STM JOURNALS

Research & Reviews: A Journal of
Immunology

ISSN: 2277-6206 (Online)
ISSN: 2349-1280 (Print)
Volume 14, Issue 1, 2024
January–April
DOI (Journal): 10.37591/RRJoI

http://stmjournals.com/index.php?journal=RRJoI&page=index

Review RRJoI

Complications with Malware Identification in IoT and an Overview of Artificial Immune Approaches

Kazi Kutubuddin Sayyad Liyakat

Abstract

An immunity is facilitated by lymphocyte T&B-cells that possess a wide range of T&B-cell receptors, respectively. These cells can identify and react to pathogens and diseased cells by presenting peptide antigens by means of significant histocompatibility complexes (MHCs). The amount of data on the repertoire of adaptive immune receptors has increased dramatically in recent years because to advancements in deep sequencing. Furthermore, the presentation of peptides with MHC has been extensively studied by proteomics approaches. These massive data sets are now enabling the training of deep learning-DL and machine learning-ML models that may be applied to the identification of intricate and multidimensional structures in immune repertoires. This article presents adaptive immune repertoires, as they relate to biological sequence data. The passage delineates a comprehensive overview of the multifaceted applications within this domain, encompassing diverse areas such as the engineering of immunotherapeutic interventions aimed at bolstering immune responses, prognostication of a host's immunological status for tailored medical interventions, and the fine-grained prediction of antigen specificity exhibited by individual receptors, thus underpinning advancements in personalized medicine and immunotherapy strategies.

Keywords: IoT, Malware, Artificial immune, B-cell receptor, T-cell receptor

INTRODUCTION

Globe is more interconnected now than it has ever been. Technology is a need for societies nowadays, as it has ingrained itself into peoples' daily existence. Internet of Things (IoT) paradigm has made data-driven technologies possible, such as smart homes, smart cities, and e-government. The COVID-19 pandemic of 2019 has created a circumstance that has expedited the deployment of these advances in a number of ways. For example, e-health apps have emerged to supplement the exhausted medical personnel and infrastructure. Widespread access to the internet raises the possibility of cyberattacks, which could disclose previously thought to be safe data. For example, there are significant security concerns due to the large volume of patient data transmitted on Internet of Medical Things (IoMT) platforms. As a result, numerous standards have been created to deal with these problems, to stop private data from leaking. Cybercrime is characterised as any unlawful activity carried out via the internet against PCs or customary crimes directed at specific individuals. Digital apps that save personal data, like Zoom or different immunisation website, are seriously threatened by cybercrime. This issue is made much more significant by the growing role that IoT devices are playing in digital applications. In this study, we explore relevant detection and prevention techniques and analyse how IoT devices raise the danger of malware assaults.

*Author for Correspondence
Kazi Kutubuddin Sayyad Liyakat
E-mail: drkkazi@gmail.com

Professor and Head, Department of Electronics and Telecommunication Engineering, Brahmdevdada Mane Institute of Technology, Solapur, Maharashtra, India

Received Date: March 02, 2024
Accepted Date: March 13, 2024
Published Date: April 24, 2024

Citation: Kazi Kutubuddin Sayyad Liyakat. Complications with Malware Identification in IoT and an Overview of Artificial Immune Approaches. Research & Reviews: A Journal of Immunology. 2024; 14(1): 53–62p.

IoT apps pose a serious risk to the security of the system since they are the frail points in IT network [1]. Worldwide risk factor was created in 2017

when the WannaCry ransomware assault impacted over 600 organisations, including health, finance, education, and governance agencies. One of the organisations that was attacked was the NHS in the UK; as a result, medical personnel in the impacted hospitals in England and Scotland were unable to access their digital systems. This incident resulted in financial expenditures, missed appointments, and deaths. One of the main security risks associated with the Internet of Things is malware attacks, and one of the continuous problems is detecting malware, particularly unknown malware files. IoT devices are often lightweight since they have little memory and processing power. This feature restricts the complexity of potential vulnerability fixes. The identification of malware attacks in the Internet of Things faces three main difficulties. First off, the complicated nature of the safety method is constrained by the limited processing capacity of the majority of IoT devices. Second, a quick-adapting detection system—which calls for sophisticated algorithms—is required due to the increase in undetected malware attacks directed towards IoT equipment. Thirdly, an extremely strong defence system is necessary due to the quick proliferation of IoT devices and the ensuing rise in security risk [2].

The malware detection systems available now, which have proven effective on customary networks, have either become too complicated to deploy on IoT networks or are not flexible and reliable enough to support secure operations. The defence mechanisms employed by the human immune system serve as the model for Artificial Immune System (AIS) techniques. They have been shown to be robust, distributed, adaptive, and low-cost computationally, making them appropriate for Internet of Things security. Consequently, the focus of this review study is to examine and evaluate the AIS techniques for identifying malicious files in the Internet of Things.

MALWARE ATTACKS AND SECURITY ISSUES IN THE INTERNET OF THINGS

Malware poses a serious danger to Internet of Things security, and one of the main obstacles is identifying unknown malware. First, attempting to deploy security solutions is made extremely difficult by the constraints of IoT devices, such as their limited computational processing and low power retention capabilities. Second, the introduction of novel network connectivity methods, such cloud services, exposes networks to a wide range of security risks, including malware attacks. Furthermore, implementing security measures becomes more difficult when new devices—like smart sensors—that were not previously a part of standard networks are connected via these novel connection techniques. These factors make conventional virus detection techniques inappropriate for Internet of Things environments [3].

We first give a quick overview of the IoT security constraints and problems in this section. A review of current techniques for assessing and identifying malware in general is then covered, along with a discussion of how these techniques relate to IoT platforms. We then look at malware related to IoT, which has become much more prevalent recently and has to be addressed right away.

Security Challenges and IoT Features

IoT is a network of networked devices, each identified by a unique number. Without requiring human intervention, the gadgets may exchange data and communicate inside a network. The IoT system is made up of individual devices, sometimes called IoT devices, that are uniquely identified and easily integrated employing sophisticated interfaces to access the information network. IoT systems are used in many different fields, including healthcare, the environment, smart cities, business, and industry. They frequently consist of lightweight, networked IoT devices. IoMT devices are IoT devices used in the healthcare industry. These are essential for remote monitoring of health and intervention and include wearable medical monitoring devices, implanted healthcare devices, and hospital-based connected medical equipment. Because of this, protecting IoMT systems and devices is essential and necessitates using strong malware detection techniques. IoT equipment like temperature, humidity sensors stay frequently battery-operated and placed in remote areas for use in agricultural and environmental applications. As a result, a computationally and energy-efficient malware detection technique is needed to prolong the battery life. Strict security measures are necessary to prevent unauthorised access to smart

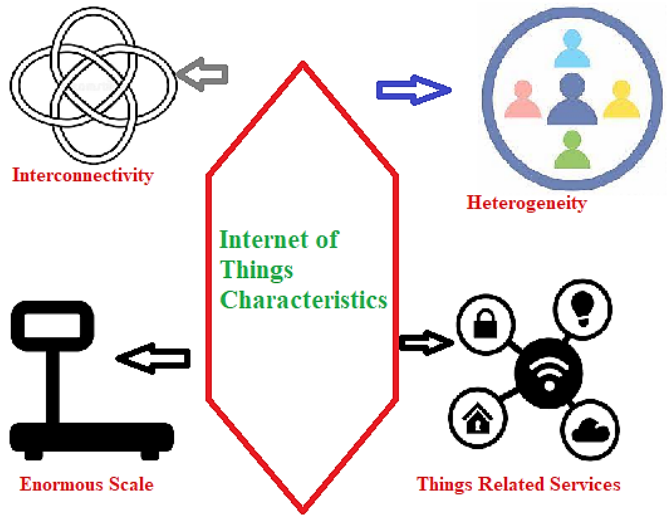

Figure 1. IoT Characteristics.

cities since they use a variety of IoT systems and devices, including surveillance cameras that record sensitive data. IoT organizations in production and Logistics where people work close to machinery run by IoT devices are referred to as industrial IoT (IIoT) applications [4, 5]. In these situations, protecting the IIoT system against malware is essential for both maintaining worker safety and IIoT operating efficiency. These IoT devices can be virtual or physical objects that communicate with one another, combining to create the IoT system with the key components shown in Figure 1.

The term “interconnectivity” describes how a device is connected to the cloud as well as other gadgets. In order to gain access to the data gathered by IoT device's sensors and to enable remote operation of the device, connectivity is a must. For instance, a patient's heart rate is monitored remotely using an IoMT equipment in heart disease prognosis. The health metrics are gathered instantly and sent to a cloud-based data centre. Thus, maintaining the security of this link is essential to safeguarding sensitive data.

Because they may have been developed on several platforms and have varying specs, IoT devices are heterogeneous. Different companies may provide the virtual components, such a cloud-based data centre, and the hardware, like a basic heart rate monitor. Different security protocols may be used by these integrated IoT devices, which could result in a lack of network standardisation. Different security protocols may be used by each connected device, each with its own set of flaws and restrictions that could expose the system to various forms of hacking.

Within the limitations of each device, virtual and physical devices can exchange services in an IoT context. Since there is no central processor or human in charge of controlling connectivity between various IoT devices, this could pose a severe hazard. A rogue device may begin to disrupt other equipment by downloading harmful data if it poses as an approved IoT device.

The quantity of data being generated by IoT devices is on the rise, with an exponential increase in their number. Between 25 billion to 50 billion IoT devices are anticipated by 2025. Massive-scale networks pose serious issues in terms of data integrity and privacy due to their sheer size. For example, enormous volumes of real-time data are created and saved in cloud by COVID-19 applications based on IoMT. However, when more data is produced, the network is under more strain, which could result in instances of incorrect interpretations [6].

Some of the smart gadgets and sensors used in the IoT run on non-chargeable batteries. Because of that, battery’s life is one of main issues with IoT security. Utilising security rules may depletee battery's

capacity. It is not advisable and dangerous to implement minimum security requirements when gadgets are able to access or are collecting sensitive data. Because these gadgets are meant to be inexpensive and lightweight, it cannot always feasible to increase the size and capacity of the battery. Among the IoT network-layer security problems are device authentication and authorization, in addition to device limiting and object identification. Because there is no worldwide root certificate authorities and a large number of linked items, it is very difficult to issue certificates to every object in the Internet of Things. Another barrier to IoT network-layer security is Domain Name System (DNS), that is employed to determine things and their properties. Here, data integrity is compromised by the potential for man-in-the-middle attacks or DNS cash poisoning. The act of posting misleading material to divert Internet traffic to malicious websites is known as an attack [7].

These security issues with IoT give rise to the potential of malware assaults. The primary line of defence for real-time malware detection is antivirus software. Nevertheless, decentralised and robust security solutions for the Internet of Things are not provided by the conventional security solutions, which have proven to be ineffective. Moving comparable solutions via conventional platforms to IoT may not be financially feasible due to IoT device limitations and processing power constraints. Since a gadget must be both secure and energy efficient, issues with battery capacity and predicted longevity limit the application of security methods. Furthermore, network resources have been incorporated into gadgets in IoT systems that were not previously thought to be a part of computer networks. The introduction of IoT devices into conventional networks presents a novel security paradigm. Aside from those aimed at IoT devices, integrated systems receive the conventional network security vulnerabilities. As a result, IoT systems need more than just typical security measures to be able to detect malware.

Analysis and Detection of Malware

Malware is characterised as harmful software that operates on a system without the consent of the user. Malware authors and creators go by many names, including hackers, crackers, and black hats. When authors create this harmful executable software, they do so for a variety of reasons, such as internal risks, governance, or competitor spying. “Traditional” malware was frequently created with defined goals and straightforward writing methods. In contrast, malware that is referred to as “next-generation” is created with many malevolent intentions and utilises technological advancements to create a more intricate design. Malware analysis and detection are becoming more crucial but also more difficult due to the combination of rapidly expanding IoT devices, malware attacks' inherent vulnerability, and their rising sophistication [8].

Examining Malware

Developing efficient malware detection strategies requires an understanding of malware analysis methodologies. In order to create an effective defence strategy, these techniques analyse the malware's functioning and operation. The same objective is accomplished by three primary malware analysis techniques: understanding how assault will affect the network and how malware operates [9].

- *Static analysis, often known as code analysis:* This method examines and evaluates the compromised file without running it. System calls, data flows, and control flow graphs (CFGs) are examples of low-level data that is extracted. Static analysis has a low rate of false-positives, which translates to a higher detection rate, is safe to employ, and analyses data quickly. Static analysis also has a global perspective because it tracks every path that could be taken, but it is unable to identify new viruses through code obfuscation.
- *Dynamic analysis, also known as behavioural analysis:* In this analysis, the malware file remains unchanged since it inspects the infected file while it is being executed, which is typically done on an unseen virtual machine. Only a small number of paths can be found using dynamic analysis based on triggered files, and it is both time-consuming and susceptible. Moreover, it has a high rate of false positives and is neither quick nor safe. Nonetheless, dynamic analysis is renowned for its effectiveness in identifying novel and unidentified viruses.

The goal of hybrid analysis was to address the drawbacks and restrictions of the earlier two methods. To enhance malware analysis, it first examines the signature descriptions of any malicious code and then integrates that information with additional dynamic parameters. Currently, cloud services enable connections in IoT networks. Most often, cloud-based malware analyses—static, dynamic, and hybrid—are used to safeguard Internet of Things devices.

Methods for Detecting Malware

Malware detection involves the application of three primary techniques: specification-based, behavior-based, and signature-based. Files are reviewed and contrasted with a previously preexisting list in the signature-based technique; if the files are included in the list, they are categorised as malicious. Since some malware contains encryption and requires a lot of computing power to extract, this strategy is not always successful in identifying all malware who enters the network. Moreover, it is ineffective against malware that is unknown or fresh [10].

Behavioral-based approach does not read the signature of the programme; instead, it observes its behaviour. The three steps of this technique are as follows: the first gathers program-related data; the second converts the input into intermediate representations for interpretation; and the third compares the intermediate representation with known behaviour signatures. This strategy can be applied in two ways: the first involves modelling the behaviour of existing programmes and comparing any new programme to it. This method detects the majority of malware, including novel varieties. Because each programme on the network behaves differently—a video reader, for instance, will utilise various services rather than a mail or web client—it is costly to develop. The second method prevents the identification of new (unknown) malware by mimicking the behaviour of known malware and comparing them with novel programmes.

The first two strategies had limitations and drawbacks, which led to the introduction of the specification-based method. This method detects malware using a variety of features, such as the following:

- *API calls:* One of the earliest groups to suggest employing system call sequences and application interfaces for malware identification.
- *OpCode:* Researchers utilise this operational code to identify malware since executable files are composed of a sequence of assembly codes.
- *N-Grams:* this technique detects malware by analysing the binary codes of executable programmes.
- *CFG:* Malware behaviour has been analysed using this graph, which shows how programmes influence flow.
- *Hybrid feature:* to improve results, researchers blend various malware detection algorithms in this machine learning method.
- Zhu, Quanyan, and T are three anomaly detection methods based on game theory. Using behavioural analysis, Başar introduced many approaches to malware detection, including consensus algorithms, detection and prevention of data exfiltration, as well as data filtered for dispersed detection
- *Prospect theoretic approaches:* The methods rely on evaluating the reliability of the system's aggregated data. The primary drawback of the specification-based approach is the challenge of precisely defining the entire set of acceptable behaviours that a system ought to display.

MALWARE ON IOT- INTERNET OF THINGS

The IoT has been protected via routing attacks centred on IPv6 route protocol by using SVELTE, an intrusion detection method based on signatures and anomalies. The malicious software detection methods discussed in the previous part have been utilised applied to the IoT. While creating a behavioral-based or specification-based method to secure the IoT is mathematically costly because of to the lengthy simulation process it requires, implementing a signature-based method for identifying

malware in IoT does not constitute the most effective approach since it is not intended to identify unfamiliar or recently created malicious files.

The two main AI approaches for IoT security are specification-based and behavior-based approaches, both of which are difficult to integrate into IoT systems. In addition to its low likelihood and instability, AI systems require a lot of resources and are computationally demanding. As a result, we examine AIS solutions for IoT security in this work that have high detection probability and are easier to implement.

IoT devices are being used more often by cybercriminals to distribute malware payloads as long as consumers and companies keep connecting gadgets to the Internet without taking the necessary security precautions. IoT attacks increased by 55% in the first half of 2019 according to SonicWall. Compared to the 12 million attacks that were discovered in the first quarter of 2018 and came from 69,000 IP addresses, this amount of attacks was nearly nine times higher. Malware attacks pose a serious danger to IoT security for all the reasons mentioned above, necessitating the need of an IoT-specific security solution.

Based on its features and architecture, the Internet of Things can be secured most effectively by putting in place a dynamic, distributed, self-monitoring & adaptive system. This prompts us to look into AIS solutions and how to use them to protect the Internet of Things from malware assaults.

IMMUNE SYSTEMS

We present the AIS techniques in this area; they are based on the immunological system of humans. After introducing the idea of AIS, we provide a quick overview of the immune system and its defence mechanisms, which serve as the basis for AIS techniques. Next, we showcase the primary AIS techniques that replicate analogous ideas.

Overview of Artificial Immune Systems

Nature finds creative solutions to issues. Over the years, computer scientists have drawn inspiration from the knowledge gleaned from studying nature to solve difficult problems. These problems are typically ones for which traditional approaches either cannot produce a workable solution or would require a complex solution requiring a lot of processing power. In situations when analytical formulations are unavailable, suboptimal solutions could be efficiently found by nature-inspired computing. Algorithms inspired by nature abstract from natural occurrences and move through computational layers or evolutionary stages before arriving at a solution. Particle swarm optimisation, artificial neural networks (ANNs), AIS, and ant colony optimisation are a few examples. The field of AIS is made up of various techniques that are influenced by numerous biological immune system ideas. Immune system is in charge of defending the body against impurities and potential threats, or “antigens.”

Overview of Immune System

The innate immune system, which consists of both internal defence layers like stomach acid and external layers like skin to defend the body, is the body's initial line of defence. Furthermore, macrophages may kill up to 100 germs before they perish, while blood cells like neutrophils can kill any harmful agent (antigen) they come into contact with before dying. Macrophages has the ability to eliminate diseased cells, including malignant ones. The adaptive immune system, it is composed of two kinds of lymphocyte cells, is triggered if innate system is unable to eradicate antigen(Threat). First, before a disease manifests, B-cells are activated in response to an antigen entering the body. They give the antigen-sticking antibodies that “mark” the antigen so the macrophages can destroy it. Memory B-cells also retain details regarding the assault for later use. T-cells, which are produced after infection, come in second. Helper T-cells and cytotoxic T-cells are the two types of T-cells. There are two categories of helper T-cells: effector T-cells, which send out an alert and offer info about antigen, and memory T-cells, which store the antigen information for later use. Cytotoxic T-cells are tasked with eliminating diseased bodily cells that are incurable.

In the lines that follow, we will go over how B-cells and T-cells collaborate to combat antigens. Comprehending these occurrences will aid in the development of an Anti-Virus System (AIS). As the primary producers of antibodies, B-cells are the backbone of our adaptive immune system. More than 100 million distinct forms of B cells in the human body, and cause of this is because each type of B cell produces a unique set of antibodies to fend off any potential attacks since different antibodies react differently to different antigens. As a result, the body begins producing more of a certain type of B-cell when an antigen that calls for that type of B-cell to handle it enters body.

In terms of how antibodies are produced, they are composed of both thick and thin chains made of various forms of deoxyribonucleic acid (DNA). The body mixes and matches various DNA strains to produce various types of antibodies that can identify any kind of antigen. This means that each B cell will end up with a unique type of antibody following the mix and match.

There are four steps in the Clonal Selection process. Initially, B-cells produce a test patch of their antibodies, which are known as B-cell receptors and travel to the surface as "bait." B-cells search for a corresponding antigen (which their particular antibodies can recognise) by floating around in their zone. Second, a B-cell that binds to a cognate antigen doubles in size and divides into two B-cells. These two B-cells then divide again, creating a total of four B-cells. Each B cell must grow and divide for up to 12 hours during this process, known as proliferation. After a week, the body will have produced enough of that particular subset of B cells to mount a significant defence against the same type of antigen. Following all of the hard work, the majority of B-cells die and the produced antibodies are sent into the bloodstream by the B-cells. Antibodies' primary function is to only identify the antigen—not to destroy it. When the antigen is finally identified by antibodies, it is the job of phagocytes—like macrophages—to consume and eliminate it. Between antigens and macrophages, the antibody creates a bridge.

Techniques for Artificial Immune Systems (AIS)

Researchers have begun to explore several techniques that mimic the mechanisms used by AIS to defend human body in order to secure computer networks. A primary purpose of AIS in security use is detection of security incidents, such as hostile actors' employment of malware, automated tools, or low-level scripts to compromise a host or network. We distinguish four techniques for creating artificial immune systems: clonal, artificial immune networks, positive selection, and negative selection.

Aiming to mimic the "process of self-tolerance of B-cells, and CLONALG, which is inspired by clonal selection theory and consists of mutation and selection processes," the negative selection approach makes use of the supervised learning classification algorithm. The detector generation phase and the matching and detection phase are the two stages of the method's operation. The process first creates detectors that are incompatible with the protected data, and it then repeatedly compares these detectors with the data. If a match is found, it indicates that the protected data has changed and that something needs to be done about it. The primary goal of this approach, which was first presented by Christley et al(2018), was to create a method with approaches akin to those of the human immune system, which is able to differentiate between self-cells (body cells) and non-self-cells (antigens). Self-cells in computer networks are mapped to approved system files, while non-self-cells are mapped to harmful files.

Data representation and matching rules are followed during the detector creation, matching, and detection stages. A significant distinction amongst several variants of negative selection algorithms is the data representation. It modifies the production of detectors, the matching rule procedure, and the detection process. Assuming that every dataset is finally implemented as binary bits, the primary data representation technique for this method is binary. Textual, boolean, category, and numeric data are examples of additional representations. Real-valued vector representation and string representation are the two categories into which these various representations can be divided. The distance measured between the generated detectors and the tested data is known as matching or recognition, and it is

defined by the matching rule. It is applied in both the detection and detector generation stages. Formally, matching rule can be described as distance measureed between k and f within a threshold for all data representations, where f is a data instance and k is a detector. With the help of said matching rule, idea of an partial matching is introduced, in which a match does not require perfect matching between the detector and the data instance in every bit. For instance, if we apply a matching distance of three to the data 11001100, matched detectors with at least five bits matching the detector's original data may be (11001100, 11001111, 11001000, 00101100, etc.).

This method operates by keeping an eye on a large network. Since each copy of detecting method is distinct, even if copy is discovered at one site, other sites will maintain their own copies. Because detection is probabilistic, distinct sets of detectors are used to safeguard distinct entities. Furthermore, the approach should identify any foreign activity instead of looking for a certain pattern.

T-cell selection process, in which only T-cells capable of identifying self-molecules (body cells) being employed in immune system, serves as the model for the positive selection approach (which is influenced by negative selection). This positive selection approach, in contrast to the negative selection strategy, will provide detectors that identify and match self-protected data. If a detector is found during the detection stage and it does not match protected data, it indicates that the protected data has undergone modifications. A generic classification technique called Positive Selection Classification algorithm(PSCA) uses classifiers that can identify self-class data to classify unknown data. PCSA was used by the authors by Leem et al. (2018) to detect malware using the following stages: learning, when the algorithm learns to categorise data into two classes viz. self and non-self, stimulation and modify stages. Lastly, unlike the typical classification strategy that uses the minimum distance between many centres, the radius is a threshold employed for categorization.

According to the clonal selection theory, which was put forth by Lima(2020), B-cells go through cloning, variation, and selection in order to develop affinity. The clonal selection hypothesis served as the inspiration for Castro and Zuben's CLONALG approach, which was first developed for optimisation and pattern recognition problems. The authors (Kovaltsuk et al.(2018)) state that defining the five primary components of CLOALG is necessary. The number of receptors, size of the receptor population, the assortment technique, affinity function which yields real-valued measures, and the function that determines the number of clones and the rate of mutation based on affinity. Cloning is made easier with the use of supervised data mining. The immune system clones the most excited lymphocytes in response to a novel antigen, but B-cells begin cloning particular antibodies for that type of antigen as it enters the body. In the same way, the CLONALG technique produces a set of receptors R that are capable of identifying a set of patterns P.

The theory of the artificial immune network (AIN) was put forth by Chicz et al(1992). Because B-cells have a self-reinforcing network, their immunological memory served as the model for AIN, an unsupervised learning system. According to this procedure, B-cells communicate with one another to maintain memory and to show active behaviour even in the absence of an immune response. Immune network theory and certain aspects of clonal selection are also imitated by AIN. Establishing a collection of repertoires for a particular problem is the aim of the AIN system process, wherein high-performing cells suppress low-similarity (comparable) cells inside the system. A natural process is used to achieve this standard, in which the population is exposed to external input and responds with both internal meta-elements of intra-population interactions and a clonal selection reaction. As a result, it evens out how the populace responds to external stimulus.

The AIS is capable of identifying and responding to harmful files that differ from the system files used during the training phase, just as the human immune system is able to recognise and react to antigens in our bodies. AIS are excellent choices for identifying unknown malware files because they have the ability to recognise anomalies in the behaviour of the system and identify attacks without

having any prior knowledge of them. We examine the most recent AIS solutions for malware detection and IoT security in the next section [6].

REVIEW OF ALGORITHMS BASED ON ARTIFICIAL IMMUNE AND IMMUNE SYSTEMS

The authors Senior et al(2020) introduced DeepDCA-AIS-based malware detection system. DeepDCA employs Self-Normalizing Neural Networks (SNN) with the danger theory technique known as dendritic cell algorithm (DCA). The suggested method concentrates on the preprocessing stage, outlining the procedures for feature selection, signal processing, SNN signal classification, and anomaly metrics. In the work, certain of categorical variables were converted using the Bot-IoT dataset, making it simple to apply feature selection method. When an approach was tested with various file characteristics, it produced an F1-Score of less than 50% when the dataset had unbalanced data. F1-Score rose to more than 90% when balanced data was used for the dataset's top ten file features. Despite having a low false negative rate and a high detection accuracy rate, this approach is not distributive enough or lightweight enough to be used in Internet of Things devices. The authors Alipanahi et al(2015) proposed Artificial-Awareness-Architecture (AWA) as a model for Artificial Immunological Ecosystems. This investigation demonstrates that while suggested method may identify intrusions in particular IoT systems, it is unable to identify anomalies or outliers [7].

Additionally, using immunology approaches, Senior et al (2020), presented a novel approach to IoT security. The suggested approach uses cyclical and dynamic defence mechanisms to counter a security threat. Security threat detection, hazard calculation, security reaction, security defence strategy design, and security defence are the five links it includes. IoT network traffic collection and analysis are handled by the first link, and subsequent links operate according to the output. The following mechanisms underpin the method's simulation of AIS techniques for intrusion detection: first, it captures data from IoT traffic and uses it to simulate data to antigens in AIS; second, it represents the detector simulation for the detection elements, like the number of recognised antigens and its living time; and third, it implements a matching mechanism to find out if a detector and an antigen match. Moreover, the differentiation of the detectors into memory, immature, and mature detectors illustrates the evolution process. The experiment simulated replay attacks, mutant replay attacks, cloning attacks, and hybrid cloning assaults. No affected files were utilized in the experiment, despite the fact that this technique may identify security threats and modify detectors to react to the changing IoT environment. Furthermore, no actual IoT scenario was used in the implementation of this study [8].

Additionally, Artificial-Immune-based technique for IoT intrusion detection was put forth by the authors Zhavoronkov et al (2019). Numerous local intrusion detection sub-models that exchange learning outcomes are used in this strategy. Using this strategy, antigens are represented as binary strings in the signature information in the IoT sensing layer. When detector sets are formed, they contain the number of antigens that the detector matches as well as the detector's generation life. The suggested method's primary drawback is that it isn't light enough to satisfy the needs of the IoT.

Lastly, an AIS-based approach for IoT intrusion detection was suggested by the authors Stokes et al (2020). It was stated that, for experimental purposes, the primary signature data on the IoT datagram is retrieved and converted to a binary character string. Immature, mature, and memory detectors are the terms used to describe the various stages of detectors. According to the authors, mature detectors develop into immature detectors, whereas immature detectors satisfy the recognition diversity of intrusion detection. Despite the fact that this work offers a novel approach to identifying unidentified malware in the Internet of Things, no simulation results were provided. Furthermore, we see that this approach requires a lot of time and memory for IoT devices [11].

CONCLUSION

IoT systems are heterogeneous, networked devices with a constrained amount of processing power. IoT applications are multiplying quickly, as is the integration of these applications into conventional

networks. As a result, there are now rapidly emerging security risks, such as malware attacks, that are not sufficiently addressed by conventional security solutions. Conventional methods for detecting IoT malware use behavioural and signature-based approaches. We have shown that these are not cost-effective for IoT applications, or inappropriate for identifying unknown malware files. The term "AIS" refers to a line of inquiry into danger detection that draws inspiration from Adaptive-Immune-System of human body. Because AIS techniques may identify unknown attacks and intelligently store record of any attempt for later use, they are generally appealing for malware detection. Furthermore, their properties match IoT system characteristics the best, making them a strong candidate in design of IoT malware detection. The characteristics of AIS methods—such as their robustness, distributed implementation, lightweight processing, and adaptivity—align with the particular needs of IoT devices. In order to do this, a survey of current AIS research for malware detection is provided in this article.

REFERENCES

1. Halli U M, "Nanotechnology in IoT Security", Journal of Nanoscience, Nanoengineering & Applications, 2022, Vol 12, issue 3, pp. 11–16
2. Liyakat, K.K.S. (2024). Machine Learning Approach Using Artificial Neural Networks to Detect Malicious Nodes in IoT Networks. In: Udgata, S.K., Sethi, S., Gao, XZ. (eds) Intelligent Systems. ICMIB 2023. Lecture Notes in Networks and Systems, vol 728. Springer, Singapore. https://doi.org/10.1007/978-981-99-3932-9_12 available at: https://link.springer.com/chapter/10.1007/978-981-99-3932-9_12
3. Christley S, Toby IT, Fonner JM, Rubelt F, Cowell LG. VDJServer: a cloud-based analysis portal and data commons for immune repertoire sequences and rearrangements. Frontiers in immunology. 2018 May 8;9:348868.
4. Leem J, de Oliveira SHP, Krawczyk K, Deane CM. 2018. STCRDab: the Structural T-Cell Receptor Database. Nucleic Acids Res. 46(D1):D406–12
5. Lima WC, Gasteiger E, Marcatili P, Duek P, Bairoch A, Cosson P. 2020. The ABCD database: a repository for chemically defined antibodies. Nucleic Acids Res. 48(D1):D261–64
6. Kovaltsuk A, Leem J, Kelm S, Snowden J, Deane CM, Krawczyk K. 2018. Observed Antibody Space: a resource for data mining next-generation sequencing of antibody repertoires. J. Immunol. 201(8):2502–9
7. Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, et al. 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. Nature 358(6389):764–68
8. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, et al. 2020. Improved protein structure prediction using potentials from deep learning. Nature 577(7792):706–10
9. Alipanahi B, Delong A, Weirauch MT, Frey BJ. 2015. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat. Biotechnol. 33(8):831–38
10. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, Terentiev VA, Polykovskiy DA, Kuznetsov MD, Asadulaev A, Volkov Y. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature biotechnology. 2019 Sep;37(9):1038-40.
11. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, MacNair CR, French S, Carfrae LA, Bloom-Ackermann Z, Tran VM. A deep learning approach to antibiotic discovery. Cell. 2020 Feb 20;180(4):688-702.

