Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

DOI: 10.37943/20VPSX8675

Tamara Zhukabayeva
PhD, Professor, Department of Information Systems
zhukabayeva_tk@enu.kz, orcid.org/0000-0001-6345-5211
L.N. Gumilyov Eurasian National University, Kazakhstan

Nurdaulet Karabayev
PhD student, Department of Information Systems
020419501012@enu.kz,orcid.org/0009-0008-6532-6382
L.N. Gumilyov Eurasian National University, Kazakhstan

Asel Nurusheva
PhD, Researcher, Department of Information Security
asselnurusheva7@gmail.com, orcid.org/0000-0001-5407-7191
L.N. Gumilyov Eurasian National University, Kazakhstan

Dina Satybaldina
Candidate of Physical and Mathematical Sciences, Associate Professor,
Research Institute of Information Security and Cryptology
satybaldina_dzh@enu.kz, orcid.org/0000-0003-0291-4685
L.N. Gumilyov Eurasian National University, Kazakhstan

A METHOD OF VULNERABILITY ANALYSIS IN WIRELESS INTERNET OF THINGS NETWORKS FOR SMART CITY INFRASTRUCTURES

Abstract: The article proposes an approach to information security vulnerability analysis and threat modeling in wireless Internet of Things networks for Smart City infrastructures. Currently, such infrastructures are becoming increasingly widespread in a variety of Smart City application areas, including industrial life support systems, pipelines, communication networks, and transportation systems. The wide coverage of end users, the critical nature of such infrastructures and the value of their inherent assets determine the increasing importance of solving problems of determining the security level of such infrastructures and the timely application of protective measures. The ultimate goal of the proposed approach is to assess the security of the infrastructure. This article analyses articles at the intersection of the subject area of vulnerability and attack analysis in information systems and networks and the area of Smart City infrastructure issues. The proposed approach includes the use of an analytical model of an intruder which, together with the analysis of the specification of a specific Smart City infrastructure, allows us to determine the current types of attacks. In order to obtain infrastructure security assessments, the CAPEC database of wireless network vulnerabilities and attack patterns is analysed. In this case, the main attributes of the attacks are identified, unified and transformed into a single format using the numerical values of the considered attributes. The feasibility of the proposed approach is also analysed and its main advantages and disadvantages are considered. In addition, the main areas of further activity and tasks related to testing and improving the proposed approach in practice are identified.

Keywords: internet of things; wireless networks; smart city infrastructure; attack; vulnerability

Received: 24.09.2024 Accepted: 17.10.2024 Published: 30.10.2024

DOI: 10.37943/20VPSX8675

Introduction

At present, various integrated systems of the Internet of Things are becoming more widespread, encompassing various embedded, mobile and autonomous electronic devices using wired and wireless communication protocols. Such systems are increasingly aimed at solving problems in areas of human activity such as intelligent energy networks, transport and logistics systems, telemedicine systems and implantable medical devices for monitoring human health, production networks for managing production processes using numerical control, etc. [1].

Modern smart city infrastructures are organised at the intersection of automation and digitalisation of the processes of using physical space and energy resources, information transfer, user and asset management [2]. At the same time, smart city infrastructures usually imply a strong collaborative aspect in the interaction of stakeholders, covering the joint implementation of distributed accumulation and exchange of primary data from sensors and device users, data processing processes, their intelligent analysis and making operational decisions on the maintenance and operation of the entire infrastructure. In addition, unlike information and communication systems and general-purpose networks, in smart city infrastructures the processes of physical movement of subjects, physical objects, electronic devices, energy exchange processes, processes of monitoring and accounting for the consumption of physical resources used within such an infrastructure acquire special importance.

In [3], the main characteristics of smart cities are the active use of modern information and network technologies, using the evolution of telecommunications and IP networks, including an automated asset management system. In particular, the most important aspects of the functioning of such smart cities are increased control over energy consumption and operating resources in order to reduce overall operating costs.

As examples, a range of works highlight areas of smart city development and specific application scenarios, such as an intelligent energy system of a smart city aimed at solving the problems of interaction between individual energy elements within the wireless network of the Internet of Things [4]; a scenario of a heterogeneous network of devices for implementing real-time video surveillance in a smart city environment [1]; a system for predicting the availability of parking spaces [5], etc. All these systems and scenarios, which assume the presence of highly specialised devices, their functions and the services they provide, and which operate in a potentially unreliable and untrusted environment, contain software and hardware vulnerabilities, including zero-day vulnerabilities [6]. Exploitation of such vulnerabilities by information security violators can lead to successful implementation of attack actions aimed at compromising the security properties of the entire smart city infrastructure. Negative consequences of such attacks include, for example, the disruption of energy generation and distribution processes, with power outages for key consumers and damage to the electrical equipment of the smart city. Another example is the disruption and actual blocking of parking processes and the creation of traffic jams of public, commercial and private transport in the city.

Therefore, all this determines the importance of studying the issues of vulnerability analysis in wireless networks of the Internet of Things within the smart city infrastructure, as well as modeling information security threats in such infrastructures for the purpose of subsequent assessment of the infrastructure security level and further forecasting of cyber-physical security incidents.

The main contribution of this article includes the analysis of relevant literature sources in the area of the study, the proposed approach to vulnerability analysis and threat modeling for obtaining assessments of the infrastructure security level, as well as the results of the analysis of the applicability of this approach. One of the peculiarities of the work is the focus of the proposed approach on considering the risks of carrying out attack actions, both with the initial knowledge of the characteristics of the target system, expected types of intruders and possible

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

sequences of their steps, and with the subsequent clarification of the posterior probability of specific actions of the attacker.

The rest of the article is organised as follows. The next section includes an analysis of relevant literature sources in the field of vulnerability analysis and threat modeling in wireless Internet of Things networks for smart city infrastructures. The next section includes an analysis of the vulnerability database of wireless networks for the Internet of Things. The next section reveals the essence and main features of the proposed approach to vulnerability analysis and threat modeling. The final section provides a summary of the results obtained and the main conclusions.

Literature review

Recent years have seen the proliferation of various wireless networks. At the same time, despite the mobility characteristics of the devices, their autonomy and the personal nature inherent to some of these devices, these characteristics make wireless networks vulnerable to attacks related to the unauthorised exploitation of the mobility and autonomy characteristics, aimed at violating the security policy of a smart city, violating the confidentiality, integrity and availability characteristics of data [4], and damaging the information services provided by the network devices [7].

In particular, unlike stationary devices such as personal computers, servers and other network devices, mobile devices, operating in a potentially untrusted, changing environment, are subject to a number of unauthorised manipulations of the data of such devices, such as modification of device configuration parameters, eavesdropping on communication channels, an attack of repeated reproduction of previous authentication commands, and so on. In addition, the autonomy of mobile network devices determines the importance of the battery life of the device, the malicious accelerated depletion of which can disrupt the operability of the device and the functionality of its services [8].

In addition, using a mobile network device that communicates with other devices on the mobile network via a wireless communication channel and that is at the disposal of the intruder, the intruder can move along the network by successive elevation of privileges to take control or control of other, not necessarily mobile, devices on the smart city network. There are a number of approaches to assessing the information security risks of the operation of mobile devices. In [9], risk is understood in terms of the likelihood of a particular type of attack and the criticality of the information and software/hardware assets that such an attack may affect. An approach to assessing the risks of running a specific software application is also studied [1], depending on the configuration parameters and permissions required by that application [10].

In [2], the concept of Digital-as-a-Service (DaaS) is disclosed, which abstracts the digital infrastructure independently of the characteristics of the real physical infrastructure and cloud structure underlying the actual execution environment of the information services provided. In particular, elements are highlighted that cover the management of digital systems, data transmission, servers and workstations in the context of smart city scenarios, such as smart public transport, building management, digital railways, emergency response, etc. In [11], the systematisation and analysis of information security risks in smart city infrastructures is carried out. At the same time, it is noted that the variability of such infrastructures complicates the process of assessment and confirmation of the quality of such assessment in real practical cases. The authors propose an approach to risk assessment based on the use of artificial intelligence methods to automate the assessment processes, limited to DDoS attacks and attacks that disrupt the processes of identifying network nodes and disrupting routing, such as Black Hole attacks, Gray Hole attacks, Sinkhole attacks, Wormhole attacks, Sybil attacks and Illusion attacks.

DOI: 10.37943/20VPSX8675

Unlike existing published work in this area, for instance [12] and [13], this article focused on different types of attacks that are currently recognised as relevant and included in the extensible CAPEC attack template database. This provides the possibility to adapt the resulting security assessment apparatus to the conditions of different types of smart city infrastructures in different applications. In addition, a limitation of the work [13] is that the authors rely primarily on network infrastructure metrics to obtain assessments.

Analysis of the wireless network vulnerability database

A distinctive feature of the proposed approach to vulnerability analysis and threat modeling in IoT wireless networks for smart city infrastructures is that it is based, on the one hand, on the specification of the wireless network under study and its devices and, on the other hand, on information from existing information security standards and open vulnerability databases. The main database used in this work is CAPEC [14] (Common Attack Pattern Enumerations and Classifications), which is an open catalogue of attack patterns. This database contains descriptions of related entities – types and categories of attacks, their representations and their various characteristics and metrics that allow to evaluate these attacks, the possibilities and probabilities of their successful implementation. In addition, the possibilities of attack modeling based on information from CAPEC are assumed. As an example, Table 1 provides partial information on the CAPEC-70: 'Try Common or Default Usernames and Passwords' and CAPEC-117: 'Interception' attack patterns, which characterise current attacks on mobile Internet of Things (IoT) networks in smart city infrastructures [14].

Table 1. CAPEC vulnerability database fragments on current attacks on mobile networks of the Internet of Things of Smart City infrastructures

Attack Pattern ID	Attack Pattern Name	Likelihood Of Attack	Typical Severity	Prerequisites	Skills Required / Resources Required	Consequences – Scope	Consequences – Impact
70	CAPEC-70: Try Common or Default Usernames and Passwords	Medium	High	The system uses one factor password based authentication. The adversary has the means to interact with the system.	[Level: Low] Common default usernames/passwords or brute force attack for the known user name.	Confidentiality Access Control Authorization	Gain Privileges
117	CAPEC-117: Interception	Low	Medium	The target must transmit data over a medium that is accessible to the adversary.	Skills to intercept information passing between the nodes of a network (e.g. tcpdump tool for TCP/IP).	Confidentiality	Read Data

In this article, the CAPEC database is used, version 3.9, which can be used directly to analyse and model the information security processes of smart city infrastructures. This database is extensible and currently contains 559 records of different attack patterns. The *Attack Pattern ID* attribute is a unique identifier of the attack pattern under which it is stored in the CAPEC database. The *Attack Pattern Name* attribute is a short, intuitive name for the attack that allows to understand its scope and specifics. In addition, this attribute can be used to pre-select the relevant attack patterns of interest for a particular case, as well as to explicitly classify patterns according to certain characteristics. For example, the CAPEC-70 pattern defines an actual attack using the default password and brute-force password selection using a specified

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

dictionary. Note that such a vulnerability is extremely relevant to Smart City infrastructures, as Internet of Things devices, such as portable Wi-Fi access points, are often poorly configurable due to organisational, technical, operational and financial constraints. In addition, an incorrectly performed standard firmware update of such a router may inadvertently result in the collection of its configuration settings with the possibility of administrative access to it using the default factory credentials. CAPEC-117 provides a basic template for a passive attack to intercept data, including IoT wireless radio channels, using the principles of mesh topology used in smart cities.

The *Likelihood Of Attack* attribute determines the probability of a successful attack, assuming it is launched by an intruder. This attribute can take one of a number of categorical, qualitative values – for the two specified attack templates this attribute takes the values *Medium* and *Low* respectively. In particular, the categorisation of possible attacks according to the probability of their implementation within the Smart City infrastructure allows the prioritisation and sequencing of measures to close vulnerabilities associated with the considered current attack types [15].

The *Typical Severity* attribute specifies a qualitative characteristic of the severity of an attack – for the two attacks considered, these are *High* and *Medium* respectively. For each attack, the *Prerequisites* attribute specifies the prerequisites for its implementation. For CAPEC-70 this is the use of single-factor password authentication, while for CAPEC-117 it is the presence of a wired or wireless medium accessible to the attacker through which data is transmitted. Note that in the general case, the attribute is a laconic textual description that clearly defines the requirements interpreted at the expert level. Thus, network security administrators and smart city content service providers can be guided in the selection of application technologies, relational and ontological DBMS systems, application servers, and web servers by the selection of assets used with the condition of minimising the *Typical Severity* attribute for all vulnerabilities underlying a given attack.

The next two attack template attributes are combined into a single field, *Skills Required / Resources Required*, which defines firstly the skills required by the attacker to perform this attack, expressed by a qualitative characteristic and a general textual description, and secondly the software, hardware and time resources and tools required to perform the attack. According to the specifics of each attack template, the qualitative assessment of skills and resources is optional and may not be explicitly specified. In particular, the CAPEC-70 template requires skills and basic tools for simple password guessing using available dictionaries, while CAPEC-117 requires passive eavesdropping on the Smart City network communication channel.

The *Consequences* attribute is divided into the following two attributes. The *Scope* attribute defines the basic information security mechanisms that may be violated by the attack in question. Examples include confidentiality, integrity, availability, authentication, authorisation and access control mechanisms. The *Impact* attribute specifies the type of impact effect that the attack provides and includes, in particular, values such as *Read Data, Modify Data* and *Gain Privileges*, which determine the acquisition of data by the attacker, the modification of data and the gain of privileges. Note that an attack pattern may include several effects at the same time, as in CAPEC-560: Use of Known Domain Credentials, where the three effects listed above may be present simultaneously, one of them, or a pair of them. In addition, if it is not possible to clearly identify a particular type of impact, the *Scope* attribute may take the value *Other*.

In addition to the attributes listed in the description, most templates also use additional attributes, including more detailed textual descriptions of the attack, links to other known types of attacks and tricks (*Weakness*), possible mitigations (*Mitigations*), and others. However, we also note that for some attack templates, such as the CAPEC-399: 'Cloning RFID Cards or Chips' template, some of the important attributes are not defined, which means that the pro-

DOI: 10.37943/20VPSX8675

cess of modeling and analysing the associated threats has to work under conditions of uncertainty. The information extracted from the CAPEC vulnerability database is a knowledge model that can be used firstly for expert analysis of possible types of attacks as applied to wireless networks of the Smart City infrastructure. In particular, based on the available values of the estimates of the extracted attributes, it is possible to calculate the information security risks of individual devices and to automate the risk calculation process. Such estimates can in turn be used to assess the security of both individual wireless network nodes and the entire smart city infrastructure.

As a way of implementing such procedures, it is also advisable to study data mining tools for advanced processing of information from the CAPEC database through intelligent processing of attributes and calculation of correlations between them [16]. It is advisable to solve such problems using machine learning and deep machine learning methods, in particular using recurrent neural networks and convolutional neural networks [17].

A vulnerability analysis and threat modeling approach to assess network security

The approach proposes a system model for assessing the security of wireless Internet of Things networks of the Smart City communication infrastructure. The sources of initial data for this model are the specification of the digital infrastructure of Smart, a CAPEC-based model of knowledge about attack patterns, and an analytical model of an intruder of the Smart City system. A generalised diagram of the proposed model is shown in Figure 1.

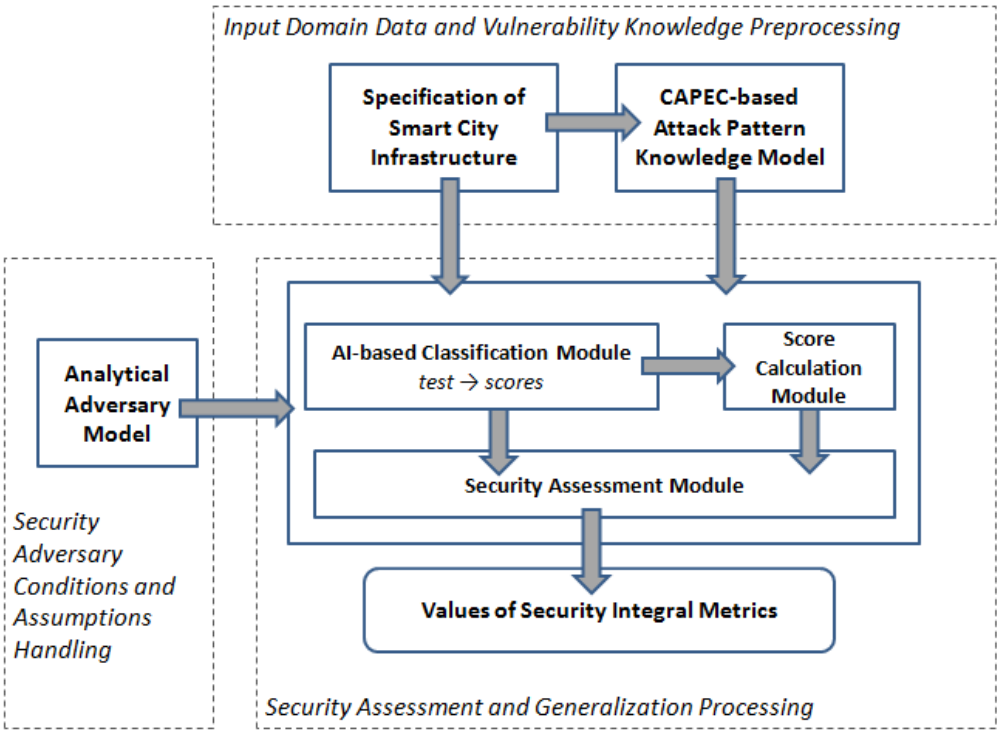

Figure 1. Model diagram for assessing the security of wireless networks of the Internet of Things of the Smart City communication infrastructure

The Smart City Digital Infrastructure Specification provides a formal description of the structure, devices, and communication channels of the infrastructure using identifiers from the Official Common Platform Enumeration (CPE) database developed by the National Institute of Standards and Technology [18]. This database includes a dictionary of software identifiers of the form *cpe:/<p>:<v>:<a>*, where *<p>* denotes the entity being described, *<v>* denotes the ven-

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

dor, and <*a*> denotes an additional extension component of the description. The <*p*> symbol can take one of the following three predefined values, namely *h* for hardware, *o* for operating system, and *a* for software application, while the <*v*> and <*a*> symbols take specific named values. The *cpe* identifier allows for the unambiguous identification of specific software and hardware samples, which is necessary for searching and combining information in various vulnerability databases, such as CAPEC. An example of a description of a Smart City infrastructure *cpe* device is the identifier '*cpe:/h:miele:xgw_3000_zigbee_gateway:-*', which specifies a ZigBee Gateway device with the XGW 3000 model, which works with the ZigBee protocol support and is manufactured by Miele. Within this identifier, any additional extended information about the software application or firmware of this device is omitted and indicated by a hyphen.

The CAPEC-based attack pattern knowledge model includes the aggregation of attack pattern data by CPE software and hardware identifiers obtained from the Smart City infrastructure specification. In fact, this knowledge model in its current form, which is sufficient to perform the Smart City infrastructure security assessment procedures, is a tabular relational representation of the corresponding attack patterns using the following map consisting of eight information fields: *{Attack Pattern ID; Attack Pattern Name; Likelihood of Attack; Typical Severity; Prerequisites; Skills/Resources Required; Consequences – Scope; Consequences – Impact}*. The main necessary explanations of the structure, format and semantics of the fields are given above.

The main types of initial data used for security assessment as well as the sources of obtaining this data, and examples of initial data of the corresponding categories are given in Table 2. Network topology data is generated by wirelessly scanning the network, while vulnerability scanners such as Nessus and OpenVAS are used to identify specific vulnerabilities. Information on the found vulnerabilities of devices is extracted automatically by using a Python script based on the analysis of data on the software and hardware of the wireless network nodes. Specific security events are obtained by analyzing the available logs of various mechanisms and utilities for analyzing and ensuring information security, such as firewalls, antiviruses and Security information and event management (SIEM).

Table 2. Characteristics of initial data for assessing network security

Initial data	Sources	Examples
Network topology data	Network scanners Zenmap, SolarWinds, Advanced IP Scanner, etc.	{Hi} – a set of hosts {Lij} – a set of connections between hosts {PL} – properties of the connections, including types of communication channels and channel bandwidth
Data on software and hardware of network devices	National Vulnerability Database CPE [18	Wireless router D-Link DIR-826L/RU/A1A Wireless N600 Dual-band Gigabit Cloud Route represented using an identifier cpe:/h:dlink:dir-826l_wireless_n600_cloud_router:a1 [19]
Vulnerability Information	Vulnerability Databases and Standards CVE MITRE [20], CAPEC [14] and CVSS [21]	Vulnerability CVE-2013-4772 allowing an adversary being located remote to to bypass authentication via [22] a direct request when an authorized session is active [23].
Security Events	Firewall, SIEM, Antivirus, etc.	- receiving a network packet on the host - event of a regular update of an operating system component - authorization of a new user in the system

The intruder model of the Smart City system is based on an approach to the analytical modeling of information security intruders, which is based on the analysis of possible attacker types by the type of access to the Internet of Things device and by the level of capabilities.

DOI: 10.37943/20VPSX8675

The 5 corresponding types of intruder access to a specific device and 3 corresponding levels of capability introduced in this approach form 15 different intruder categories. Main types and levels used to characterise a potential intruder and its characteristics will be explained. The intruder is divided into types according to the following areas of influence: Type 0 – exploits social engineering attacks; Type 1 – remote access using global Internet protocols; Type 3 – remote access using IR, Bluetooth, RFID, NFC, etc.; Type 4 – access via wired device interfaces such as RS-232, microUSB, microcontroller hardware pins, external devices, etc.; Type 5 – full direct access to the device [24]. The intruder is also divided into three levels of capability: level 1 – exploiting known vulnerabilities and attack scenarios, as well as using simple, widely accepted utilities such as Nmap; level 2 – using specialised attack tools; level 3 – essentially a group of level 2 intruders with unlimited financial, computing and hardware resources [25]. As a result, mapping intruder categories to specific assets of the smart city infrastructure allows filtering out only those categories of potential intruders that are relevant to the infrastructure under consideration. In other words, it becomes possible to select for subsequent analysis only those intruder categories whose inherent attacks have corresponding points of application to critical assets of the smart city infrastructure.

It should be noted that such a comparison can be carried out expertly, i.e. manually, on the basis of the existing expert qualification. At the same time, such a comparison can be carried out using a system of rules that specify the process of such selection. In addition, on the basis of the already selected categories of intruders, it is possible to prioritise the attacks inherent in each of the categories of intruders in accordance with the objectives, limitations of the target infrastructure devices, as well as the value of the information assets.

The main module of the security assessment system consists of the following three components (see Figure 1) The classification module, based on machine and deep learning methods, which allows transforming the following three attributes of attack patterns CAPEC *Attack Pattern Name, Prerequisites*, *Skills Required / Resources Required*, extracted from the obtained knowledge model, into categorical qualitative values suitable for automating the assessment of the security level of individual devices and the infrastructure as a whole. In the current work, the following classical machine learning methods are used for these purposes, including the Support Vector Machine (SVM), Random Forest and AdaBoost classifiers, as well as the LSTM recurrent neural network. The results of the operation of this module on specific samples of the original data are fed into the score calculation module for combination and unification with the existing categorical data. The conversion of categorical values into numerical values is performed on the basis of setting normalised scores and using the one-hot encoding method [26]. Score Calculation Module produces a transformation of categorical values of attributes to normalized numerical values, while Security Assessment Module is responsible for calculating the level of security and risks according to the following formulas. The calculation of the security analysis is fulfilled separately for each wireless network device according to the following formula:

$$sc_{AS}(dev_i) = \frac{1}{R_a(dev_i) + R_i(dev_i) + R_c(dev_i) + R_n(dev_i)} \quad (1)$$

where $R_a(dev_i)$, $R_i(dev_i)$, $R_c(dev_i)$, $R_n(dev_i)$ are values of risks of violation of the properties of availability, integrity, confidentiality and non-repudiation of device dev_i. The higher the total risks of violating the four main properties of information security, the lower the level of security of the device in question. Therefore, the security of the entire wireless network *net*, consisting of k nodes, can be calculated using the following formula

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

$$sc_{AS}(net) = \frac{1}{\sum_{i=1}^{k} R_a(dev_i) + R_i(dev_i) + R_c(dev_i) + R_n(dev_i)} \quad (2)$$

According to the common consideration, a risk is understood as the product of the probability of a successful attack on the device in question and the amount of damage that can be caused by this attack to this device

$$R_*(dev_i) = p(att) \cdot dam(dev_i, att), \quad (3)$$

where p is the probability of the attacker achieving the attack goals *att*, whereas *dam* is the amount of damage caused. In turn, the damage can be calculated using the following formula

$$dam(dev_i) = crt(dev_i) \cdot CAPEC\_Typical\_Severity(att) \cdot CAPEC\_Attack\_Consequences(att), \quad (4)$$

where *crt* determines the criticality of the device dev_i, expressed using a vector $< A = 0.0..10, I = 0.0..10, C = 0.0..10, N = 0.0..10 >$, while the second and third factors of the formula are chosen in accordance with the fields *Typical Severity* and *Attack Consequences Scope* of attack templates CAPEC (see Table 1).

In the framework of the modeling, taking as an assumption the fact of independence of the facts of attacks on the device, we will determine the general probability that the device will be attacked by at least one of the possible variations of attacks using the following formula

$$p(att) = 1 - (1 - p(att_1)) \cdot (1 - p(att_2)) \cdot \ldots \cdot (1 - p(att_m)), \quad (5)$$

the probability of a particular attack $p(att_i)$ is calculated as a prodact of CAPEC fields *Prerequisites (PS)* and *Skills Required / Resources Required (SR / RR)*, calculated in quantitative terms using a normalizer *norm* and constructed classifier *classificator*:

$$p(att_i) = norm(classificator(PS) \cdot classificator(SR / RR)) \quad (6)$$

Thus, the existing refined categorical values of attack attributes are converted into numerical values. Then, within the Security Assessment module, these data are converted into corresponding numerical values of security using the apparatus for calculating the probability of an attack on specific infrastructure devices and the associated numerical values of risk, taking into account the rules according to the CVSS (Common Vulnerability Scoring System) standard [21]. The final results of the operation of this system will be integral assessments of the security of the entire infrastructure, collected as a convolution of private assessments of the security of all devices.

Discussion and Experiments

In this section, it will be analysed the feasibility of the proposed approach to vulnerability analysis and threat modeling in wireless networks of the Internet of Things for smart city infrastructures in terms of the feasibility and prospects of this approach, as well as the correctness of the assessments of the level of infrastructure security obtained at the output. The experiments are based on an application scenario covering a smart industrial lighting subsystem and including controlled wireless devices operating via the Wi-Fi protocol and specialized server equipment that allows for the management of network devices, data flows, information services, and also provides remote control of the main functions of the system and ensures the information security of devices.

The main functions of the analyzed smart city infrastructure include the regulation of smart lighting devices in order to optimize energy consumption processes through automation and intellectualization of service and application processes, the use of specialized sensors on de-

DOI: 10.37943/20VPSX8675

vices, and the recording of significant volumes of operational infrastructure data. Table 3 exposes the main types of entities of the infrastructure under consideration and their characteristics.

Table 3. Characteristics of the entities of the smart lighting subsystem of the smart city infrastructure

Entiry type	Characteristics
Smart lights	The main components of the system are equipped with motion sensors, light sensors and video cameras. Brightness is adjusted depending on the time of day, the presence of people or transport and other parameters.
Central controller	It collects data from all lights and sensors, analyzes it and makes operational decisions about the current functioning and safety of the system.
Communication channels	A wireless network that links the system's devices. It allows lights and sensors to transmit data and logs to the central controller and receive commands.
Environmental sensors	In addition to motion and light sensors, temperature, humidity and air quality sensors can be used to adapt lighting based on weather conditions, traffic and other conditions.
User control interface	Mobile and web applications through which operators can manage the system, view statistics and configure operating parameters.
Software adapters for integration with other subsystems ot the infrastructure	Smart lighting can be integrated with traffic management systems, security and other smart city elements to use resources more efficiently, reduce time costs and financial resources.

Based on expert analysis, a limited samples of input data within the framework of the above scenario were collected to configure the component by means of machine learning. Currently, weselected 120 CAPEC attack patterns describing attacks straightforwardly related to wireless sensor networks and applicable within Smart City infrastructures. Table 4 exposes the categories of attack patterns used in the experiments, as well as the number of attack pattern samples taken from each category.

Table 4. Categories and cardinality of the data sample used

Attack Pattern Category	Number of Sumples Used
Software	30
Hardware	24
Communications	36
Supply Chain	12
Social Engineering	11
Physical Security	7

Table 5 discloses the text values of the field, using an example of five heterogeneous attack patterns *Prerequisites*, as well as its marking in the form of a categorical value, established on a 5-point scale, traditional for the standard CAPEC – *VERY LOW*, *LOW*, *MEDIUM*, *HIGH*, *VERY HIGH* [14]. Value None means the absence of any obvious preconditions for the attack to be performed and is classified as *VERY LOW*. For some templates, it is difficult to determine the qualitative value unambiguously, as for example, for the template «CAPEC-158: Sniffing Network Traffic», therefore, such a data sample was not included in the sample we analyzed. Note that the markup in accordance with the requirements, restrictions were defined, and assumptions about the corresponding attacks as applied specifically to Smart City infrastructures.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

Table 5. Example of source data markup for classification

Attack Pattern Name	Prerequisites	Prerequisites Scores
CAPEC-114: Authentication Abuse	An authentication mechanism or subsystem implementing some form of authentication such as passwords, digest authentication, security certificates, etc. which is flawed in some way.	*MEDIUM*
CAPEC-533: Malicious Manual Software Update	Advanced knowledge about the download and update installation processes. Advanced knowledge about the deployed system and its various software subcomponents and processes.	*HIGH*
CAPEC-609: Cellular Traffic Intercept	None	*VERY LOW*
CAPEC-690: Metadata Spoofing	Identification of a resource whose metadata is to be spoofed	*HIGH*
CAPEC-699: Eavesdropping on a Monitor	Victim should use an external monitor device Physical access to the target location and devices	*VERY HIGH*

It should be noted that the need to classify a number of attack template description fields is related to the fact that, in accordance with the structure and content of the current version of the CAPEC attack template database, the *Prerequisites* and *Skills Required / Resources Required* fields are often presented only as text descriptions, usually interpreted by experts. At the same time, for some of the *Skills Required / Resources Required* fields, categorical values in the CAPEC database can be given, but this information is incomplete. In addition, for some attack templates, such as for the «CAPEC-699: Eavesdropping on a Monitor», two variants of values are defined for the *Skills Required* subfield at once, so such ambiguity must be eliminated. The hypothesis of the classification is in TF-IDF metric, which operates with the properties of the presence and importance of specified words in the context of the analyzed text description of the corresponding field, and it can be used to evaluate the *Prerequisites* and *Skills Required / Resources fields* [27]. The following three supervised machine learning methods were used in the experiments, namely SVM, Logistic Regression, and Random Forest with default hyperparameter values. Training was performed on 80% of the used data samples, while the remaining data were used to test the constructed classifiers. The training was performed using Sklearn and Pandas libraries of the Python programming language. According to the results of the experiments on the used test samples, the value of the Weighted Average F1-measure for the specified 5 classes was 0.89, which confirms the possibility of obtaining reasonable categorical representations for CAPEC text fields. It should be noted that, within the current stage of the work, machine learning was performed to check the feasibility of the proposed approach. However, in further work, in order to improve the quality of classification, a more detailed and comprehensive analysis of the used feature space and the construction of feature sets that will increase the F1-measure indicator is assumed. In addition, it is also planned to conduct empirical assessments of non-functional performance indicators, including resource consumption and time complexity indicators.

Despite the fact that this approach currently uses a fixed list of available attack patterns, one of the advantages of this approach is the potential extensibility of the CAPEC attack pattern database. Thus, with each expansion of the database, the process of learning and calculating probabilities, risks and the integral security metric can be reproduced to obtain updated values. In addition, as the wireless network operates, the tracking of actual events in the system can indicate the attacker's progress through the network, allowing the probabilities of defeating subsequent nodes to be determined a posteriori.

It is noted that in order to further improve the quality of the classification module, which allows converting such text attributes of attacks as *Prerequisites* and *Skills Required / Resources Required* into categorical ones, it is advisable to use hyperparameter optimisation tools, such as GridSearch and Optuna [28], as well as tools for combining classifiers built on the basis of different methods of machine and deep learning. As a combination, it is useful to use ensemble with stacking, majority voting, as well as sampling to balance the power of the classes observed during the expert analysis of the CAPEC database [29].

The advantages of this approach also include the ability to work with incomplete, sometimes missing, similar information about specific vulnerabilities and related attack pattern characteristics (such as missing data on the *Likelihood Of Attack* and *Consequences-Scope* attributes), which is found in some CAPEC attack patterns. In general, the lack of such data contributes to a reduction in the quality of the assessment results, but nevertheless allows for a quick and early express assessment of smart city devices and the overall infrastructure until the missing information is clarified.

Limitations of this approach include its focus on using the CAPEC attack pattern database, which primarily covers wireless and wired networks of devices interacting via TCP/IP protocols. In particular, for smart city networks built with heterogeneous devices, the security assessment will only cover smart devices, routers, servers and other devices that represent separately dedicated network hosts, while end devices indirectly connected to the network through wireless protocols such as LoRa and LPWAN will only be considered in the assessment process indirectly – through their indirect impact on the functioning of the rest of the network [30]. This limitation applies in particular to corporate and private networks, where various peripheral devices of relatively simple functionality may present vulnerabilities, the exploitation of which may allow an attacker to gain access to important nodes of the wireless network [31]. It should be noted that in the future, when improving and developing the proposed approach, a combined model is needed to eliminate this drawback, taking into account in particular information from the CVE MITRE database [20], enriched with CVSS metrics and other attributes.

As a further perspective for the work, we highlight the further development of this approach towards a combined system of static and dynamic security assessment, based on an automated and intelligent analysis of, firstly, the specification of the target Smart City system, secondly, existing open standards in the field and vulnerability databases such as CAPEC, CVE MITRE and CVSS, and, thirdly, dynamic information on the functioning of the infrastructure devices for recalculating security assessments in real time, taking into account security events that have already occurred within the infrastructure up to the time in question [32].

Conclusion

The article presents an approach to vulnerability analysis and threat modeling for smart city infrastructure in the wireless Internet of Things network. The ultimate goal of the proposed approach is to assess the security of the infrastructure. The proposed approach includes the use of an analytical model of the attacker, which, using the specifications of the smart city network devices, allows you to determine the current types of attacks. The CAPEC database of vulnerabilities and attack patterns applicable to wireless networks was analysed to obtain infrastructure security assessments. In the process of security analysis, the main attributes of attacks are identified, unified and then converted into a single format using numerical values of attributes. Finally, the article analyses the feasibility of the proposed approach and identifies the main tasks in this scientific direction as further steps.

Acknowledgment

This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992852).

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

References

[1] Sánchez, L., Elicegui, I., Cuesta, J., Muñoz, L., & Lanza, J. (2013). Integration of utilities infrastructures in a future internet enabled smart city framework. *Sensors, 13*(11), 14438-14465.
[2] Serrano, W. (2018). Digital systems in smart city and infrastructure: Digital as a service. *Smart cities, 1*(1), 134-154.
[3] Al-Hader, M., & Rodzi, A. (2009). The smart city infrastructure development & monitoring. *Theoretical and Empirical Researches in Urban Management, 4*(2 (11), 87-94.
[4] Kasznar, A. P. P., Hammad, A. W., Najjar, M., Linhares Qualharini, E., Figueiredo, K., Soares, C. A. P., & Haddad, A. N. (2021). Multiple dimensions of smart cities' infrastructure: A review. *Buildings, 11*(2), 73.
[5] Nam, T., & Pardo, T. A. (2011, June). Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times* (pp. 282-291).
[6] Stellios, I., Kotzanikolaou, P., & Psarakis, M. (2019). Advanced persistent threats and zero-day exploits in industrial Internet of Things. *Security and Privacy Trends in the Industrial Internet of Things*, 47-68.
[7] Dvinsky, M.B., Drobyshev, I.A., Nepomnyaschaya, N.V., & Pavluchenko, T.V. (2017). Smart city."smart" infrastructure, networks and communications.
[8] Al-Hader, M., Rodzi, A., Sharif, A.R., & Ahmad, N. (2009, September). Smart city components architicture. In 2009 International Conference on Computational Intelligence, Modelling and Simulation (pp. 93-97). IEEE.
[9] Theoharidou, M., Mylonas, A., & Gritzalis, D. (2012). A risk assessment method for smartphones. In *Information Security and Privacy Research: 27th IFIP TC 11 Information Security and Privacy Conference, SEC 2012, Heraklion, Crete, Greece, June 4-6, 2012. Proceedings 27* (pp. 443-456). Springer Berlin Heidelberg.
[10] Jing, Y., Ahn, G.J., Zhao, Z., & Hu, H. (2014, March). Riskmon: Continuous and automated risk assessment of mobile applications. In *Proceedings of the 4th ACM Conference on Data and Application Security and Privacy* (pp. 99-110).
[11] Kalinin, M., Krundyshev, V., & Zegzhda, P. (2021). Cybersecurity risk assessment in smart city infrastructures. *Machines*, 9(4), 78.
[12] Lupton, B., Zappe, M., Thom, J., Sengupta, S., & Feil-Seifer, D. (2022, January). Analysis and prevention of security vulnerabilities in a smart city. In *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0702-0708). IEEE.
[13] Pertence, A.A., Mini, R.A., & Marques-Neto, H.T. (2020, September). Vulnerability Analysis of the Urban Transport System in the Context of Smart Cities. In *2020 IEEE International Smart Cities Conference (ISC2)* (pp. 1-8). IEEE.
[14] CAPEC. Common Attack Pattern Enumeration and Classification. A Community Resource for Identifying and Understanding Attacks. https://capec.mitre.org (accessed on 2024.10.04).
[15] Butun, I., Österberg, P., & Song, H. (2019). Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures. *IEEE Communications Surveys & Tutorials, 22(1)*, 616-644.
[16] Riera, T. S., Higuera, J. R. B., Higuera, J. B., Herraiz, J. J. M., & Montalvo, J. A. S. (2022). A new multi-label dataset for Web attacks CAPEC classification using machine learning techniques. *Computers & Security, 120*, 102788.
[17] An, J.H., Wang, Z., & Joe, I. (2023). A CNN-based automatic vulnerability detection. *EURASIP Journal on Wireless Communications and Networking, 2023*(1), 41.
[18] NIST. Official Common Platform Enumeration (CPE) Dictionary. https://nvd.nist.gov/products/cpe (accessed on 2024.10.04).
[19] National Vulnerability Database. CPE Summary https://nvd.nist.gov/products/cpe/detail/F130C305-BFA4-4EB5-97F3-AB42E1CDB188 (accessed on 2024.10.12).
[20] CVE. Common Vulnerabilities and Exposures. https://cve.mitre.org (accessed on 2024.10.04).
[21] NIST. Product Integration using NVD CVSS Calculators. https://nvd.nist.gov/vuln-metrics/Calculator-Product-Integration (accessed on 2024.10.04).

[22] National Vulnerability Database. CVE-2013-4772 Detail. https://nvd.nist.gov/vuln/detail/CVE-2013-4772 (accessed on 2024.10.12).
[23] Common Vulnerabilities and Exposures. CVE-2013-4772. https://cve.mitre.org/cgi-bin/cvename.cgi?name=2013-4772 (accessed on 2024.10.12).
[24] Rae, A., & Wildman, L. (2003). A taxonomy of attacks on secure devices. In *Australia Information Warfare and Security Conference*. (pp. 251-264).
[25] Abraham, D.G., Dolan, G.M., Double, G.P., & Stevens, J.V. (1991). Transaction security system. *IBM systems journal*, 30(2), 206-229.
[26] Hussein, A.Y., Falcarin, P., & Sadiq, A.T. (2021). Enhancement performance of random forest algorithm via one hot encoding for IoT IDS. *Periodicals of Engineering and Natural Sciences, 9*(3), 579-591.
[27] Yuan, H., Tang, Y., Sun, W., Liu, L. (2020) A detection method for android application security based on TF-IDF and machine learning. *PLOS ONE* 15(9): e0238694.
[28] An open source hyperparameter optimization framework to automate hyperparameter search. https://optuna.org (accessed on 2024.10.04).
[29] Fang, Y., Liu, Y., Huang, C., & Liu, L. (2020). FastEmbed: Predicting vulnerability exploitation possibility based on ensemble machine learning algorithm. *Plos one, 15*(2), e0228439.
[30] Nurbatsin, A., Kireyeva, A., Gamidullaeva, L., Abdykadyr, T. (2023). Spatial analysis and technological influences on smart city development in Kazakhstan. *Journal of Infrastructure, Policy and Development*, 8.
[31] Urdabayev, M., Kireyeva, A., Vasa, L., Digel, I., Nurgaliyeva, K., Nurbatsin, A. (2024). Discovering smart cities' potential in Kazakhstan: A cluster analysis. *PLOS ONE*, 19. e0296765.
[32] Zhakiyev, N., Kalenova, A., Khamzina, A. (2022). The Energy Sector of the Capital of Kazakhstan: Status Quo and Policy towards Smart City. *International Journal of Energy Economics and Policy*, 12(4), 414–423.