041	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
37	0.041	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
38	0.042	fe80::212:7405:5:505	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
39	0.24	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
40	0.241	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)

Figure 21. Benign "radiolog.csv"—1 to 40 records.

"radiologICMPv6.csv": The generated benign "radiologICMPv6.csv" file consists of 7975 records and its last 28 records (i.e., 7948–7975) are depicted below in Figure 22.

No	Time (sec)	Source Address (IPv6)	Destination Address (IPv6)	Protocol	Length (bytes)	Info
7948	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7949	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7950	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7951	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7952	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7953	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7954	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7955	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7956	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7957	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7958	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7959	1383.446	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7960	1384.025	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7961	1384.025	fe80::212:7402:2:202	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7962	1388.914	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7963	1388.914	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7964	1388.914	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7965	1388.914	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7966	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7967	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7968	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7969	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7970	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7971	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7972	1389.531	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7973	1389.532	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7974	1389.532	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)
7975	1389.532	fe80::212:7403:3:303	fe80::212:7401:1:101	ICMPv6	102	RPL Control (DODAG Information Object)

Figure 22. Benign "radiologICMPv6.csv"—7948 to 7975 records.

"radiologUDP.csv": The generated benign "radiologUDP.csv" file consists of 104,048 records and its last 37 records (i.e., 104,012–104,048) are depicted below in Figure 23.

No	Time (sec)	Source Address (IPv6)	Destination Address (IPv6)	Protocol	Length (bytes)	Info
104012	5160.069	2002:db8::212:7401:1:101	2002:db8::212:7404:4:404	UDP	61	Source port: rrac Destination port: ultraseek-http
104013	5228.195	2002:db8::212:7401:1:101	2002:db8::212:7404:4:404	UDP	61	Source port: rrac Destination port: ultraseek-http
104014	5288.296	2002:db8::212:7401:1:101	2002:db8::212:7404:4:404	UDP	61	Source port: rrac Destination port: ultraseek-http
104015	5338.452	2002:db8::212:7401:1:101	2002:db8::212:7404:4:404	UDP	61	Source port: rrac Destination port: ultraseek-http
104016	5384.086	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104017	5404.824	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104018	5472.868	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104019	5499.575	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104020	5537	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104021	5577.016	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104022	5604.155	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104023	5641.794	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104024	5673.504	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104025	5705.082	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104026	5735.509	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104027	5771.839	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104028	5850.894	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104029	5877.398	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104030	5909.601	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104031	5936.792	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104032	5967.579	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104033	5994.686	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104034	6027.008	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104035	6059.489	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104036	6094.091	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104037	6149.474	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104038	6185.05	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104039	6245.208	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104040	6279.464	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104041	6316.108	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104042	6362.969	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104043	6393.244	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104044	6427.186	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104045	6457.901	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104046	6522.564	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104047	6591.672	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac
104048	6647.425	2002:db8::212:7405:5:505	2002:db8::ff:fe00:1	UDP	53	Source port: ultraseek-http Destination port: rrac

Figure 23. Benign "radiologUDP.csv"—104,012 to 104,048 records.

5. Generation of Malicious IoT/IIoT Datasets

In this Section, we provide a detailed description of the approach followed to generate a set of malicious datasets by implementing a UDP flooding attack scenario in the Cooja simulator, as shown in Figure 24. Similar to the approach followed for the generation of the benign datasets in Section 4, the generated IoT/IIoT-specific information from the simulated attack scenario was captured from the Contiki plugin "powertrace" (i.e., features such as CPU consumption) and the Cooja tool "Radio messages" (i.e., network traffic features) in order to generate the "powertrace" dataset and the network traffic dataset for the simulated UDP flooding attack scenario.

The network topology of the simulated UDP flooding attack scenario in the Cooja simulator environment consists of 4 yellow (benign) UDP-client motes (i.e., motes 2, 3, 4 and 5), the violet (malicious) UDP-client mote (i.e., mote 6) and the green (benign) UDP-sever mote (i.e., mote 1), as depicted in Figure 24. The simulation duration was set to 60 min and the motes' outputs were printed out in the respective window (e.g., Mote output) while simulations run, as shown in Figure 25. Moreover, the 4 yellow (benign) UDP-client motes were configured to send text messages every 10 s, approximately, to the UDP-sever mote that was configured to provide a corresponding response. On the other hand, the violet (malicious) UDP-client mote (i.e., mote 6) was compromised with malicious code in order to send UDP packets within a very short period of time (i.e., every 200 ms). Finally, it is noteworthy to say that similar to the benign network scenario, the UDP protocol was used at the Transport Layer, the IPv6 at the network layer, and the type of motes was the Tmote Sky in the UDP flooding attack scenario.

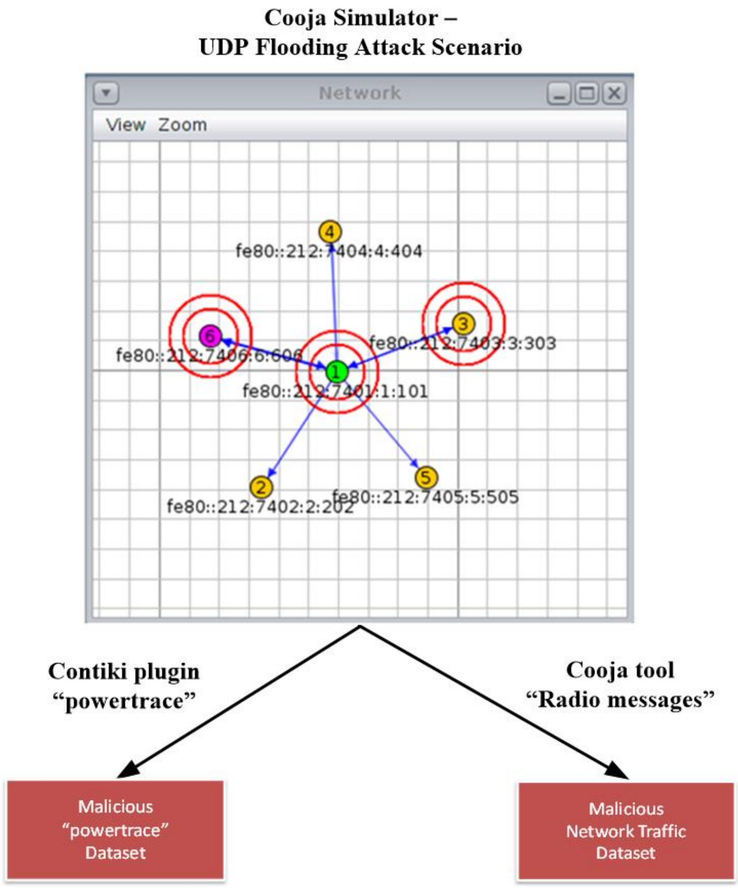

Figure 24. Malicious datasets generation by utilizing the Cooja simulator.

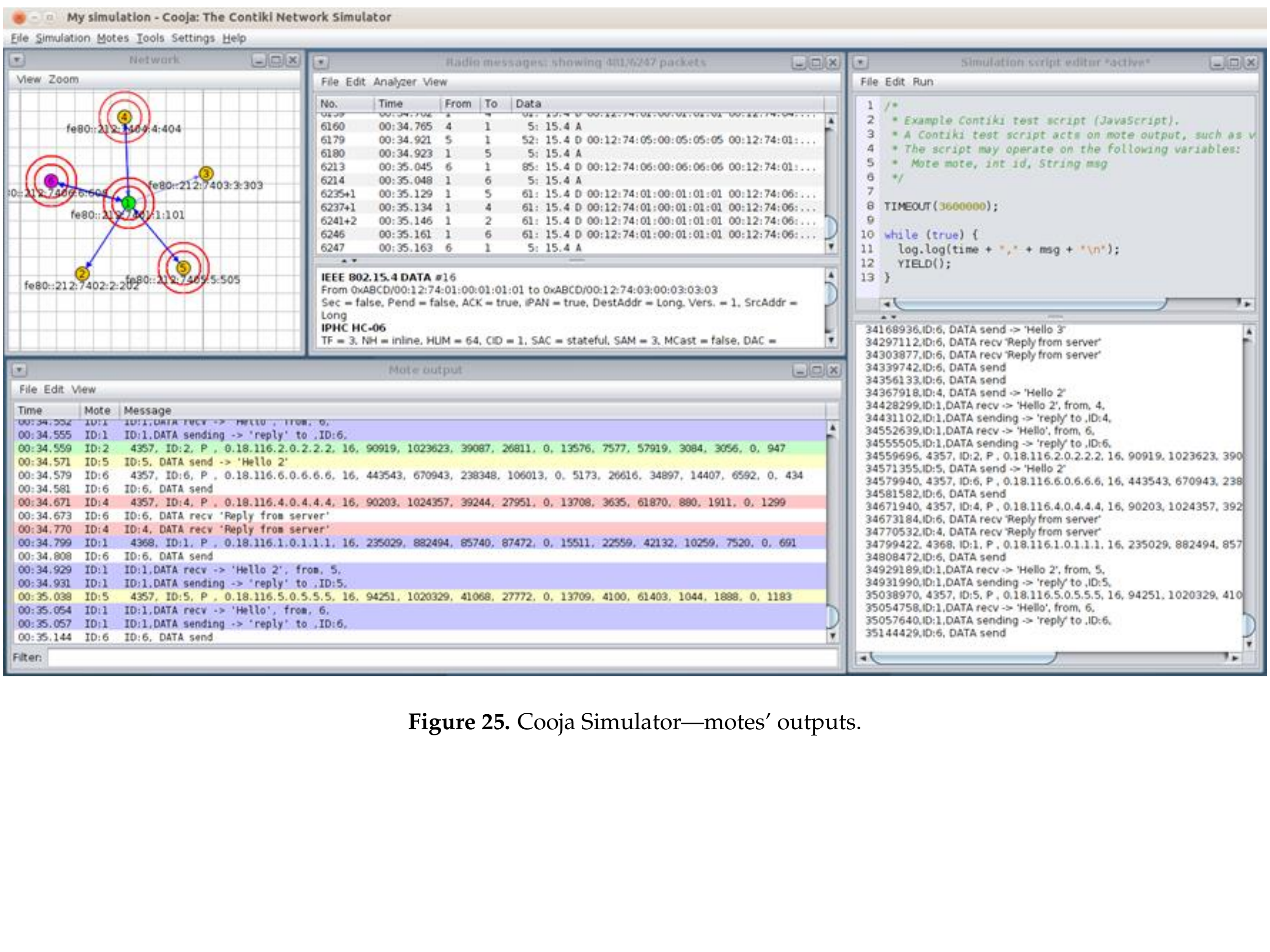

Figure 25. Cooja Simulator—motes’ outputs.

5.1. Malicious "Powertrace" Dataset Generation

5.1.1. Malicious "Powertrace" Dataset Generation

The approach followed for the "powertrace" dataset generation from the UDP flooding attack scenario was similar to the approach followed for the "powertrace" dataset generation from the benign IIoT network scenario in Section 4.1.1. In addition, the "powertrace" plugin was similarly enabled for collecting "powertrace" related features, summarised in Table 1, from the motes of the attack scenario every two seconds. In Figure 26, the depicted mote output window displays the captured "powertrace" information every two seconds and also the messages sent and received by each mote during the simulation time (60 min).

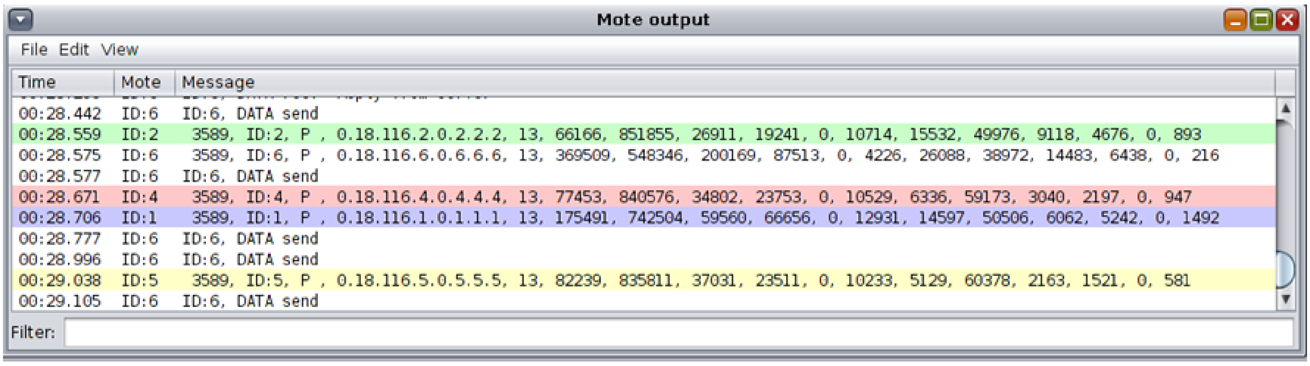

Figure 26. Cooja Simulator—Mote output window.

When the timeout occurred, the simulation stopped, and all the captured information and prints were stored in the "COOJA.testlog" file. Afterwards, the "IoT_Simul.sh" file, described in Section 4.1.1, created (a) a new root folder named as "2020-12-09-14-59-59", and (b) the "log" folder, inside the "2020-12-09-14-59-59" folder, where the "COOJA.testlog" file was copied from the " . . . /cooja/build" folder located in the Cooja Simulator. Then, the "IoT_Simul.sh" file following the same process, as described in Section 4.1.1, extracted the required "powertrace" information from the "COOJA.testlog" file and saved it in the "pwrtrace.csv" file in the "dataset" folder that was created by the batch file inside the "2020-12-09-14-59-59" folder, as shown below in the left part of Figure 27. In the "dataset" folder, apart from the "pwrtrace.csv" file, the "IoT_Simul.sh" file generated two more files (i.e., the "recv.csv" file and the "send.csv"), following the same process as in Section 4.1.1. The "recv.csv" file and the "send.csv" file include the "received" and "sent" messages printed by the motes, respectively.

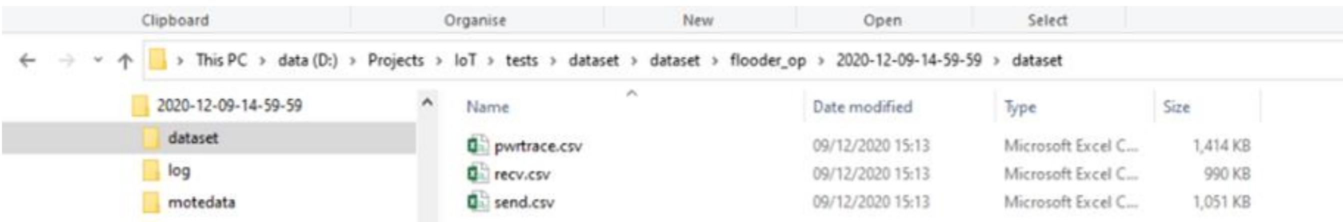

Figure 27. Location of the generated "pwrtrace.csv", "recv.csv", and "send.csv" files by the "IoT_Simul.sh" bash file.

Finally, similar to the benign "powertrace" dataset generation approach in Section 4.1.1, the "IoT_Simul.sh" file extracted the information related to each mote from the "pwrtrace.csv" file and generated one csv file for each mote with the corresponding information from the "pwrtrace.csv" file. The generated six csv files (i.e., mote1.csv, mote2.csv, mote3.csv, mote4.csv, mote5.csv, and mote6.csv) were stored in the "motedata" folder, created also by the "IoT_Simul.sh" file, as shown in the left part of Figure 27.

5.1.2. Malicious "powertrace" Datasets—Results

Malicious "pwrtrace.csv": The generated malicious "pwrtrace.csv" file consists of 10,794 records and its first 38 records (i.e., 1–38) and its last 38 records (10,757–10,794) are depicted in Figures 28 and 29, respectively.

B	D	E	F	G	H	I	K	L	M	N	O	P	R	S	T	U	V	W
							Total measurements from the begining of the simulation						Measurements for each of the 2-sec monitoring period					
No	Real time [us]	Clock time (in ticks)	ID		Rime Address	seq no	all_cpu (in ticks)	all_lpm (in ticks)	all_transmit (in ticks)	all_listen (in ticks)	all_idle_transmit (in ticks)	all_idle_listen (in ticks)	cpu (in ticks)	lpm (in ticks)	transmit (in ticks)	listen (in ticks)	idle_transmit (in ticks)	idle_listen (in ticks)
1	2555692	261	ID:2	P	0.18.116.2.0.2.2.2	0	6742	59714	2589	442	0	364	6742	59714	2589	442	0	364
2	2570487	261	ID:6	P	0.18.116.6.0.6.6.6	0	7709	58725	2590	442	0	364	7709	58725	2590	442	0	364
3	2665753	261	ID:4	P	0.18.116.4.0.4.4.4	0	2189	64265	0	390	0	390	2189	64265	0	390	0	390
4	2699493	261	ID:1	P	0.18.116.1.0.1.1.1	0	2817	63639	0	999	0	744	2817	63639	0	999	0	744
5	3034683	261	ID:5	P	0.18.116.5.0.5.5.5	0	6742	59714	2589	442	0	364	6742	59714	2589	442	0	364
6	3216735	261	ID:3	P	0.18.116.3.0.3.3.3	0	2189	64265	0	390	0	390	2189	64265	0	390	0	390
7	4554978	517	ID:2	P	0.18.116.2.0.2.2.2	1	7904	124063	2589	858	0	780	1159	64349	0	416	0	416
8	4575548	517	ID:6	P	0.18.116.6.0.6.6.6	1	10228	121854	2590	1159	0	767	2516	63129	0	717	0	403
9	4671767	517	ID:4	P	0.18.116.4.0.4.4.4	1	3574	128552	0	1104	0	1056	1382	64287	0	714	0	666
10	4702609	517	ID:1	P	0.18.116.1.0.1.1.1	1	8551	123417	2980	1467	0	1134	5731	59778	2980	468	0	390
11	5034813	517	ID:5	P	0.18.116.5.0.5.5.5	1	8255	123715	2589	1136	0	1010	1510	64001	0	694	0	646
12	5217991	517	ID:3	P	0.18.116.3.0.3.3.3	1	3658	128314	0	1090	0	1043	1466	64049	0	700	0	653
13	6555863	773	ID:2	P	0.18.116.2.0.2.2.2	2	9577	187908	2589	1471	0	1170	1670	63845	0	613	0	390
14	6573145	773	ID:6	P	0.18.116.6.0.6.6.6	2	40545	156877	20450	8529	0	988	30315	35023	17860	7370	0	221
15	6666980	773	ID:4	P	0.18.116.4.0.4.4.4	2	4960	192521	0	1520	0	1472	1383	63969	0	416	0	416
16	6704432	773	ID:1	P	0.18.116.1.0.1.1.1	2	12693	184842	2980	3685	0	2194	4140	61425	0	2218	0	1060
17	7036198	773	ID:5	P	0.18.116.5.0.5.5.5	2	14278	183202	5575	1605	0	1400	6020	59487	2986	469	0	390
18	7217945	773	ID:3	P	0.18.116.3.0.3.3.3	2	5047	192434	0	1506	0	1459	1386	64120	0	416	0	416
19	8557499	1029	ID:2	P	0.18.116.2.0.2.2.2	3	15580	247416	5574	1940	0	1560	6000	59508	2985	469	0	390
20	8574202	1029	ID:6	P	0.18.116.6.0.6.6.6	3	72195	190733	39240	14939	0	1222	31648	33856	18790	6410	0	234
21	8670462	1029	ID:4	P	0.18.116.4.0.4.4.4	3	21137	241852	9460	4759	0	1810	16174	49331	9460	3239	0	338
22	8702861	1029	ID:1	P	0.18.116.1.0.1.1.1	3	15882	247108	2980	6503	0	3838	3186	62266	0	2818	0	1644
23	9037531	1029	ID:5	P	0.18.116.5.0.5.5.5	3	25136	237851	11495	4573	0	1738	10855	54649	5920	2968	0	338
24	9221415	1029	ID:3	P	0.18.116.3.0.3.3.3	3	19298	243688	8345	4245	0	1823	14248	51254	8345	2739	0	364
25	10558220	1285	ID:2	P	0.18.116.2.0.2.2.2	4	25340	303155	10934	4604	0	1924	9757	55739	5360	2664	0	364
26	10574593	1285	ID:6	P	0.18.116.6.0.6.6.6	4	102520	225896	56293	22039	0	1456	30322	35163	17053	7100	0	234
27	10668636	1285	ID:4	P	0.18.116.4.0.4.4.4	4	22607	305875	9460	5175	0	2226	1468	64023	0	416	0	416
28	10707377	1285	ID:1	P	0.18.116.1.0.1.1.1	4	21475	307168	2980	10148	0	4695	5590	60060	0	3645	0	857
29	11035707	1285	ID:5	P	0.18.116.5.0.5.5.5	4	26575	301905	11495	4989	0	2154	1437	64054	0	416	0	416
30	11219597	1285	ID:3	P	0.18.116.3.0.3.3.3	4	20726	307753	8345	4661	0	2239	1426	64065	0	416	0	416
31	12557488	1541	ID:2	P	0.18.116.2.0.2.2.2	5	27170	366840	10934	5773	0	3048	1828	63685	0	1169	0	1124
32	12669462	1541	ID:4	P	0.18.116.4.0.4.4.4	5	24354	369643	9460	6363	0	3383	1745	63768	0	1188	0	1157
33	12700632	1557	ID:6	P	0.18.116.6.0.6.6.6	5	134474	263557	73964	30327	0	1940	31951	37661	17671	8288	0	484
34	12821697	1556	ID:1	P	0.18.116.1.0.1.1.1	5	45913	351915	15135	17366	0	5621	24436	44747	12155	7218	0	926
35	13036484	1541	ID:5	P	0.18.116.5.0.5.5.5	5	28370	365626	11495	6481	0	3331	1792	63721	0	1492	0	1177
36	13219734	1541	ID:3	P	0.18.116.3.0.3.3.3	5	22495	371498	8345	5528	0	2806	1766	63745	0	867	0	567
37	14557450	1797	ID:2	P	0.18.116.2.0.2.2.2	6	28686	430834	10934	6773	0	4048	1513	63994	0	1000	0	1000
38	14590601	1799	ID:6	P	0.18.116.6.0.6.6.6	6	167799	292124	92841	37747	0	2083	33323	28567	18877	7420	0	143

Figure 28. Malicious "pwrtrace.csv"—1 to 38 records.

B	D	E	F	G	H	I	K	L	M	N	O	P	R	S	T	U	V	W
							Total measurements from the begining of the simulation						Measurements for each of the 2-sec monitoring period					
No	Real time [us]	Clock time (in ticks)	ID		Rime Address	seq no	all_cpu (in ticks)	all_lpm (in ticks)	all_transmit (in ticks)	all_listen (in ticks)	all_idle_transmit (in ticks)	all_idle_listen (in ticks)	cpu (in ticks)	lpm (in ticks)	transmit (in ticks)	listen (in ticks)	idle_transmit (in ticks)	idle_listen (in ticks)
10757	3.587E+09	459018	ID:5	P	0.18.116.5.0.5.5.5	1792	6484924	110972410	1976106	3249351	0	2067142	13864	52920	7065	5266	0	1092
10758	3.587E+09	459013	ID:3	P	0.18.116.3.0.3.3.3	1792	6407343	111044806	1988080	2342008	0	1181632	1615	63875	0	416	0	416
10759	3.589E+09	459269	ID:2	P	0.18.116.2.0.2.2.2	1793	6288419	111233629	1859570	3180790	0	2071651	10964	54533	5355	4048	0	908
10760	3.589E+09	459269	ID:6	P	0.18.116.6.0.6.6.6	1793	49272148	68222749	26428032	12698225	0	487982	21797	41017	11343	5525	0	234
10761	3.589E+09	459269	ID:4	P	0.18.116.4.0.4.4.4	1793	6077237	111445104	1735004	3122961	0	2078867	1654	63857	0	960	0	960
10762	3.589E+09	459269	ID:1	P	0.18.116.1.0.1.1.1	1793	37354505	80163010	16447901	12709259	0	1274462	15538	49969	6420	5486	0	976
10763	3.589E+09	459269	ID:5	P	0.18.116.5.0.5.5.5	1793	6486773	111034789	1976106	3250322	0	2067906	1846	62379	0	971	0	764
10764	3.589E+09	459269	ID:3	P	0.18.116.3.0.3.3.3	1793	6408983	111108661	1988080	2342621	0	1182245	1637	63855	0	613	0	613
10765	3.591E+09	459525	ID:2	P	0.18.116.2.0.2.2.2	1794	6293423	111294132	1861337	3182661	0	2072205	5002	60503	1767	1871	0	554
10766	3.591E+09	459528	ID:6	P	0.18.116.6.0.6.6.6	1794	49303975	68257291	26445531	12706237	0	488374	31824	34542	17499	8012	0	392
10767	3.591E+09	459525	ID:4	P	0.18.116.4.0.4.4.4	1794	6078847	111509005	1735004	3123744	0	2079650	1607	63901	0	783	0	783
10768	3.591E+09	459540	ID:1	P	0.18.116.1.0.1.1.1	1794	37376330	80210698	16456056	12717221	0	1274968	21822	47688	8155	7962	0	506
10769	3.591E+09	459525	ID:5	P	0.18.116.5.0.5.5.5	1794	6488393	111098680	1976106	3251466	0	2069050	1617	63891	0	1144	0	1144
10770	3.591E+09	459525	ID:3	P	0.18.116.3.0.3.3.3	1794	6413239	111169907	1989162	2344529	0	1183238	4253	61246	1082	1908	0	993
10771	3.593E+09	459781	ID:2	P	0.18.116.2.0.2.2.2	1795	6295156	111357899	1861337	3183818	0	2073362	1730	63767	0	1157	0	1157
10772	3.593E+09	459782	ID:6	P	0.18.116.6.0.6.6.6	1795	49329509	68296718	26458484	12713264	0	488746	25532	39427	12953	7027	0	372
10773	3.593E+09	459781	ID:4	P	0.18.116.4.0.4.4.4	1795	6080517	111572831	1735004	3125078	0	2080984	1667	63826	0	1334	0	1334
10774	3.593E+09	459781	ID:1	P	0.18.116.1.0.1.1.1	1795	37397949	80250568	16465241	12724409	0	1275229	21616	39870	9185	7188	0	261
10775	3.593E+09	459781	ID:5	P	0.18.116.5.0.5.5.5	1795	6490075	111162496	1976106	3252623	0	2070207	1679	63816	0	1157	0	1157
10776	3.593E+09	459781	ID:3	P	0.18.116.3.0.3.3.3	1795	6414946	111233697	1989162	2345345	0	1184054	1704	63790	0	816	0	816
10777	3.595E+09	460037	ID:2	P	0.18.116.2.0.2.2.2	1796	6296887	111421667	1861337	3185493	0	2075037	1728	63768	0	1675	0	1675
10778	3.595E+09	460037	ID:6	P	0.18.116.6.0.6.6.6	1796	49355789	68335752	26472262	12720153	0	488850	26277	39034	13778	6889	0	104
10779	3.595E+09	460037	ID:4	P	0.18.116.4.0.4.4.4	1796	6082233	111636611	1735004	3126792	0	2082698	1713	63780	0	1714	0	1714
10780	3.595E+09	460037	ID:1	P	0.18.116.1.0.1.1.1	1796	37428570	80285455	16481427	12733982	0	1275681	30619	34887	16186	9573	0	452
10781	3.595E+09	460037	ID:5	P	0.18.116.5.0.5.5.5	1796	6491782	111226285	1976106	3254337	0	2071921	1704	63789	0	1714	0	1714
10782	3.595E+09	460037	ID:3	P	0.18.116.3.0.3.3.3	1796	6416585	111297551	1989162	2346148	0	1184857	1636	63854	0	803	0	803
10783	3.597E+09	460293	ID:2	P	0.18.116.2.0.2.2.2	1797	6303068	111480993	1863731	3188667	0	2076549	6178	59326	2394	3174	0	1512
10784	3.597E+09	460296	ID:6	P	0.18.116.6.0.6.6.6	1797	49386703	68371029	26489236	12727987	0	489439	30911	35277	16974	7834	0	589
10785	3.597E+09	460293	ID:4	P	0.18.116.4.0.4.4.4	1797	6088623	111695721	1737682	3129879	0	2084216	6387	59110	2678	3087	0	1518
10786	3.597E+09	460293	ID:1	P	0.18.116.1.0.1.1.1	1797	37456122	80323411	16495172	12743176	0	1276539	27549	37956	13745	9194	0	858
10787	3.597E+09	460293	ID:5	P	0.18.116.5.0.5.5.5	1797	6493481	111290084	1976106	3255887	0	2073471	1696	63799	0	1550	0	1550
10788	3.597E+09	460293	ID:3	P	0.18.116.3.0.3.3.3	1797	6427436	111352202	1994524	2350173	0	1185746	10848	54651	5362	4025	0	889
10789	3.599E+09	460549	ID:2	P	0.18.116.2.0.2.2.2	1798	6304739	111544834	1863731	3189634	0	2077516	1668	63841	0	967	0	967
10790	3.599E+09	460549	ID:6	P	0.18.116.6.0.6.6.6	1798	49415032	68407317	26505403	12735002	0	489824	28326	36288	16167	7015	0	385
10791	3.599E+09	460549	ID:4	P	0.18.116.4.0.4.4.4	1798	6103717	111746125	1745707	3135503	0	2085151	15092	50404	8025	5624	0	935
10792	3.599E+09	460563	ID:1	P	0.18.116.1.0.1.1.1	1798	37476871	80371940	16505059	12750749	0	1277638	20746	48529	9887	7573	0	1099
10793	3.599E+09	460549	ID:5	P	0.18.116.5.0.5.5.5	1798	6499100	111349971	1978161	3258443	0	2074569	5616	59887	2055	2556	0	1098
10794	3.599E+09	460549	ID:3	P	0.18.116.3.0.3.3.3	1798	6433759	111411377	1997201	2352157	0	1186162	6321	59175	2677	1984	0	416

Figure 29. Malicious "pwrtrace.csv"—10,757 to 10,794 records.

Malicious "recv.csv": The generated malicious "recv.csv" file consists of 21,573 records and its first 27 records (i.e., 1–27) are depicted below in Figure 30.

B	C	D	E	F	G
No	Real time [us]	ID (Receiver)	Received message		ID (Sender)
1	4928647	ID:1	DATA recv -> 'Hello'	from	6
2	5179252	ID:1	DATA recv -> 'Hello'	from	6
3	5427442	ID:1	DATA recv -> 'Hello'	from	6
4	5555622	ID:1	DATA recv -> 'Hello'	from	6
5	5803531	ID:1	DATA recv -> 'Hello'	from	6
6	5926813	ID:1	DATA recv -> 'Hello'	from	6
7	6305514	ID:1	DATA recv -> 'Hello'	from	6
8	6428582	ID:1	DATA recv -> 'Hello'	from	6
9	6677519	ID:1	DATA recv -> 'Hello'	from	6
10	7303554	ID:1	DATA recv -> 'Hello'	from	6
11	8930177	ID:1	DATA recv -> 'Hello'	from	6
12	8939620	ID:1	DATA recv -> 'Hello'	from	6
13	9178554	ID:1	DATA recv -> 'Hello'	from	6
14	9680082	ID:1	DATA recv -> 'Hello'	from	6
15	9690472	ID:1	DATA recv -> 'Hello'	from	6
16	9928890	ID:1	DATA recv -> 'Hello'	from	6
17	9938388	ID:1	DATA recv -> 'Hello'	from	6
18	10178017	ID:1	DATA recv -> 'Hello'	from	6
19	10430699	ID:1	DATA recv -> 'Hello'	from	6
20	10440280	ID:1	DATA recv -> 'Hello'	from	6
21	10680895	ID:1	DATA recv -> 'Hello'	from	6
22	10930537	ID:1	DATA recv -> 'Hello'	from	6
23	11180697	ID:1	DATA recv -> 'Hello'	from	6
24	11193252	ID:1	DATA recv -> 'Hello'	from	6
25	11796952	ID:6	DATA recv 'Reply from server'		
26	11803849	ID:6	DATA recv 'Reply from server'		
27	12179677	ID:1	DATA recv -> 'Hello'	from	6

Figure 30. Malicious "recv.csv"—1 to 27 records.

5.2. Malicious Network Traffic Dataset Generation

5.2.1. Malicious Network Traffic Dataset Generation

The approach followed for the network traffic dataset generation from the UDP flooding attack scenario was similar to the approach followed for the network traffic dataset generation from the benign IIoT network scenario in Section 4.2.1. The "Radio messages" tool, provided by the Cooja simulator, was similarly used for collecting data related to the corresponding network traffic features (e.g., source/destination IPv6 address, packet size, and communication protocol) from the network of the attack scenario. During the simulation, the network traffic information was being shown in the top part of the "Radio messages" output window as depicted in the top part of Figure 31.

When the simulation stopped, the generated pcap file was saved as "radiolog-1607519517066.pcap" within the " . . . /cooja/build" folder. Afterwards, the "IoT_Simul.sh" file, described in Section 4.2.1, created (a) a new root folder named as "2020-12-09-14-59-59", and (b) the "nettraffic" folder, inside the "2020-12-09-14-59-59" folder, where the "radiolog-1607519517066.pcap" file was copied from the " . . . /cooja/build" folder located in the Cooja Simulator. The "nettraffic" folder inside the root folder "2020-12-09-14-59-59" and the copy of the "radiolog-1607519517066.pcap" file in the "nettraffic" folder are shown in Figure 32.

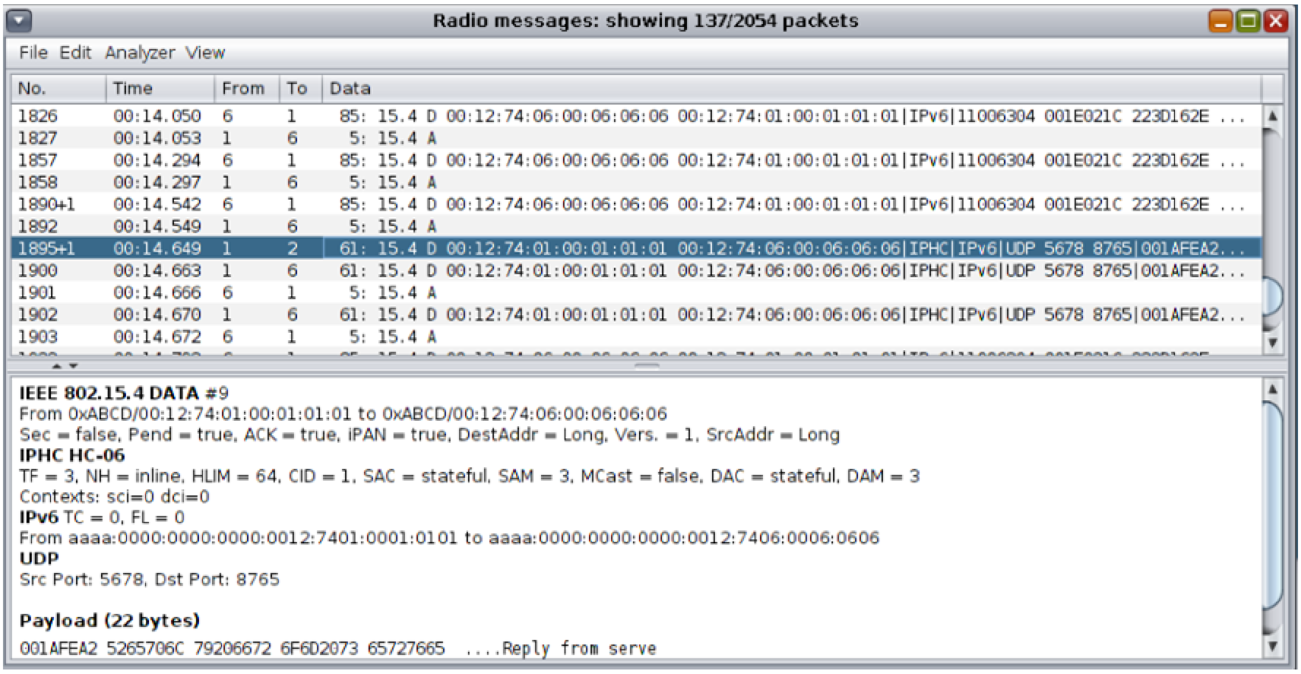

Figure 31. Network traffic information from the attack scenario in the "Radio messages" output window.

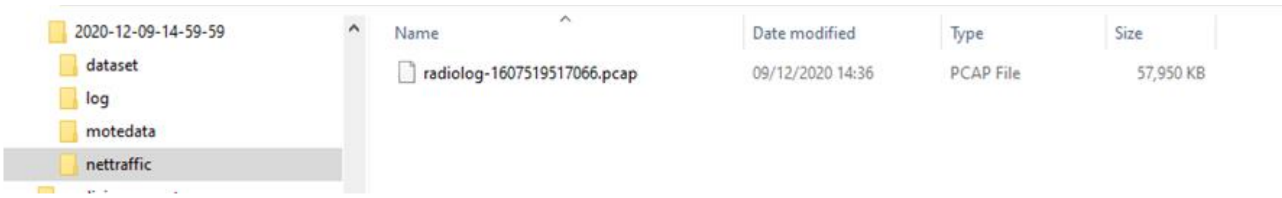

Figure 32. The "nettraffic" folder inside the root folder "2020-12-09-14-59-59" and the copy of the "radiolog-1607519517066.pcap" file.

Then, following the same process, as described in Section 4.2.1, we used Wireshark to extract the stored network traffic information from the "radiolog-1607519517066.pcap" file to the "radiolog.csv" file stored in the "nettraffic" folder as shown in Figure 33.

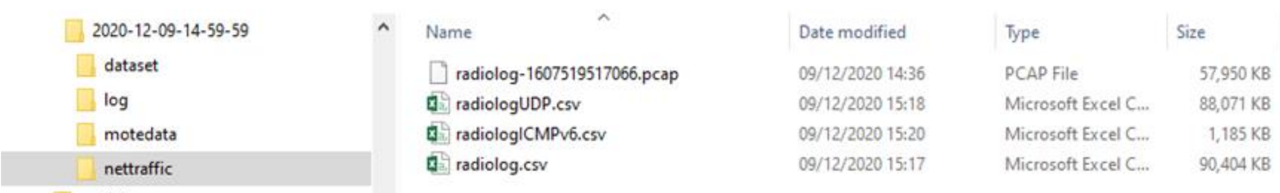

Figure 33. The "nettraffic" folder inside the root folder "2020-12-09-14-59-59" and its included files.

In the "nettraffic" folder, apart from the "radiolog.csv" file, we also used Wireshark, following the same process as in Section 4.2.1, to generate two more files (i.e., the "radiologICMPv6.csv" file and the "radiologUDP.csv" file) from the "radiolog-1607519517066.pcap" file.

5.2.2. Malicious Network Traffic Datasets—Results

"radiolog.csv": The generated malicious "radiolog.csv" file consists of 702,332 records and its first 25 records (i.e., 1–25) are depicted below in Figure 34.

B	C	D	E	F	G	H
No.	**Time**	**Source**	**Destination**	**Protocol**	**Length**	**Info**
1	0	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
2	0.032	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
3	0.033	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
4	0.067	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
5	0.1	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
6	0.175	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
7	0.176	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
8	0.197	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
9	0.199	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
10	0.201	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
11	0.203	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
12	0.26	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
13	0.262	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
14	0.329	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
15	0.33	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
16	0.332	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
17	0.333	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
18	0.391	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
19	0.397	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
20	0.441	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
21	0.459	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
22	0.497	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
23	0.498	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
24	0.499	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
25	0.5	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)

Figure 34. Malicious "radiolog.csv"—1 to 25 records.

"radiologICMPv6.csv": The generated malicious "radiologICMPv6.csv" file consists of 9908 records and its first 25 records (i.e., 1–25) are depicted below in Figure 35.

B	C	D	E	F	G	H
No.	**Time**	**Source**	**Destination**	**Protocol**	**Length**	**Info**
1	0	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
2	0.032	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
3	0.033	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
4	0.067	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
5	0.1	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
6	0.175	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
7	0.176	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
8	0.197	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
9	0.199	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
10	0.201	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
11	0.203	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
12	0.26	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
13	0.262	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
14	0.329	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
15	0.33	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
16	0.332	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
17	0.333	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
18	0.391	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
19	0.397	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
20	0.441	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
21	0.459	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
22	0.497	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
23	0.498	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
24	0.499	fe80::212:7406:6:606	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)
25	0.5	fe80::212:7402:2:202	ff02::1a	ICMPv6	64	RPL Control (DODAG Information Solicitation)

Figure 35. Malicious "radiologICMPv6.csv"—1 to 25 records.

"radiologUDP.csv": The generated malicious "radiologUDP.csv" file consists of 670,671 records and its first 25 records (i.e., 1–25) are depicted below in Figure 36.

B	C	D	E	F	G	H
No.	Time	Source	Destination	Protocol	Length	Info
1	1.234	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
2	1.235	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
3	1.236	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
4	1.236	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
5	1.237	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
6	1.238	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
7	1.239	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
8	1.24	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
9	1.24	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
10	1.241	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
11	1.242	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
12	1.242	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
13	1.243	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
14	1.243	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
15	1.244	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
16	1.245	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
17	1.245	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
18	1.246	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
19	1.246	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
20	1.247	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
21	1.248	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
22	1.248	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
23	1.249	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
24	1.25	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac
25	1.25	aaaa::212:7406:6:606	aaaa::ff:fe00:1	UDP	85	Source port: ultraseek-http Destination port: rrac

Figure 36. Malicious "radiologUDP.csv"—1 to 25 records.

6. Discussion on the Generated Datasets

The generated benign and malicious "pwrtrace" datasets, presented in Sections 4.1.2 and 5.1.2, respectively, include information about raw features (e.g., all_cpu, all_lpm, all_transmit, all_listen) which can be used to derive new features more informative, in terms of the behaviour of each mote, and non-redundant. These new features are intended to constitute valuable features for training and evaluating AIDS for IoT/IIoT networks. Towards this direction, the total energy consumption of a mote in an IoT/IIoT network can be considered as a valuable feature for detection of a UDP flooding attack and its source as the compromised mote carrying out the attack is characterised by high total energy consumption, as demonstrated below.

Based on [29,30], the total energy consumption of each mote, at the reading (i.e., record) *i*, is given by the sum of (a) the energy consumption in the CPU state; (b) the energy consumption in the LPM state; (c) the energy consumption in the Tx state; and the average power consumption Listen state, at the reading (i.e., record) *i*, as shown in the equation below:

$$\begin{aligned} \mathrm{E}_{\mathrm{total}_i}(\mathrm{mj}) &= \mathrm{E}_{\mathrm{cpu}_{\mathrm{total}_i}} + \mathrm{E}_{\mathrm{lpm}_{\mathrm{total}_i}} + \mathrm{E}_{\mathrm{tx}_{\mathrm{total}_i}} + \mathrm{E}_{\mathrm{rx}_{\mathrm{total}_i}} = \\ &= \left(\mathrm{I}_{\mathrm{cpu}} \times \mathrm{V}_{\mathrm{cpu}} \times \mathrm{T}_{\mathrm{cpu}_i}\right) + \left(\mathrm{I}_{\mathrm{lpm}} \times \mathrm{V}_{\mathrm{lpm}} \times \mathrm{T}_{\mathrm{lpm}_i}\right) + \left(\mathrm{I}_{\mathrm{tx}} \times \mathrm{V}_{\mathrm{tx}} \times \mathrm{T}_{\mathrm{tx}_i}\right) + \left(\mathrm{I}_{\mathrm{rx}} \times \mathrm{V}_{\mathrm{rx}} \times \mathrm{T}_{\mathrm{rx}_i}\right) \end{aligned} \tag{1}$$

where

$\mathrm{I}_{\mathrm{cpu}}$: the nominal current in the CPU state;
$\mathrm{I}_{\mathrm{lpm}}$: the nominal current in the LPM state;
I_{tx}: the nominal current in the TX state;
I_{rx}: the nominal current in the RX state;
$\mathrm{V}_{\mathrm{cpu}}$: the nominal voltage in the CPU state;

V_{lpm}: the nominal voltage in the LPM state;
V_{tx}: the nominal voltage in the TX state;
V_{rx}: the nominal voltage in the RX state;
$T_{cpu_i} = \frac{cpu_i\ (\#\ ticks)}{RTIMER\_ARCH\_SECOND} = \frac{cpu_i\ (\#\ ticks)}{32{,}768}$
$T_{lpm_i} = \frac{lpm_i\ (\#\ ticks)}{RTIMER\_ARCH\_SECOND} = \frac{lpm_i\ (\#\ ticks)}{32{,}768}$
$T_{tx_i} = \frac{tx_i\ (\#\ ticks)}{RTIMER\_ARCH\_SECOND} = \frac{tx_i(\#\ ticks)}{32{,}768}$
$T_{rx_i} = \frac{rx\ _i(\#\ ticks)}{RTIMER\_ARCH\_SECOND} = \frac{rx_i\ (\#\ ticks)}{32{,}768}$

Based on Equation (1) and Table 2 that provides the typical operating conditions for a Tmote Sky mote, the total energy consumption, at the reading (i.e., record) i, is given by Equation (2):

$$\begin{aligned} E_{total_i}(mj) &= 1.8 \times 3 \times \left(\frac{cpu_i\ (\#\ ticks)}{32{,}768}\right) \\ &+ 0.0545 \times 3 \times \left(\frac{lpm_i\ (\#\ ticks)}{32{,}768}\right) \\ &+ 19.5 \times 3 \times \left(\frac{tx_i(\#\ ticks)}{32{,}768}\right) \\ &+ 21.8 \times 3 \times \left(\frac{rx_i\ (\#\ ticks)}{32{,}768}\right) \end{aligned} \quad (2)$$

Table 2. Typical Operating Conditions for Tmote Sky motes.

	MIN	NOM (Typical)	MAX	UNIT
Supply voltage	2.1	3.0	3.6	V
Supply voltage during flash memory programming	2.7	3.0	3.6	V
Operating free air temperature	−40		85	ºC
Current Consumption: MCU on, Radio RX		21.8	23	mA
Current Consumption: MCU on, Radio TX		19.5	21	mA
Current Consumption: MCU on, Radio off		1800	2400	µA
Current Consumption: MCU idle, Radio off		54.5	1200	µA
Current Consumption: MCU standby		5.1	21.0	µA

Based on Equation (2) and the following features, from the generated benign "powertrace" dataset, for each mote: (a) all_cpu; (b) all_lpm; (c) all_transmit; and (d) all_listen, the total energy consumption by each mote, during the simulation time (i.e., 60 min = 3600 s) is shown below in Figure 37.

On the other hand, based on Equation (2) and the same features (i.e., all_cpu, all_lpm, all_transmit; and all_listen) for each mote, from the generated malicious "powertrace" dataset, the total energy consumption by each mote, during the simulation time (i.e., 60 min = 3600 s) is shown below.

As shown in Figure 38, mote6, which is the compromised client that carried out the UDP flooding attack, consumed much more energy than any other legitimate client and the legitimate server in the UDP flooding attack scenario. Moreover, mote6 in the UDP flooding attack consumed much more energy than the energy it consumed in the benign scenario as demonstrated in Figure 37.

Furthermore, the generated benign and malicious network traffic datasets, presented in Sections 4.2.2 and 5.2.2, respectively, include information about raw features, such as source/destination address, protocol, which can be used to derive new features more informative, in terms of the behaviour of the network traffic, and non-redundant. These new features are also intended to constitute valuable features for training and evaluating AIDS for IoT/IIoT networks. From the network traffic point of view, the total RPL (Routing Protocol for Low-Power and Lossy Networks) messages overhead of the IoT/IIoT network can be considered as a feature for detection of a UDP flooding attack as an IoT/IIoT network under a UDP flooding attack is characterised by low total RPL messages overhead because of the huge amount of the UDP messages flooding the network, as shown below.

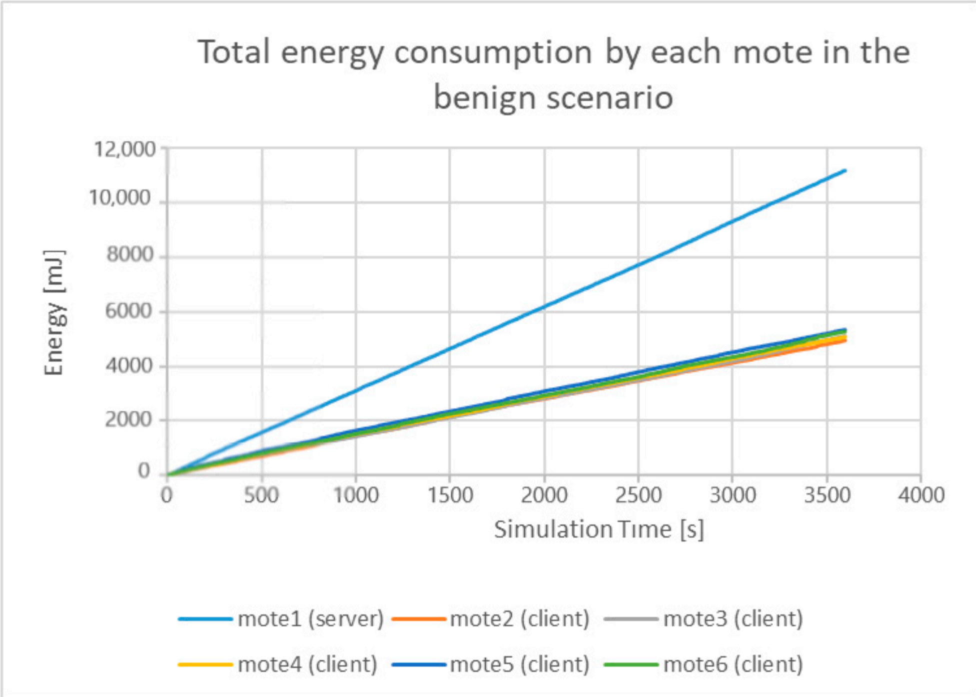

Figure 37. Total energy consumption by each mote in the benign scenario.

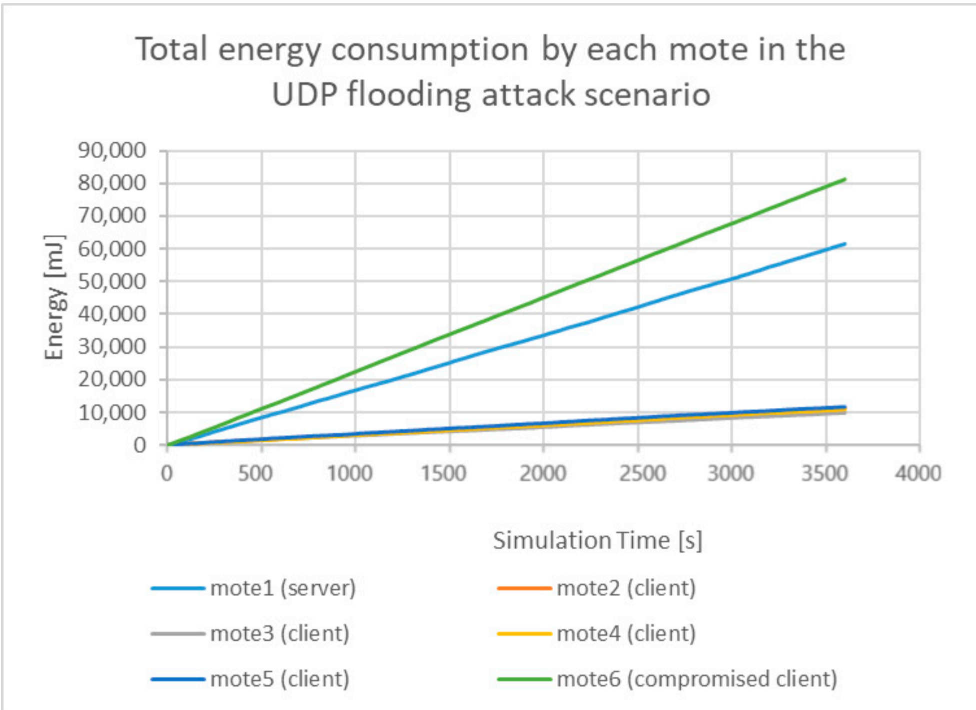

Figure 38. Total energy consumption by each mote in the UDP flooding attack scenario.

Table 3 was extracted from the benign network traffic dataset (i.e., benign "radiolog.csv") and shows, in the last column, the percentage of the RPL messages overhead per mote which is calculated as follows: the number of RPL messages per mote over the total number of exchanged messages within the network during the simulation time (i.e., 116,463 messages). The last row of Table 3 contains the total number of RPL messages (7975), UDP messages (104,048), and other protocol messages (4440) exchanged within the network, and the total RPL messages overhead (%).

Table 3. RPL messages overhead of the IoT/IIoT network in the benign scenario.

	RPL Messages Overhead			
	Number of RPL Messages	**Number of UDP Messages**	**Number of Other Messages**	**RPL Overhead (%)**
Mote 1	290	43,804	N/A	0.25
Mote 2	1982	11,621	N/A	1.70
Mote 3	1621	11,883	N/A	1.39
Mote 4	1604	11,827	N/A	1.38
Mote 5	1308	12,556	N/A	1.12
Mote 6	1170	12,357	N/A	1.00
Total	7975	104,048	4440	6.85

Based on the information included in Table 3, the calculated RPL messages overhead per mote and the total RPL messages overhead are depicted in Figure 39.

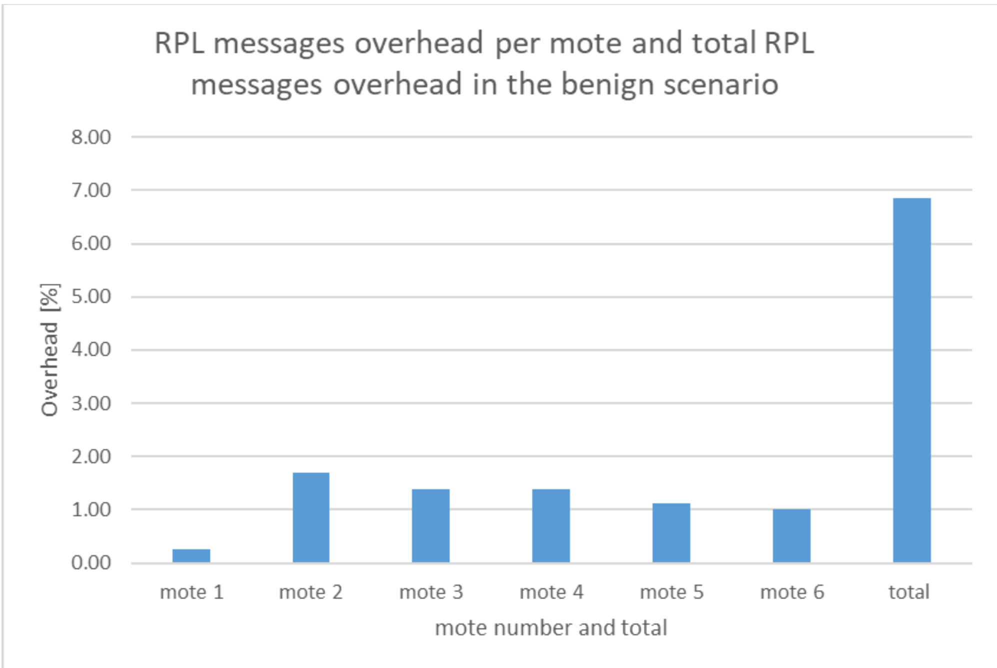

Figure 39. RPL messages overhead per mote and total RPL messages overhead in the benign scenario.

On the other hand, Table 4 was extracted from the malicious network traffic dataset (i.e., malicious "radiolog.csv") reflecting the UDP flooding attack scenario. Similar to Table 3, Table 4 shows, in the last column, the percentage of the RPL messages overhead per mote which is calculated as follows: the number of RPL messages per mote over the total number of exchanged messages within the network during the simulation time (i.e., 702,332 messages). The last row of Table 4 contains the total number of RPL messages (9908), UDP messages (670,671), and other protocol messages (21,753) exchanged within the network, and the total RPL messages overhead (%).

Table 4. RPL messages overhead of the IoT/IIoT network in the benign scenario.

RPL Messages Overhead				
	Number of RPL Messages	Number of UDP Messages	Number of Other Messages	RPL Overhead (%)
Mote 1	203	254,796	N/A	0.03
Mote 2	2228	28,953	N/A	0.32
Mote 3	2768	30,238	N/A	0.39
Mote 4	1976	27,260	N/A	0.28
Mote 5	2084	31,247	N/A	0.30
Mote 6	6490	298,177	N/A	0.09
Total	9908	670,671	21,753	1.41

Based on the information included in Table 4, the calculated RPL messages overhead per mote and the total RPL messages overhead are depicted in Figure 40.

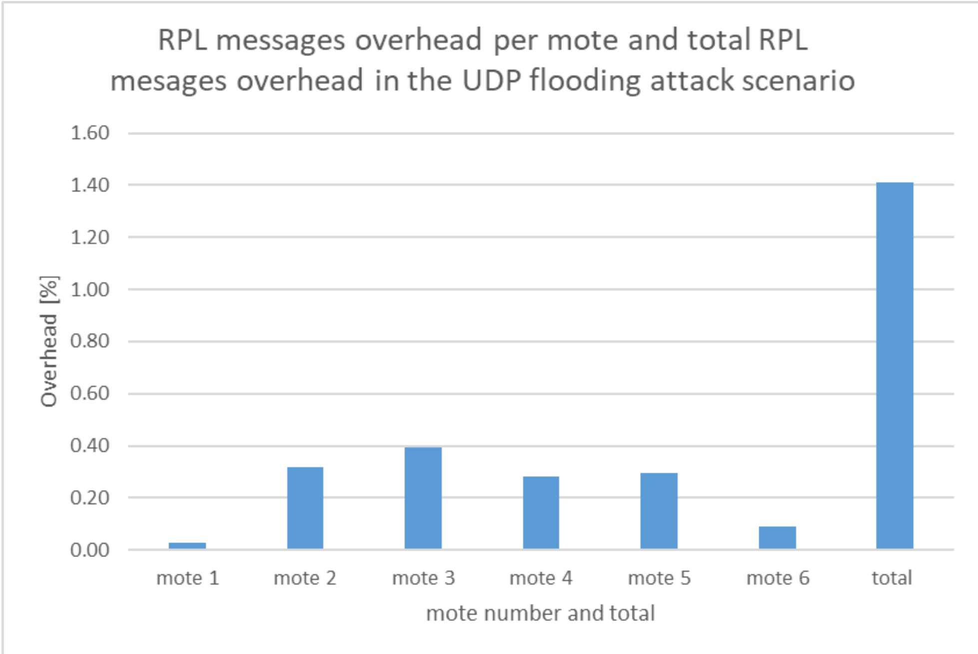

Figure 40. RPL messages overhead per mote and total RPL messages overhead in the malicious scenario.

As shown in Figures 39 and 40, the total RPL messages overhead (1.41%) in the malicious scenario is much less than the total RPL messages overhead in the benign scenario (6.85%) because of the huge amount of the UDP messages flooding the network in the malicious scenario.

7. Conclusions

Due to the urgent need for up-to-date, representative and well-structured IoT/IIoT-specific datasets which are publicly available and constitute benchmark datasets for training and evaluating ML models used in AIDSs for IoT/IIoT networks, we target the generation of new labelled IoT/IIoT datasets that will be publicly available to the research community and include (i) events reflecting multiple benign and attack scenarios from current IoT/IIoT network environments, (ii) sensor measurement data, (iii) network-related information (e.g., packet-level information and flow-level information) from the IoT/IIoT network, and (iv) information related to the behaviour of the IoT/IIoT devices deployed within the IoT/IIoT network. In this context, this paper we presented an initial set of datasets with these significant characteristics for effective training and testing of ML models used in AIDSs for protecting IoT/IIoT networks. In particular, the provided set of datasets consists of (a) benign IoT/IIoT datasets (i.e., around 11,000 records of the benign "powertrace" dataset and around 116,000 records of the benign network traffic dataset), and (b) malicious

IoT/IIoT datasets (i.e., around 11,000 records of the malicious "powertrace" dataset and around 700,000 records of the malicious network traffic dataset).

In addition, in this paper, we presented in detail the approach that we adopted to generate the initial set of benign IoT/IIoT and malicious IoT/IIoT datasets by utilising the Cooja simulator that was the simulation environment where the corresponding benign and attack scenarios were implemented. It is worthwhile to highlight that for the first time and to the best of our knowledge, that the Cooja simulator, which is the companion network simulator of Contiki OS (one of the most popular OSs for resource constrained IoT devices), was used in a systematic way in order to generate IoT/IIoT datasets. In particular, we provided a comprehensive description of the whole approach we followed in order to acquire the generated datasets within csv files from the captured raw information residing in the Cooja simulator environment. Then, the generated datasets in csv format are ready to feed ML algorithms for training and testing purposes.

Our goal is that the new labelled IoT/IIoT datasets generated by utilizing the Cooja simulator should not to be considered as a replacement of datasets captured from real IoT/IIoT networks or real IoT/IIoT testbeds, but instead to be considered as complementary datasets that will contribute to fill the gap in the lack of publicly available up-to-date, representative and well-structured IoT/IIoT-specific datasets that constitute benchmark datasets for training and evaluating ML models used in AIDSs for IoT/IIoT networks.

As future work, we plan to continue working on the implementation of more benign IoT/IIoT network scenarios and various types of IoT/IIoT network attack scenarios, with more motes, in Cooja simulator in order to generate richer benign and malicious datasets for more effective training and testing of ML algorithms used in AIDSs for protecting IoT/IIoT networks such as the one described in [31]. Our intention is to make the generated rich datasets publicly available to the research community. In addition, we will also make publicly available the Cooja-based framework that will have been developed in order to generate the rich datasets. This will allow researchers to reproduce datasets as well as generate new datasets for their own scenarios without having to "reinvent the wheel". Furthermore, we intend to analyse the generated datasets to select the most appropriate features for accurate and efficient detection of different types of attacks within an IoT/IIoT network. Finally, we plan to apply a number of common ML algorithms (e.g., support vector machines (SVMs), Naïve Bayes, k-nearest neighbour, logistics regression, etc.) to evaluate their performance on the new generated datasets when these algorithms are used for anomaly detection in AIDSs.

Author Contributions: Conceptualization and methodology, G.M., J.C.R., and I.E.; software, J.C.R. and I.E.; validation, J.C.R. and G.M.; investigation, I.E., J.C.R., and G.M.; resources, I.E., J.C.R., and M.P.; writing—original draft preparation, I.E., M.P., and G.Z.; writing—review and editing, I.E., G.M., and J.R.; visualization, J.C.R. and I.E.; supervision, G.M. and J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The research work leading to this publication has received funding through the Moore4Medical project under grant agreement H2020-ECSEL-2019-IA-876190 within ECSEL JU in collaboration with the European Union's H2020 Framework Programme (H2020/2014-2020) and Fundação para a Ciência e Tecnologia (ECSEL/0006/2019).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Xu, L.D.; He, W.; Li, S. Internet of Things in Industries: A Survey. *IEEE Trans. Ind. Informatics* **2014**, *10*, 2233–2243. [CrossRef]
2. Zarpelão, B.B.; Miani, R.S.; Kawakani, C.T.; de Alvarenga, S.C. A survey of intrusion detection in Internet of Things. *J. Netw. Comput. Appl.* **2017**, *84*, 25–37. [CrossRef]

3. Sisinni, E.; Saifullah, A.; Han, S.; Jennehag, U.; Gidlund, M. Industrial Internet of Things: Challenges, Opportunities, and Directions. *IEEE Trans. Ind. Informatics* **2018**, *14*, 4724–4734. [CrossRef]
4. Alsaedi, A.; Moustafa, N.; Tari, Z.; Mahmood, A.; Anwar, A. TON_IoT Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-Driven Intrusion Detection Systems. *IEEE Access* **2020**, *8*, 165130–165150. [CrossRef]
5. Chaabouni, N.; Mosbah, M.; Zemmari, A.; Sauvignac, C.; Faruki, P. Network Intrusion Detection for IoT Security Based on Learning Techniques. *IEEE Commun. Surv. Tutorials* **2019**, *21*, 2671–2701. [CrossRef]
6. KDD Cup 1999 Data. Available online: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (accessed on 19 September 2020).
7. Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA 2009, Ottawa, ON, Canada, 8–10 July 2009; pp. 1–6.
8. Moustafa, N.; Slay, J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Proceedings of the 2015 Military Communications and Information Systems Conference, MilCIS 2015, Canberra, ACT, Australia, 10–12 November 2015; pp. 1–6.
9. Sharafaldin, I.; Lashkari, A.H.; Ghorbani, A.A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In Proceedings of the ICISSP2018, Funchal, Madeira, Portugal, 22–24 January 2018; pp. 108–116.
10. Suthaharan, S.; Alzahrani, M.; Rajasegarar, S.; Leckie, C.; Palaniswami, M. Labelled data collection for anomaly detection in wireless sensor networks. In Proceedings of the 2010 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, ISSNIP 2010, Brisbane, QLD, Australia, 7–10 December 2010; pp. 269–274.
11. Sivanathan, A.; Gharakheili, H.H.; Loi, F.; Radford, A.; Wijenayake, C.; Vishwanath, A.; Sivaraman, V. Classifying IoT Devices in Smart Environments Using Network Traffic Characteristics. *IEEE Trans. Mob. Comput.* **2019**, *18*, 1745–1759. [CrossRef]
12. 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. *Futur. Gener. Comput. Syst.* **2019**, *100*, 779–796. [CrossRef]
13. Hamza, A.; Gharakheili, H.H.; Benson, T.A.; Sivaraman, V. Detecting Volumetric Attacks on IoT Devices via SDN-Based Monitoring of MUD Activity. In Proceedings of the SOSR 2019—Proceedings of the 2019 ACM Symposium on SDN Research, San Jose, CA, USA, 3–4 April 2019; Association for Computing Machinery, Inc: New York, NY, USA, 2019; pp. 36–48.
14. Österlind, F.; Dunkels, A.; Eriksson, J.; Finne, N.; Voigt, T. Cross-level sensor network simulation with COOJA. In Proceedings of the Proceedings—Conference on Local Computer Networks, LCN, Tampa, FL, USA, 14–16 November 2006; pp. 641–648.
15. ITU-T. Recommendation ITU-T Y.2060 "Overview of the Internet of Things". 2012. Available online: https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=y.2060 (accessed on 15 December 2020).
16. Qi, Q.; Tao, F. A Smart Manufacturing Service System Based on Edge Computing, Fog Computing, and Cloud Computing. *IEEE Access* **2019**, *7*, 86769–86777. [CrossRef]
17. Lin, J.; Yu, W.; Zhang, N.; Yang, X.; Zhang, H.; Zhao, W. A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet Things J.* **2017**, *4*, 1125–1142. [CrossRef]
18. Ferrag, M.A.; Maglaras, L.; Argyriou, A.; Kosmanos, D.; Janicke, H. Security for 4G and 5G cellular networks: A survey of existing authentication and privacy-preserving schemes. *J. Netw. Comput. Appl.* **2018**, *101*, 55–82. [CrossRef]
19. Makhdoom, I.; Abolhasan, M.; Lipman, J.; Liu, R.P.; Ni, W. Anatomy of Threats to the Internet of Things. *IEEE Commun. Surv. Tutorials* **2019**, *21*, 1636–1675. [CrossRef]
20. Hassija, V.; Chamola, V.; Saxena, V.; Jain, D.; Goyal, P.; Sikdar, B. A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Access* **2019**, *7*, 82721–82743. [CrossRef]
21. Newsome, J.; Shi, E.; Song, D.; Perrig, A. The Sybil attack in sensor networks: Analysis & defenses - IEEE Conference Publication. In Proceedings of the Third International Symposium on Information Processing in Sensor Networks, Berkeley, CA, USA, 27 April 2004.
22. El-hajj, M.; Fadlallah, A.; Chamoun, M.; Serhrouchni, A. A Survey of Internet of Things (IoT) Authentication Schemes. *Sensors* **2019**, *19*, 1141. [CrossRef] [PubMed]
23. Frustaci, M.; Pace, P.; Aloi, G.; Fortino, G. Evaluating critical security issues of the IoT world: Present and future challenges. *IEEE Internet Things J.* **2018**, *5*, 2483–2495. [CrossRef]
24. Moustafa, N.; Turnbull, B.; Choo, K.K.R. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet Things J.* **2019**, *6*, 4815–4830. [CrossRef]
25. Clarence, C.; David, F. *Machine Learning and Security [Book]*; O'Reilly Media, Inc.: Newton, MA, USA, 2018.
26. Hodo, E.; Bellekens, X.; Hamilton, A.; Dubouilh, P.L.; Iorkyase, E.; Tachtatzis, C.; Atkinson, R. Threat analysis of IoT networks using artificial neural network intrusion detection system. In Proceedings of the 2016 International Symposium on Networks, Computers and Communications, ISNCC 2016, Yasmine Hammamet, Tunisia, 11–13 May 2016.
27. Moteiv Corporation Tmote Sky—Ultra Low Power IEEE 802.15.4 Compliant Wireless Sensor Module. 2006. Available online: http://www.crew-project.eu/sites/default/files/tmote-sky-datasheet.pdf (accessed on 5 December 2020).
28. Wireshark Go Deep. Available online: https://www.wireshark.org/ (accessed on 28 November 2020).
29. Amirinasab Nasab, M.; Shamshirband, S.; Chronopoulos, A.; Mosavi, A.; Nabipour, N. Energy-Efficient Method for Wireless Sensor Networks Low-Power Radio Operation in Internet of Things. *Electronics* **2020**, *9*, 320. [CrossRef]

30. Bandekar, A.; Javaid, A.Y. Cyber-attack Mitigation and Impact Analysis for Low-power IoT Devices. In Proceedings of the 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, CYBER 2017, Honolulu, HI, USA, 31 July–4 August 2018; pp. 1631–1636.
31. Amir Alavi, S.; Rahimian, A.; Mehran, K.; Alaleddin Mehr Ardestani, J. An IoT-Based Data Collection Platform for Situational Awareness-Centric Microgrids. In Proceedings of the Canadian Conference on Electrical and Computer Engineering, Quebec City, QC, Canada, 13–16 May 2018; Volume 2018.