City Research Online

City St George's, University of London

Citation: Bobrovnikova, K., Lysenko, S., Popov, P. T., Denysiuk, D. & Goroshko, A. (2021). Technique for IoT cyberattacks detection based on the energy consumption analysis. CEUR Workshop Proceedings, 2853, ISSN 1613-0073

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/26372/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Technique for IoT Cyberattacks Detection Based on the Energy Consumption Analysis

Kira Bobrovnikova[a], Sergii Lysenko[a], Peter Popov[b], Dmytro Denysiuk[a] and Andrii Goroshko[a]

[a] *Khmelnytskyi National University, Institutska str., 11, Khmelnytskyi, 29016, Ukraine*
[b] *City University of London, Northampton Square, London EC1V 0HB, United Kingdom*

Abstract

Abstract – Today Smart Home is a system for managing the basic life support processes of both small systems (commercial, office premises, apartments, cottages) and large automated complexes (commercial and industrial complexes). One of the important tasks to be solved by the concept of a modern Smart Home is the problem of preventing the malware spread and the usage of IoT infrastructure. One of the possible approaches for abnormal behavior of the IoT devices and IoT cyberattack detection is the monitoring of the energy consumption.

Thus, an effective control and monitoring of heating, ventilation, air conditioning, more efficient use of traditional appliances and the introduction of energy-efficient equipment in the building are important to ensure and decision making in the terms of cybersecurity. In addition, improving the efficiency of energy management and monitoring is the approach to increasing effectiveness of the IoT cyberattack detection in the IoT infrastructure.

The paper presents a technique for IoT attacks detection based on the IoT devices energy consumption analysis, which take into account the energy consumption related user's preference modes. With aim to improve the accuracy of IoT cyberattacks detection and localize the IoT malware on these IoT devices the IoT software opcodes sequences analysis is applied. The proposed approach allows detecting the performing of the IoT devices such attacks, for example, as DoS/DDoS with high efficiency, at a level of about 99.88% and localizing malicious IoT software on these devices with accuracy of about 99.66%.

Keywords

Internet of things, cyberattack, DDoS, malware detection, energy consumption, sequential pattern mining, opcodes analysis.

1. Introduction

The Internet of Things and Smart Home conception have become an important part of modern society. In the other hand, the growing number of IoT devices, which are often released without any security features, makes them a desirable target for cybercriminals [1, 2].

Unprotected IoT devices join the ranks of botnets that are most often used to launch DDoS attacks or as VPN exit nodes. Cryptomining is another popular way to monetize compromised IoT devices. Since the limited battery capacity of smartphones does not allow them to be used for profit, that smart TVs, set-top boxes and other IoT devices are popular with cybercriminals. Almost any smart IoT device connected to the Internet, for example, gas, water and electricity meters, can become objects of interest for cybercriminals [1].

IntelITSIS'2021: 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 24–26, 2021, Khmelnytskyi, Ukraine
EMAIL: bobrovnikova.kira@gmail.com (K. Bobrovnikova); sirogyk@ukr.net (S. Lysenko); p.t.popov@city.ac.uk (P. Popov); web.developer.den@gmail.com (D. Denysiuk); iftomm@ukr.net (A. Goroshko)
ORCID: 0000-0002-1046-893X (K. Bobrovnikova); 0000-0001-7243-8747 (S. Lysenko); 0000-0002-3434-5272 (P. Popov); 0000-0002-1386-2326 (A. Goroshko)
© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-2853/paper54.pdf

According to forecasts of the GSMA [3], by 2025 the number of connected IoT devices will double and reach almost 25 billion worldwide, and as the popularity of IoT increases, the risk of cyberattacks will increase.

Today, the efficient use of energy resources is another one of the most important tasks. At the same time, almost a third of the total energy consumption is made up of certain losses, i.e. the energy is consumed not on purpose [4]. Further growth in energy consumption is also expected. Increasing attention to the problems of energy efficiency and energy saving also contributes to the development of the concept of a modern smart home. Furthermore, if at first this concept was to connect sensors, devices and devices over a network for the purpose of remote monitoring, access and control of the living environment and provide the necessary services to users, then at the present stage it also involves the optimal use of energy in buildings, as well as the malware and IoT cyberattack detection in Smart Home infrastructure.

IoT devices energy consumption monitoring is a possible way to detect those performing attacks, which require significant energy consumption [5], for example DDoS and cryptomining. In addition, energy consumption analysis based approach is more secure in cases the kernel of the device is already compromised, so far as once the device is compromised, the data integrity cannot be guaranteed.

2. Related works

Today, scientific sources widely present various approaches aimed at ensuring energy efficiency and energy saving in the smart home system [6-12]. In [13] it is noted that in recent years the main direction of energy efficiency policy has been to promote the use of more efficient appliances and components. However, home automation control plays an important role in efficient and sustainable operation: (1) by identifying and eliminating energy losses; (2) by using energy only in the right amount, place and only at the time when it is needed; (3) by exercising correct control of the functional level of the system for correct application in the right place.

Today, there is also a shortage of operating systems that would provide the ability to integrate the devices that make up the smart home environment. The problem stems from the fact that smart devices are based on self-service modules and use independent IoT platforms developed by various manufacturers. This leads to the need to control each device separately, which reduces the energy efficiency of the home and increases the amount of traffic on the network. To solve this problem, an integrated control system is proposed in [14], which combines IoT devices into a single system.

In [15] an analysis of the benefits and risks of smart home technologies from different points of view is carried out. One of the risks is the lack of attention of developers in the field of smart home technologies to measures to increase consumer confidence in data security and privacy.

In [16] a various approaches to intelligent control of home systems in order to reduce energy consumption are considered. One such approach is feed-forward control. Such a system directly compensates for interference factors such as external temperature, wind, solar radiation, internal heat gain by measuring interference factors in real time to implement appropriate measures based on known parameters.

Another approach is model-based predictive control (MPC) [16], which is a structured approach that predicts future system behavior based on models and adjusts the system accordingly. Fuzzy logic control does not require a complex mathematical model to control the system and can be based directly on the quality user experience. The disadvantage of this approach is the complexity of determining the optimal rules and membership functions for such systems [16].

Another well-known approach for building control systems for a home is artificial neural networks (ANNs), which are widely used to model and predict energy use in buildings. Artificial neural networks are capable of simulating non-linear processes, constantly adapting to new data and learning from this data in order to solve complex problems [16].

Also known are hybrid approaches based on the use of fuzzy logic and artificial neural networks, combining the advantages of both approaches - imitation of human logic and the ability to learn. Adaptive neuro-fuzzy (ANF) systems implement neural network learning algorithms for tuning membership functions in a fuzzy system. In a control system based on agents, which are virtual or

physical modules, agents cooperate with the environment by perceiving and influencing parameters using artificial intelligence [16]. Such systems are able to balance energy consumption, cost and comfort by measuring and interacting with the environment and controlling heating, ventilation and air conditioning systems and electrical appliances.

The study [17] analyzed known home energy management systems in order to identify key differences in their functionality and quality, and identified opportunities for energy savings (both behavioral and operational). It is also noted that in many cases, potential benefits related to convenience, comfort or safety can limit the implementation of energy saving scenarios.

Also are known a number of approaches based on monitoring the IoT devices energy consumption devoted to detecting IoT cyberattacks. In the [18] quantitatively studied the impact of DDoS and E-DDoS attacks on smart home IoT devices and on them energy consumption and the underlying reasons for these devices' various response types were analyzed.

In [19] a machine learning based method wish allows to detect ransomware attacks by monitoring energy consumption patterns for different processes of Android devices was presented.

In the paper [20] a dynamic technique to detect malware on Android platform was proposed. This technique uses a set of 38 energy related features belonging to three different categories: CPU, Memory and Network, which can be symptomatic of abnormal battery consumption.

The paper [21] is focused on malware detection using power consumption and network traffic data collected. With this aim seven power-based and eighteen network traffic-based features were applied.

In the work [22] an IoT attack detection framework based on energy consumption analysis was proposed. The proposed framework processes the energy consumption of IoT devices and classifies the attack status (not only cyberattacks, but also physical attacks) of the monitored devices. A two-stage strategy is proposed: applying a short time window for rough attack detection, and a long time window to the fine attack detection.

Nonetheless, in the paper [5] energy consumption analysis approaches were evaluated and concluded that these approaches are not applicable to such devices as, for examples, smartphones. This is due to the fact that the typical energy consumption of such devices is varies quite a lot in practice, as well due to the noise introduced into the system by unpredictable user and environment interactions. These nuances will lead to a lot of false alarms. Also empirical tests were conducted and they showed that the additional power consumed by both artificial and real-world malicious applications is too small to be detectable with the mean error rates of state-of-the art measurement tools. However, it was noted, that such attacks as DDoS can be detected by analyzing the energy consumption of similar devices.

IoT devices total energy consumption monitoring cannot provide an answer to the question of localizing malware as a source of IoT cyberattack. One of the possible approaches to identifying suspicious programs with aim its localization is to analyze programs opcodes.

In [23] an approach based on analysis opcode N-gram sequences to classifying ransomware was proposed. To select feature N-grams Term frequency-Inverse document frequency (TF-IDF) for each of them is calculated. Of the TF values of the feature N-grams the feature vectors are constructed and by machine-learning methods are processed to perform ransomware classification.

In [24] a deep learning based technique for Internet Of Battlefield Things malware detection which uses class-wise selection of opcodes sequence as a feature for classification task was presented. The opcodes are transmuted into a vector space and a graph of selected features was created for each sample. To classify malicious and benign application a deep Eigen space learning approach was applied.

In the paper [25] combining sequential pattern mining algorithm with machine learning techniques to detect most frequent opcodes sequences of malicious IoT applications was applied.

In [26] a multi-view learning method that uses multiple views including opcodes, bytecodes, header information, permission, attacker's intent and API call to detection malware. With aim to detection optimization in different environment the proposed system automatically assigns different weights to these views.

In the paper [27] a malware detection approach based on the opcodes analysis by using the evolutionary algorithm. According this approach the label of suspicious instance is defined based on the most similar graph obtained from the evolutionary algorithm with each family of malware and benign applications.

Despite the large number of different developed methods for detecting and preventing cyberattacks and malware, as well as newel data analysis approaches [28-44], IoT devices is still incredibly vulnerable and suffered a wide range of cyberattacks and their financial and public relations consequences. Therefore, there is a need to develop new approaches for the IoT malware and IoT cyberattack detecting.

3. Technique for IoT cyberattacks detection based on the energy consumption analysis

The proposed technique for IoT cyberattacks detection uses analysis of the IoT devices energy consumption footprints and also applies analysis of the IoT software opcodes sequences to improve the accuracy of IoT attacks detection and localize the IoT malware on these IoT devices.

To effectively build IoT devices energy consumption footprints, it is necessary to take into account the different energy consumption related user preference modes (let denote it as *UPM*) for HVAC systems (Heating, Ventilation, & Air Conditioning), lightening and functioning of different IoT devices. Let us denote the set of IoT user's preference modes for certain IoT device as

$$P_d = \{p_i\}_{i=1}^{N_P}, \tag{1}$$

where d – the certain IoT device, $d \in D$, $D = \{d_i\}_{i=1}^{L}$ – the set of IoT devices in the IoT network, L – the amount number of IoT devices in the network;

$p_i \in \{"very\ low", "low", "normal", "high", "very\ high"\}$ and define the energy consumption related UPM, such as temperature, lighting, humidity, air quality modes etc. and functioning modes of the different IoT devices d;

N_P – the number of UPM for certain IoT device, $N_P \geq 1$.

Also let define the energy consumption control function φ, which keep up the energy consumption of the IoT device according a given UPM as

$$\varphi : \left\{ d \mid n_d \neq n_{d,p} \right\} \rightarrow n_{d,p}, \tag{2}$$

where n_d – the current energy consumption of the IoT device d;

$n_{d,p}$ – the energy consumption of the IoT device d in certain user preference mode p.

The proposed technique consists of two stages: learning and detection stages. In turn learning stage includes the energy consumption analysis and the opcodes sequences analysis. The steps of the learning stage of the energy consumption analysis are presented following.

1. The IoT devices energy consumption footprints for different energy consumption related user's preference modes in the absence of the IoT cyberattacks building, normalization and labelling.
2. The IoT devices energy consumption footprints for different energy consumption related user's preference modes in the presence of the IoT cyberattacks building, normalization and labelling.
3. Labeled and unlabeled data matrix of the IoT devices energy consumption footprints building.
4. Semi-supervised learning of the fuzzy c-means classifier by using the labeled data matrix of the IoT devices energy consumption footprints.
5. Testing of the fuzzy c-means classifier by using unlabeled data matrix of the IoT devices energy consumption footprints.
6. Evaluating of the effectiveness of the energy consumption analysis.

As noted above, to improve the accuracy of IoT attacks detection based on energy consumption analysis and localize the IoT malware on IoT devices opcodes sequences analysis is applied. The steps of the learning stage of the opcodes sequences analysis are presented following.

1. The assembly representation extraction from the benign and malicious IoT binary executable.

2. The opcodes maximal sequential patterns (MSP) mining in the all binary executable assembly representation.
3. The MSP selection for constructing feature vectors.
4. The relevance calculation for each of the selected MSP.
5. The feature vectors of opcodes MSP relevance construction and labelling for each binary executable assembly representation.
6. Labeled and unlabeled data matrix of the feature vectors of MSP building.
7. Semi-supervised learning of the fuzzy c-means classifier by using the labeled data matrix of the feature vectors of MSP.
8. Testing of the fuzzy c-means classifier by using unlabeled data matrix of the feature vectors of MSP.
9. Evaluating of the effectiveness of the opcodes sequences analysis.

On the detection stage of the proposed technique the energy consumption of IoT devices is measured and analyzed. If the IoT device has an abnormally high energy consumption, this may indicate that it has carried out cyberattacks. But IoT device total energy consumption monitoring cannot provide an answer to the question of localizing malware on this device as a source of IoT cyberattack. Therefore, it is necessary to analyze suspicious software on the IoT device in order to localize it. So identifying the suspicious programs with aim its localization software opcodes sequences analysis was performed. With this aim opcodes maximal sequential patterns, MSP, mining in the assembly representation of suspicious binary executable is performed. For obtained MSP their relevance calculated and feature vectors for these suspicious software is built and analyzed.

The scheme of the technique for IoT attacks detection based on the energy consumption analysis presented in Fig. 1.

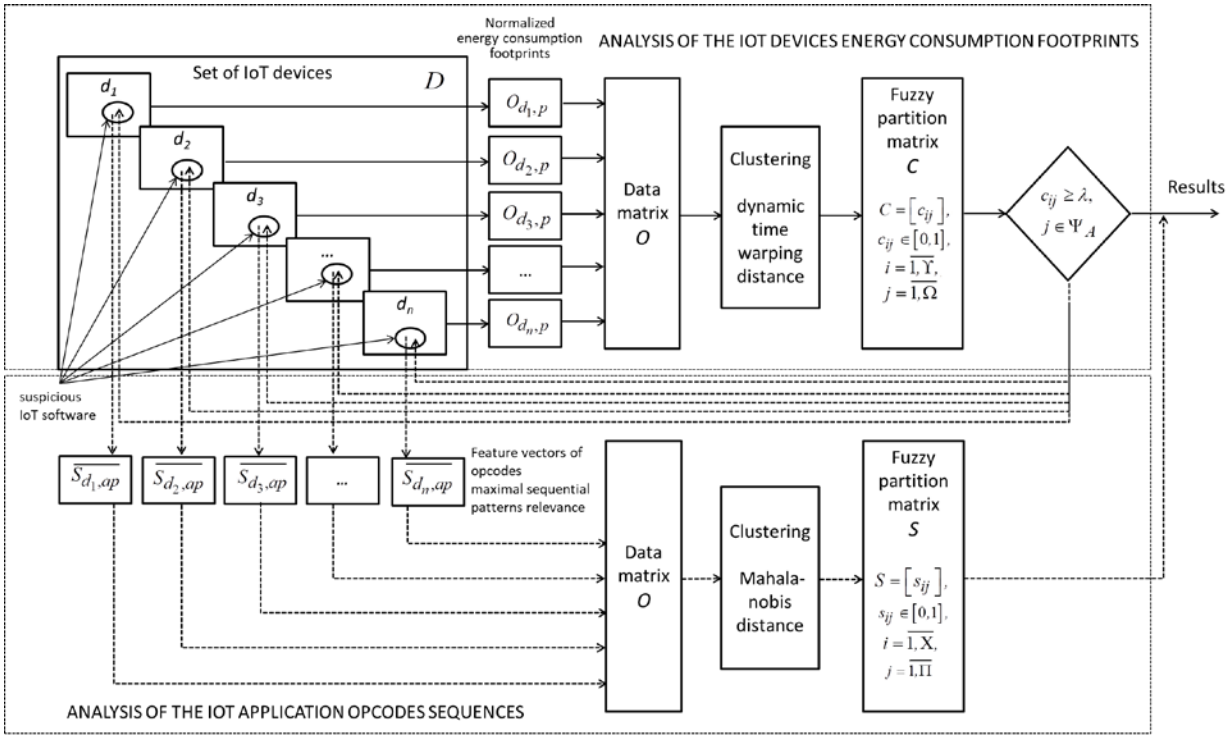

Figure 1: Scheme of the technique for IoT attacks detection based on the energy consumption analysis

Let us consider the main steps of the learning stage of technique for IoT cyberattacks detection based on the energy consumption analysis.

3.1. Energy consumption footprints building

With aim the IoT cyberattacks detection at the learning stage the energy consumption of each IoT device in the IoT network for different IoT UPM in the absence of IoT cyberattacks is measured at a certain interval and at equal sub-intervals of time. Based on these measurements, the set of IoT devices energy consumption footprints $N_{d,p}$ are constructed, part of them labelled as "normal" footprints and entered into the labeled data matrix D_l, rest of them entered on the unlabeled data matrix D_{unl} (Fig. 2).

Let us describe the energy consumption footprints in the absence of IoT cyberattacks taking into account the set of UPM as

$$N_{d,p} = (n_{d,p,i})_{i=1}^{K}, \tag{3}$$

where $n_{d,p,i}$ – the normalized measurement of the whole IoT device d energy consumption at a point in time in the absence of IoT cyberattacks for IoT user's preference mode p, $n_{d,p,i} \in [0,1]$, where 0 indicates lack of energy consumption and 1 presents the maximum of energy consumption in the absence of IoT cyberattacks;

K – the number of measurements in the time interval.

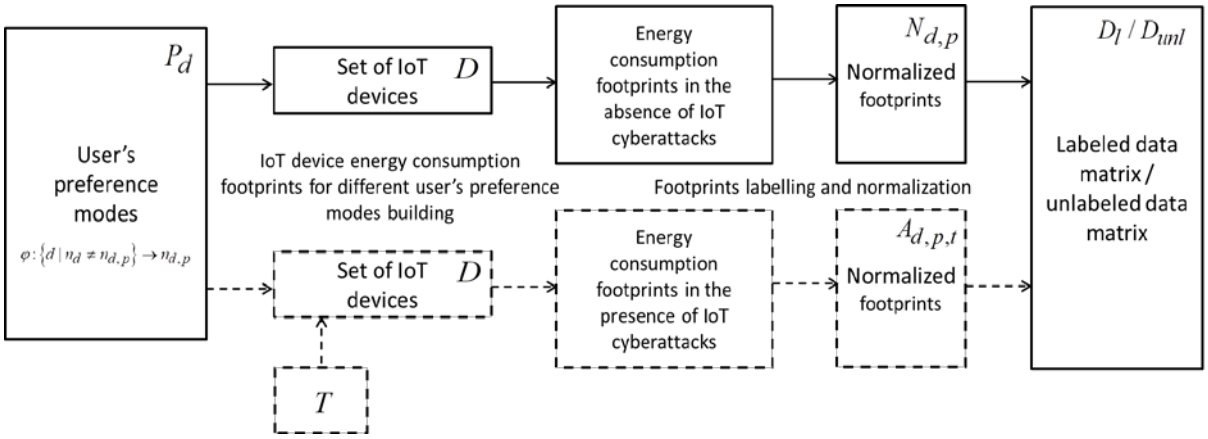

Figure 2: IoT devices energy consumption footprints for different UPM in the absence and presence IoT cyberattack building

Also the IoT devices energy consumption footprints in the presence different types of IoT cyberattacks with taking into account different IoT UPM should be built. With this aim these IoT devices were infected with malicious IoT software, which were able to carry out these types of IoT cyberattacks. After that the energy consumption of each IoT device for different IoT UPM in the presence of IoT cyberattacks is measured at a certain interval and at equal sub-intervals of time. Then based on these measurements, the set of IoT devices energy consumption footprints in the presence of IoT cyberattacks $A_{d,p,t}$ are constructed, part of them labelled as energy consumption footprints for certain type of IoT cyberattacks and entered into the labeled data matrix D_l, rest of them entered on the unlabeled data matrix D_{unl} (Fig. 2).

Let us describe the set of IoT devices energy consumption footprints in the presence of IoT cyberattacks taking into account the set of UPM as

$$A_{d,p,t} = (a_{d,p,t,i})_{i=1}^{K}, \tag{4}$$

where $a_{d,p,t,i}$ – the normalized measurement of IoT device energy consumption at a point in time in the presence of certain type cyberattacks, $a_{d,p,t,i} \in [0,1]$, where 0 indicates lack of energy consumption and 1 presents the maximum of energy consumption in the presence of IoT cyberattacks;
$t \in T$ – the type of IoT cyberattacks, T – the set of IoT cyberattacks type;
K – the number of measurements in the time interval.

After that the semi-supervised learning of the fuzzy c-means classifier by using the labeled data matrix D_l of energy consumption footprints $N_{d,p}$ and $A_{d,p,t}$ are performed.

The main particularities of applied classification algorithm are described below in Section 3.3.

To evaluate the effectiveness of the IoT cyberattack detection based on energy consumption testing the fuzzy c-means classifier by using unlabeled data matrix D_{unl} of energy consumption footprints was performed.

3.2. Feature vectors of opcodes MSP relevance building

With aim the IoT devices opcodes sequences analysis at the learning stage the assembly representation from the benign and malicious IoT binary executable examples are extracted. From these assembly representations opcodes MSP are extracted by applying of sequential patterns mining algorithm.

For each MSP the inverse document frequency value, IDF, which reduces the weight of commonly used MSP, is calculated as

$$IDF(MSP, Z) = \log \frac{|Z|}{|\{z_i \in Z \mid MSP \in z_i\}|}, \quad (5)$$

where $|Z|$ – the total number of the executables $z \in Z$, $Z = Z_b \cup Z_m$, were Z_b – set of benign IoT software, Z_m – set of malicious IoT software;

$|\{z_i \in Z \mid MSP \in z_i\}|$ – the number of the executables z in the set Z, in which appears MSP.

To determine the order of MSP in the feature vectors, the MSP are sorted ascending values $IDF(MSP, Z)$:

$$R = (IDF(MSP_i, Z))_{i=1}^{N_R},\ IDF(MSP_i, Z) < IDF(MSP_{i+1}, Z), \quad (6)$$

where N_R – the total number of different MSP.

To assess the MSP relevance for each MSP weighted term frequency (WTF) values [45] are calculated as following.

Weighted term frequency (WTF) is the result of weighting the term frequency, TF with the relevance of each opcode o and are computed as the product of sequence frequency and the calculate weight of every opcode o in the sequence *MSP*:

$$WTF(MSP, z) = TF(MSP, z) \times \prod_{o \in MSP} \frac{W(o)}{100}, \quad (7)$$

where $W(o)$ – the calculated weight, by means of mutual information gain, for the opcode o;
$TF(MSP, z)$ – the MSP frequency measure within the IoT software.

Term frequency, $TF(MSP, z)$, assessed the importance of a MSP within an IoT software executable and can be calculated as

$$TF(MSP, z) = \frac{f_{MSP,z}}{\sum_{MSP' \in z} f_{MSP',z}}, \quad (8)$$

where $f_{MSP,z}$ – the number of times the MSP appears in an executable z; $\sum_{MSP' \in z} f_{MSP',z}$ – the total number of opcodes sequences in the executable z.

The Mutual Information $I(F;\Psi)$, on which the calculation $W(o)$ is based, is measure of the statistical dependence of the two variables, in this case they are the single opcode o and whether or not the software was malware:

$$I(F;\Psi)=\sum_{\psi\in\Psi}\sum_{f\in F}p(f,\psi)\log\left(\frac{p(f,\psi)}{p(f)\times p(\psi)}\right), \quad (9)$$

where F – the opcode frequency;
Ψ – the class of the file;
$p(f,\psi)$ – is the joint probability distribution function of F and Ψ;
$p(f)$ and $p(\psi)$ – the marginal probability distribution functions of F and Ψ.

From obtained for each MSP WTF value for each IoT software feature vector of opcodes MSP relevance is built. In the process vectors, whose length is greater than the median length H of the all obtained vectors, are truncated. Vectors, whose length is less than H, are padded with zeros.

Let us denote feature vector of opcodes MSP relevance as

$$\overline{S_{d,ap}}=(s_{d,ap,i})_{i=1}^{H}, \quad (10)$$

where $s_{d,ap,i}$ – the IoT software opcodes MSP relevance;
H – the number of IoT software opcodes MSP.

The part of constructed vectors labelled respectively as "benign" or "malicious" and entered into the labeled data matrix W_l, rest of them entered on the unlabeled data matrix W_{unl} (Fig. 3).

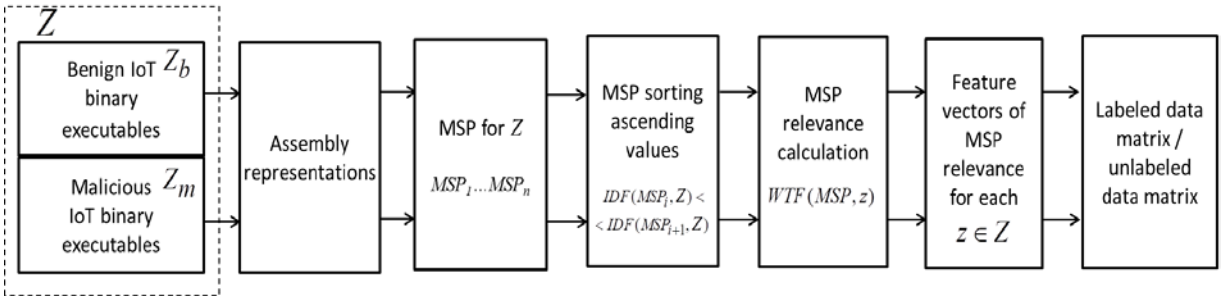

Figure 3: Feature vectors of opcodes MSP relevance building

3.3. Data classification

With aim to detect a IoT cyberattacks in proposed approach the semi-supervised fuzzy c-means classifier was applied. The advantage of the using fuzzy clustering is the weakening of the requirement for unambiguous clustering of objects, it becomes possible due to the applying of membership functions to the fuzzy clusters, that take values in the interval $[0,1]$. This allows increasing the accuracy and information completeness of the clustering results in cases where clustering objects are located at the boundaries of the clusters.

The applying of semi-supervised learning allows specifying the initial centers of clusters, which improves the quality of clustering results. The initial centers of the clusters were determined on the basis of a training sample, the volume of which was 10% of the data collected for analysis.

As clustering objects are the IoT devices energy consumption footprints. But instead of the Euclidean distance, which is used in the basic c-means algorithm, for IoT devices energy consumption footprints clustering as a distance measure the dynamic time warping, DTW, was applied.

The use of Euclidean distance has a significant drawback: if two time series are the same, but one of them is slightly displaced in time (along the time axis), then the Euclidean metric may consider that the series are different from each other.

The DTW algorithm was introduced in order to overcome this disadvantage and provide a measurement of the distance between rows, without paying attention to both global and local shifts on the timeline.

The result of the IoT devices energy consumption footprints clustering is a fuzzy partition matrix C, where each element of the matrix c_{ij} determines the degree of belonging of the i-th element (the energy consumption footprint of IoT device) to the j-th cluster: $C=\left[c_{ij}\right], c_{ij}\in[0,1], i=\overline{1,\Upsilon}, j=\overline{1,\Omega}$,

$\sum_{j=\overline{1,\Omega}} c_{ij} = 1$, where Υ – the number of energy consumption footprints, Ω – the number of the clusters. Thus, each clustering objects with a certain degree of affiliation belongs to each of the Ω clusters, each of which denotes normal energy consumption in a specific user mode or increased power consumption, indicating an attack.

For the feature vectors of opcodes MFP relevance classification instead of the Euclidean distance the Mahalanobis distance was used. It makes it possible to form clusters in the form of hyperelipsoids with axes oriented in arbitrary directions, which allows taking into account the possible presence of outliers in the classified data, that is, observation results that stand out from the general sample.

The result of clustering is a fuzzy partition matrix S, where each element of the matrix s_{ij} determines the degree of belonging of the i-th element of the set of clustering objects to the j-th cluster: $S = \left[s_{ij} \right], s_{ij} \in [0,1], i = \overline{1,\mathrm{X}}, j = \overline{1,\Pi}$, $\sum_{j=\overline{1,\Pi}} s_{ij} = 1$, where X – the number of the feature vectors of opcodes MFP relevance, Π – the number of the clusters. Thus, each feature vector with a certain degree of affiliation belongs to each of the Π clusters, each of which denotes benign software or certain type of malware.

Let's take λ as the threshold values of clustering object belonging to the cluster, at which the clustering object is considered as malicious. If $c_{ij} \geq \lambda$, then the clustering object belongs to a j cluster.

Also let us denote the set of all clusters as $\Psi = \Psi_N \cup \Psi_A$, where Ψ_N is a subset of clusters that correspond to benign clustering objects, Ψ_A is a subset of clusters that correspond to malicious clustering objects.

4. Experimental results

In order to assess the effectiveness of the proposed approach, a number of experiments were carried out. The ARM platform was chosen as the target IoT platform for the experiments, since it is one of the most common IoT platforms. Thus variety of ARM-based IoT devices (such as smart TVs, camcorders and routers) have been used. Also 284 corresponding samples of benign software from [46] and 297 malicious software samples [47], including 91 polymorphic malware samples were generated from these malware using the open-source polymorphic malware creation tool [48], have been used.

The IoT devices used in the experiments were infected with malicious software and were used to carry out DDoS attacks on a target on an isolated network. During the experiments, the energy consumption footprints of these IoT devices were obtained under normal operating conditions, as well as when these IoT devices carry out cyberattacks. Each energy consumption footprint was obtained by taking measurements after 0.5 s. within 3 minutes when the IoT device is performing an attack and normal operation. A total of 1253 energy consumption footprints of both in the presence of attacks and normal functioning IoT devices were built.

Also, using the proposed approach, opcodes sequences were extracted and analyzed from malicious software samples that carried out these DDoS attacks. These software samples were disassembled by using the IDA Pro [49] to obtain its opcodes. For opcodes sequences mining hash-based partition sequential pattern mining algorithm (HPSPM) [50] was used.

Some of these data (about 10%) were used for training, the rest of the data were used as testing data to assess the effectiveness of the proposed approach. For the purpose of classifying the malicious samples Support Vector Machine (SVM) [51, 52], K Nearest Neighbor (KNN), Decision Tree, Random Forest and Semi-Supervised Fuzzy C-Means [53] classifiers were applied.

In order to assess the effectiveness of the proposed approach, the following metrics were applied.

Accuracy is as a statistical measure which defined the proportion of correct predictions (both true positives and true negatives) among the total number of cases examined:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \quad (11)$$

where TP (true positive) – correctly classified malware samples;
TN (true negative) – correctly classified samples are benign;
FN (false negative) – malicious samples, erroneously classified as benign;
FP (false positive) – benign samples, erroneously classified as malicious.

Another measure of a test's accuracy is F-measure (or balanced F-score, F_1 score), which defined as the harmonic mean of precision and recall:

$$F_1 = 2 \times \frac{PREC \times REC}{PREC + REC}, \quad (12)$$

where $PREC = \frac{TP}{TP + FP}$, $REC = \frac{TP}{TP + FN}$.

The experimental results, which showed accuracy and F-measure values for IoT cyberattack detection based on energy consumption analysis and opcodes sequences analysis presented in Table 1, 2.

The results of the experiments showed a high efficiency of IoT cyberattacks detection based on the energy consumption analysis (Table 1). At the same time, as it is showed from the Table 2, the analysis of the opcodes sequences of suspicious software will allow localizing the program on the IoT device, which is the source of the IoT cyberattack, with high efficiency. As can be seen from the experimental results, the highest efficiency was achieved using Semi-Supervised Fuzzy C-Means clustering.

Table 1
Experimental results: accuracy and F-measure values for IoT cyberattack detection based on energy consumption analysis

Classifier	TP	TN	FN	FP	ACC	F_1
Decision Tree	1235	1239	18	14	98.72	98.72
K Nearest Neighbor	1236	1241	17	12	98.84	98.84
Random Forest	1242	1247	11	6	99.32	99.32
Support Vector Machine	1248	1249	5	4	99.64	99.64
Semi-Supervised Fuzzy C-Means clustering	1251	1252	2	1	99.88	99.88

Table 2
Experimental results: accuracy and F-measure values for IoT cyberattack detection based opcodes sequences analysis

Classifier	TP	TN	FN	FP	ACC	F_1
Random Forest	285	276	12	8	96.56	96.61
Decision Tree	292	278	5	6	98.11	98.15
K Nearest Neighbor	293	281	4	3	98.80	98.82
Support Vector Machine	295	280	2	4	98.97	98.99
Semi-Supervised Fuzzy C-Means clustering	296	283	1	1	99.66	99.66

5. Conclusions

Thus, taking into account, that high IoT device's energy consumption may indicate that the IoT device is carrying out a cyberattacks, which require increased energy consumption, a new technique for IoT attacks detection based on the IoT devices energy consumption analysis was proposed. These

technique take into account the energy consumption related user's preference modes. Therefore with aim cyberattacks detection the energy consumption of IoT devices is measured and analyzed. For the purpose of localizing the software on the IoT device that performs the cyberattacks these software opcodes analysis was performed. With this aim opcodes maximal sequential patterns, MSP, mining in the assembly representation of suspicious binary executable is performed. For obtained MSP their relevance calculated and feature vectors for this suspicious software is built and analyzed.

The experimental results show that the proposed approach allows detecting the performing by the IoT devices such attacks, as, for example, DoS/DDoS, with high efficiency, at a level of about 99.88% and localizing malicious IoT software on these devices with accuracy of about 99.66%.

6. References

[1] Trend Micro. Inside the Smart Home: IoT Device Threats and Attack Scenarios. URL: https://www.trendmicro.com/vinfo/it/security/news/internet-of-things/inside-the-smart-home-iot-device-threats-and-attack-scenarios.

[2] McAfee Labs Threats Report. URL: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-nov-2020.pdf

[3] Global System for Mobile Communications. URL: https://www.gsma.com/

[4] U.S. Energy Information Administration (EIA). International Energy Outlook, 2019. URL: https://www.eia.gov/outlooks/ieo/. – 2.07.2020 p.

[5] J. Hoffmann, S. Neumann, T. Holz, Mobile malware detection based on energy fingerprints – a dead end?. In International Workshop on Recent Advances in Intrusion Detection. Springer, Berlin, Heidelberg (2013) 348-368.

[6] A. R. Al-Ali, I. A. Zualkernan, M. Rashid, R. Gupta & M. Alikarar, A smart home energy management system using IoT and big data analytics approach. IEEE Transactions on Consumer Electronics. 2017, Vol. 63(4), pp. 426-434.

[7] M. S. Hossain, M. A. Rahman & G. Muhammad, Cyber–physical cloud-oriented multi-sensory smart home framework for elderly people: An energy efficiency perspective. Journal of Parallel and Distributed Computing. 2017, Vol. 103, pp. 11-21.

[8] M. Isnen, S. Kurniawan & E. Garcia-Palacios, A-SEM: An adaptive smart energy management testbed for shiftable loads optimisation in the smart home. Measurement. 2020, Vol. 152, 107285.

[9] M. A. Paredes-Valverde, G. Alor-Hernández, J. L. García-Alcaráz, M. D. P. Salas-Zárate, L. O. Colombo-Mendoza & J. L. Sánchez-Cervantes, IntelliHome: An internet of things-based system for electrical energy saving in smart home environment. Computational Intelligence. 2020, Vol. 36 (1), pp. 203-224.

[10] A. De Paola, P. Ferraro, G. L. Re, M. Morana & M. Ortolani, A fog-based hybrid intelligent system for energy saving in smart buildings. Journal of Ambient Intelligence and Humanized Computing. 2020, Vol. 11 (7), pp. 2793-2807.

[11] M. Killian, M. Zauner & M. Kozek, Comprehensive smart home energy management system using mixed-integer quadratic-programming. Applied energy. 2018, Vol. 222, pp. 662-672.

[12] I. Machorro-Cano, G. Alor-Hernández, M. A. Paredes-Valverde, L. Rodríguez-Mazahua, J. L. Sánchez-Cervantes & J. O. Olmedo-Aguirre, HEMS-IoT: A big data and machine learning-based smart home system for energy saving. Energies. 2020, Vol. 13 (5), 1097.

[13] V. Fabi, G. Spigliantini & S. P. Corgnati, Insights on smart home concept and occupants' interaction with building controls. Energy Procedia. 2017, Vol. 111, pp. 759-769.

[14] H. Jo, Y. I. Yoon, Intelligent smart home energy efficiency model using artificial TensorFlow engine. Human-centric Computing and Information Sciences. 2018, Vol. 8 (1), pp. 1-18.

[15] C. Wilson, T. Hargreaves, R. Hauxwell-Baldwin, Benefits and risks of smart home technologies. Energy Policy. 2017, Vol. 103, pp. 72-83.

[16] L. C. Felius, F. Dessen & B. D. Hrynyszyn, Retrofitting towards energy-efficient homes in European cold climates: a review. Energy Efficiency. 2020, Vol. 13 (1), pp. 101-125.

[17] R. Ford, M. Pritoni, A. Sanguinetti & B. Karlin, Categories and functionality of smart home technology for energy management. Building and environment. 2017, Vol. 123, pp. 543-554.

[18] B. Tushir, Y. Dalal, B. Dezfouli & Y. Liu, A Quantitative Study of DDoS and E-DDoS Attacks on WiFi Smart Home Devices. IEEE Internet of Things Journal, 2020.
[19] A. Azmoodeh, A. Dehghantanha, M. Conti & K. K. R. Choo, Detecting crypto-ransomware in IoT networks based on energy consumption footprint. Journal of Ambient Intelligence and Humanized Computing. 2018, Vol. 9 (4), pp. 1141-1152.
[20] F. Fasano, F. Martinelli, F. Mercaldo & A. Santone, Energy consumption metrics for mobile device dynamic malware detection. Procedia Computer Science. 2019, Vol. 159, pp. 1045-1052.
[21] J. H. Jimenez, K. Goseva-Popstojanova, Malware detection using power consumption and network traffic data. In 2019 2nd International Conference on Data Intelligence and Security (ICDIS). IEEE (2019) 53-59.
[22] Y. Shi, F. Li, W. Song, X. Y. Li & J. Ye, Energy audition based cyber-physical attack detection system in IoT. In Proceedings of the ACM Turing Celebration Conference-China (2019) 1-5.
[23] H. Zhang, X. Xiao, F. Mercaldo, S. Ni, F. Martinelli & A. K. Sangaiah, Classification of ransomware families with machine learning based on N-gram of opcodes. Future Generation Computer Systems. 2019, Vol. 90, pp. 211-221.
[24] A. Azmoodeh, A. Dehghantanha & K. K. R. Choo, Robust malware detection for internet of (battlefield) things devices using deep eigenspace learning. IEEE transactions on sustainable computing. 2018, Vol. 4 (1), pp. 88-95.
[25] H. Darabian, A. Dehghantanha, S. Hashemi, S. Homayoun & K. K. R. Choo, An opcode-based technique for polymorphic Internet of Things malware detection. Concurrency and Computation: Practice and Experience. 2020, Vol. 32 (6), e5173.
[26] H. Darabian, A. Dehghantanha, S. Hashemi, M. Taheri, A. Azmoodeh, S. Homayoun & R. M. Parizi, A multiview learning method for malware threat hunting: windows, IoT and android as case studies. World Wide Web. 2020, Vol. 23 (2), pp. 1241-1260.
[27] F. Manavi, A. Hamzeh, A new approach for malware detection based on evolutionary algorithm. In Proceedings of the Genetic and Evolutionary Computation Conference Companion (2019) 1619-1624.
[28] C. Shu, D. Dosyn, V. Lytvyn, V. Vysotska, A. Sachenko & S. Jun, Building of the predicate recognition system for the NLP ontology learning module. In 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) 2019, Vol. 2, pp. 802-808.
[29] R. Kochan, K. Lee, V. Kochan & A. Sachenko, Development of a dynamically reprogrammable NCAP [network capable application processor]. In Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No. 04CH37510). 2004, Vol. 2, pp. 1188-1193.
[30] S. Lysenko, K. Bobrovnikova, P. T. Popov, V. Kharchenko, D. Medzatyi, Spyware detection technique based on reinforcement learning. In CEUR Workshop Proceedings. 2020, Vol. 2623, 307-316.
[31] S. Lysenko, K. Bobrovnikova, R. Shchuka, O. Savenko, A Cyberattacks Detection Technique Based on Evolutionary Algorithms. In 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT) IEEE (2020) 127-132.
[32] A. Drozd, M. Kuznietsov, S. Antoshchuk, A. Martynyuk, M. Drozd & J. Sulima, Evolution of a Problem of the Hidden Faults in the Digital Components of Safety-Related Systens. In 2018 IEEE East-West Design & Test Symposium (EWDTS) IEEE (2018) 1-5.
[33] M. Zuzcak, T. Sochor, Behavioral analysis of bot activity in infected systems using honeypots. In International Conference on Computer Networks. Springer, Cham (2017) 118-133.
[34] T. Sochor, M. Zuzcak, High-interaction linux honeypot architecture in recent perspective. In International Conference on Computer Networks. Springer, Cham (2016) 118-131.
[35] O. Barmak, Y. Krak, E. Manziuk, Diversity as The Basis for Effective Clustering-Based Classification. ICST 2020 (2020) 53-67.
[36] A. Melnyk, V. Melnyk, Remote Synthesis of Computer Devices for FPGA-Based IoT Nodes. 2020 10th International Conference on Advanced Computer Information Technologies, ACIT 2020 – Proceedings 9208882 (2020) 254-259.
[37] A. Melnyk, V. Melnyk, Specialized Processors Automatic Design Tools-the Basis of Self-Configurable Computer and Cyber-Physical Systems. 2019 IEEE International Conference on

Advanced Trends in Information Theory, ATIT 2019 – Proceedings (2019) 326-335. doi:10.1109/ATIT49449.2019.9030481
[38] S. Lysenko, K. Bobrovnikova & O. Savenko, A botnet detection approach based on the clonal selection algorithm. In 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT). IEEE (2018) 424-428.
[39] R. Leizerovych, G. Kondratenko, I. Sidenko and Y. Kondratenko, "IoT-complex for Monitoring and Analysis of Motor Highway Condition Using Articial Neural Networks," 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), Kyiv, Ukraine (2020) 207-212. doi:10.1109/DESSERT50317.2020.9125004.
[40] O. Pomorova, O. Savenko, S. Lysenko & A. Kryshchuk, Multi-agent based approach for botnet detection in a corporate area network using fuzzy logic. In International Conference on Computer Networks. Springer, Berlin, Heidelberg (2013) 146-156.
[41] A. Drozd, M. Al-Dhabi, S. Antoshchuk, A. Martinyuk & M. Drozd, Models and methods checking mantissas by inequalities for on-line testing of digital circuits in critical applications. In 2017 IEEE East-West Design & Test Symposium (EWDTS). IEEE (2017) 1-5.
[42] S. Lysenko, K. Bobrovnikova, S. Matiukh, I. Hurman & O. Savenko, Detection of the botnets' low-rate DDoS attacks based on self-similarity. International Journal of Electrical & Computer Engineering, 2020, 10, 2088-8708.
[43] K. Bobrovnikova, S. Lysenko & P. Gaj, Technique for IoT Cyberattacks Detection Based on DNS Traffic Analysis. CEUR, 2623 (2020) 19.
[44] S. Lysenko, O. Savenko, K. Bobrovnikova & A. Kryshchuk, Self-adaptive system for the corporate area network resilience in the presence of botnet cyberattacks. In International Conference on Computer Networks. Springer, Cham (2018) 385-401.
[45] I. Santos, F. Brezo, X. Ugarte-Pedrero & P. G. Bringas, Opcode sequences as representation of executables for data-mining-based unknown malware detection. Information Sciences. 2013, Vol. 231, pp. 64-82.
[46] Packages Search for Linux and Unix. URL: https://pkgs.org/
[47] VirusTotal. URL: http://www.virustotal.com
[48] Obfuscatoin-for-ARM-disassembled-binary. URL: https://github.com/darabian/Obfuscatoin-for-ARM-disassembled-binary
[49] Hex Rays. IDA Pro. URL: https://www.hex-rays.com/products/ida/
[50] R. Millham, I. E. Agbehadji, H. Yang, Pattern Mining Algorithms. In Bio-inspired Algorithms for Data Streaming and Visualization, Big Data Management, and Fog Computing. Springer, Singapore (2021) 67-80.
[51] S. Lysenko, K. Bobrovnikova, O. Savenko & A. Kryshchuk, BotGRABBER: SVM-based self-adaptive system for the network resilience against the botnets' cyberattacks. In International Conference on Computer Networks. Springer, Cham (2019) 127-143.
[52] S. Lysenko, K. Bobrovnikova, A. Nicheporuk, R. Shchuka, SVM-based technique for mobile malware detection. In CEUR Workshop Proceedings (2019) 85-97.
[53] S. Lysenko, O. Savenko & K. Bobrovnikova, DDoS Botnet Detection Technique Based on the Use of the Semi-Supervised Fuzzy c-Means Clustering. In ICTERI Workshops (2018) 688-695.