Article

Modelling and Mitigation Strategy of IoT Botnet Propagation

Mohammed Ibrahim[1,*], Mohd Taufik Abdullah [2], Azizol Abdullah [3]and Thinagaran Perumal [4]

[1] Faculty of Computer Science and Information Technology, University Putra, 43400 Serdang, Malaysia; m.ibrahim47@yahoo.com; mibrahima47@gmail.com

[2] Faculty of Computer Science and Information Technology, University Putra, 43400 Serdang, Malaysia; taufik@upm.edu.my

[3] Faculty of Computer Science and Information Technology, University Putra, 43400 Serdang, Malaysia; azizol@upm.edu.my

[4] Faculty of Computer Science and Information Technology, University Putra, 43400 Serdang, Malaysia; Thinagaran@upm.edu.my

* Correspondence: m.ibrahim47@yahoo.com; mibrahima47@gmail.com; Tel.: +601-158631773, +234-8076331586

Abstract: Nodes in wireless sensor networks (WSN) are characterized particularly by their limited power and memory capabilities. Limited memory is an important parameter as it defines the size of the operating system and the processing code. As established previously, energy and memory efficiency is the most important evaluation factors of WSNs as they are directly related to data loss and network lifetime. However, based on our simulation results, memory efficiency determines the selection or abandon of nodes by the botmaster for the propagation of bots in an IoT infrastructure. Consequently, the node's memory efficiency determined the spread of bots in the network and provides defense actors with an insight of the botmaster behavior for mitigation of the attack. Conventional botnet propagation and mitigation models did not consider the impact of node's memory efficiency in the IoT platform. To address this gap, we build IoT-SIEF, a novel propagation model with forensic capability that will analyze command and control propagation behavior based on the perspective of the node's memory efficiency. IoT-SIEF model used to explore the dynamics of propagation using numerical simulation with more than 50% outperform other models in mitigating the number of secondary bots. Consequently, it can serve as a basis for assisting the planning, design, and defense of such networks from the investigator's point of view.

Keywords: IoT (Internet of Things); bot; botnet; propagation; nodes; sensor; infectious; mitigation

1. Introduction

The penetration and adoption of IoT application into various aspects of humanity is becoming alarming. The network of IoT is growing to be the next digital landscape [1]. By utilizing the capabilities of the wireless sensor network (WSN), IoT penetration is not only restricted to scientific research but has great potentials in civil and military applications. The sensor nodes in the IoT platform are made to sense, observe and transmit the observable data from an environment to the processing or control center [2]. The sensed data is to be utilize in improving efficiency, learning about customers and increases services [1]. However, sensor nodes are attributed to low power and radio communication capabilities [2]. The limited power capabilities of sensor nodes resulted in poor processing capacity that in turn affect the security of the WSN. Also, limitation in transmission range resulted in data generated by the distant sensor nodes from the sink (serve as a user interface) to relay the data along with the intermediate nodes [2]. Similarly, the need for data collection resulted in a dense accumulation of the sensor nodes to ensure sufficient data coverage. This will lead to the unique characteristics that have an advance effect on the malware propagation and the exploitation of the botnet threats potency.

The issues of IoT botnets threats and its emergence in the late 2016s have justified the need for not only studying its mode of propagation but how to mitigates and establish facts from its attack surface. Mirai as publicly and popularly known botnet in the late 2016s has deceived more than 600,000 devices to spread itself and cause a series of large-scale DDoS Attacks [3]. Another issue surrounding the Mirai botnet was that the source code was made free and open for public consumption and exploitation. In this regard, the question is what can be done to stalk the tide of these new types of IoT botnet attacks? [4]. Challenges related to detecting, observing and mitigating malware propagation remain critical in IoT based networks. On the other hand, sensor nodes are resource-constrained, they are mostly attributed to poor defense capabilities and are becoming the focal point for software attacks (like virus or worm attacks on the Internet), particularly when the nodes are distributed in an aggressive environment [2]. Furthermore, new technology like IoT and IPv6 remain strange in the sight of the users, understanding their operations and the mode of malware propagation is still at immature stages. As such, users are aware of the attack only when an existing infection is identified [1]. This necessitates the development of exclusive models for such wireless IoT networks which can capture the spatial distribution of the devices and the dynamic processes of malware infiltration, control command propagation, and device patching by the defender [5].

Contrary to the existing approaches our model focused on the constraint nature of nodes functionality like memory availability alongside command and control propagation behavior in IoT botnet formation. Also, to mitigate the botnet formation, recovered nodes are obtained through patching or vaccination in the existing approaches, however, recovered nodes are liable to the same malware reinfections since bots ofttimes tides with new exploits [1]. Thus, it is pragmatic to conceive the recovered nodes as an object of forensic interest that can be subjected to forensic analysis. Therefore, we model and examine the node's resource constraint with respect to the control command propagation behavior in determining botnet formation in a wireless-based IoT environment.

In line with our approach, an IoT-SIEF model was proposed in this paper, IoT-SIEF is aimed at modeling the botnet propagation in an IoT WSN. The model incorporates memory efficiency as influential factors in not only examining the mode of malware propagation but mitigating the menace for further forensic analysis. IoT-SIEF was inspired by epidemiological concept SI (Susceptible, Infectious) but the model was modified to take into account the latent infections and the effect of the forensic aspect of IoT.

Consequently, IoT-SIEF (IoT- Susceptible, Infectious, Expose and Forensic) stands to examine the control command propagation behavior with respect to the resource constraints of bots under consideration as well to determine the most influential factor from a forensic perspective. This will include examining the bot activity and mitigation of its propagation within the IoT network environment.

2. Related Works

Advances in internet technology and the corresponding cyber threats like malware propagation necessitate a push of propagation models from conventional and mobile-based to WSN and IoT based models. The wide range of literature reveals the models followed a state-based transition by modifying the concept of epidemic theory. Our proposed model tends to follow the same trends of state-based transition using epidemic modeling. WSN propagation models considered transmission range/radius,node mobility, energy usage topological variances impacting node density as an environmental factor in addition to user's awareness and recovery rates [1].

Mishra et al. [6], proposed a susceptible- exposed- infectious-recover-susceptible with vaccination (SEIRS-V) to describe the dynamics of worm propagation with respect to time in WSN. Based on a certain number of parameters that include birth, death, recovery and vaccination rates among others. Basic reproduction number, equilibria, and their stability were found. If the reproduction number is less than one, the infected fraction of the sensor nodes disappear and if the reproduction number is greater than one, the infected fraction persists. The authors performed numerical simulation using MATLAB and reveal that intensifying on vaccination and recovery can importantly mitigate the scale of infections by fascinating many S nodes into the R and S-V States. Consequently, the susceptibility of nodes to further infections is diminished.

Feng et al. [7] proposed an improved Susceptible -infectious- recover- susceptible (SIRS) by the emphasis on communication radius and distributed density of nodes as well assume a uniform distribution of nodes in a 2D space. The model investigates the dynamics of worm propagation over time in WSNs and the basic reproductive number that determines global dynamics of worm propagation in WSNs is obtained. Based on the reproductive number, a threshold is defined for transmitting radius in such way that if the value is less than the specified threshold and $R_0 \leq 1$,the worm can be eliminated. Also, the node's density threshold is defined, if for any given value of the node density is less than the specified threshold and $R_0 \leq 1$, the network will remain in "worm-free equilibrium" state. Finally, Numerical simulations results show that decreasing the value of communication radius or reducing the distributed density of nodes is an effective method to prevent worms spread in WSNs.

Khanh [8], used epidemic theory to propose a susceptible - infectious- quarantine - recovered (SIQR) model to describe dynamics of worms propagation with quarantine in the wireless sensor network. Similarly, mathematical analysis shows that the dynamics of the spread of worms determined by their threshold R_0. If $R_0 \leq 1$, the worm-free equilibrium is globally asymptotically stable, and if $R_0 > 1$, the worm-endemic equilibrium is globally asymptotically unstable. A numerical investigation is carried out to confirm the analytical results, however, based on the results of parameter analysis, some effective strategies for eliminating worms are suggested that decrease the contact and transformation parameters of the model slow down the malware propagation.

Wang et al. [9] emphasized the influence of mobile actuators in WSN and proposed a microscopic mathematical model to describe the propagation dynamics of the sensor worm. The model follows the state transition scheme of a typical susceptible-infected (S-I) infection model, but can microscopically compute the prior probability of each sensor being infected by the worm. Based on the model simulated results and comparison with the other models, various results were generated from various density values. Based on their findings, the involvement of infected mobile actuators reinforced worm propagation across various number of tests.

Gardner et al. [4] proposed an IoT-BAI (IoT- Botnet Awareness Information) model based on the concept of the epidemic model (SEIRS). The model considered the possibility of mitigating the frequency of IoT botnet attacks with improved user information that may positively affect user behavior. Nodes can be transit to R state from any state, and the recovery rates growth for a limited period following an attack. The simulation of the IoT-BAI model revealed some important aspects of the model that include changes in Infected peak periods due to changes in the various model parameters. The authors indicate that the constant flow of new nodes subject the IoT network more and more vulnerable to the botnet attack. However, increasing user awareness pushes the Botnet Reduction Phase earlier and delays the time between epidemics, theoretically decreasing botnet impact.

Ji et al. [10] overview the architecture of Mirai botnet and analyze its life cycle. In the analysis phase of malware propagation, and SIR format was used with N represents the total number of IoT devices in a region. Based on the simulation result performed on the estimated population of US IoT-enabled camera, the result showed the efficiency of Mirai botnets SYN weak password and the rate of infection transmission is less affected by the initial infection. Conclusively, the infected node cannot transmit the infections to other nodes in the network.

Acarali et al. [1] analyzed IoT botnet formation based on specific characteristics that include energy, limited processing power and the density of a node to proposed IoT-SIS model. IoT-SIS model built on the concept of epidemic model that focused on specific characteristics of IoT that distinguish from the other network. The model's starting scenario assumes that the network is currently infected and seek to measure how rapidly and widely the infections will disperse. The infection method of the model has various attack surfaces based on the network structures of inter-WSN, Intra-WSN and between neighbors. From the simulation results, the IoT-SIS model shows that the available attack space and the choosing propagation method determine the impact of spacial distribution of bot nodes. Also, pushing the malware to propagate harder or faster resulting in node's finite energy consumption while threaten the longevity and consistency of the botnet. Finally, the findings revealed that maximizing the effectiveness of transmission probability can determine the botnet propagation strategies in IoT networks than concentrating on the contact rate.

The existing works focused mainly on worm propagation in WSN. However, few works considered botnet formation in IoT base WSN with worm similarities in terms of propagation. While energy and memory efficiency is the most important evaluation factors of WSNs due to data loss and network lifetime [11]. Previous studies emphasized on energy and other influential factors to determine botnet formation and its dynamic propagation without due consideration to node's memory efficiency. Node's memory efficiency determines the free memory space of a node that can speedily process malware packets and avoid packet loss during botnet propagation. Consequently, it will affect the botmaster's command and control propagation behavior in prioritizing nodes that can widely and speedily propagate bot's infection. Similarly, existing models mitigate botnet attack via immunization, vaccination, patching or user's awareness to transfer bots to recover state without considering memory efficiency as well as command and control propagation behavior. However, recover nodes are vulnerable to reinfection since bots oftentimes get updates with new exploits.

In this paper, we propose a model of IoT botnet with an emphasis on memory efficiency as the influential factor using a system of differential equations. The model follows a state base epidemic

doi:10.20944/preprints201912.0097.v1

model but can determine the selection or abandoning of bots during botnet propagation. Contrary to previous works, we design our model with forensic capabilities as an add-on to capture bots that are of interest to the botmaster for forensic analysis. We first determine the probability of the abandon bots by the botmaster as a complement of the probability of the forensic nodes that can be transferred to forensic class. Our major contributions are listed below:

- We propose an IoT-SIEF, a novel state base epidemic model for IoT botnet that incorporates memory efficiency in describing propagation dynamics. This model can estimate the number of abandon nodes during propagation which distinguishes from previous models.
- The add-on of forensic class that can mitigate the infectious peak value and peak period of the number of secondary bots generated during botnet propagation.
- We performed a simulation using Matlab to evaluate the validity of the IoT-SIEF model. The simulation shows that the proposed model mitigates infectious peak value and delayed the propagation peak period without the addition of control strategies.

3. Background

3.1. Epidemic Modelling Approach

Epidemic modeling was developed from the concept of epidemic theory. In the medical perspective, epidemic modeling studied disease incidence in a population with respect to time. With the epidemic model, medical professionals can analyze the dynamics of disease spread and also measure potential immunization strategies [1]. However, with the dynamic nature of the individual systems, immunity may be long-lasting, temporary and permanent for some individuals. In this regard, various epidemic models such as SIR (Susceptible-Infectious-Recovery) [12]were developed to characterize the spread of infections. Apart from medical professionals, area of specializations such as social and behavioral scientists have been applying epidemic models in dealing with research questions. In the context of computing, epidemic modeling was introduced to cybersecurity by Kephart and White [13] while studying viruses in computer system, they defined populations of computer systems, replaced malware for diseases, and analyzed using the network contact communication graphs. This has advanced to utilize to different types of malware with botnets inclusively [1].

The epidemic model takes the population at a given time. The model categorized the population into states or classes along with some transitional conditions that a node can exploit to changes from one state to another. In this regard, the number of states indicates the amount of possible roles that nodes may accommodates. Over time, the model takes into consideration the rates of change of state by generating the amount of nodes within each compartment or state. Considering the state's changes, transitions from one state to another are typically formed as a system of differential equations, usually attributing elements such as the contact rates, infection, recovery, births, and deaths [1].

The epidemic theory has been built on the concept of infection rate along with susceptible and infectious individuals. The model is commonly abbreviated as SI (Susceptible-Infectious). 'Susceptible' determined a state where individuals are vulnerable to infections but not yet infected, while 'Infectious' described the state of which individuals are the carrier of the pathogen and capable of transmitting to susceptible individuals. In this regard, giving an infection rate, a node can transit from a state S to a state I by contacting the infection rate. Subsequent models followed SI by adding other classes, SIR is one of the models among many models that add recovery individuals R which can develop immunity after recovering from being infected. The basic building versions of epidemic modeling are deterministic based that may include probabilistic elements [1].

In this paper, we consider probabilistic elements into consideration a likelihood of infectious contact, abandon (when a bot is no more suitable for the control command), latent and forensic states. Therefore, our work proposed additional classes into SI to take into consideration expose (latent) nodes E and forensic nodes F at the point of botnet formation. This is because the control command prioritizes nodes with sufficient memory to propagates an attack while abandoning the nodes with low memory capabilities. However, to mitigates botnet formation, nodes of high interest to the botmaster's command and control will be identified as object of forensic interest that can be moved to the forensic class. Consequently, this can slow down the propagation rate which will likewise mitigate botnet formation.

3.2. IoT Wireless Sensor Network

IoT network is usually attributed to WSN. WSN consists of small sensing devices with constricted bandwidth, power and computational capacities [11]. Devices made of WSN collaboratively sense and respond to the environment [14]. in this regard, information captured from the source node can be sent to its neighbor nodes within its signal transmission range. The neighbor nodes also pass on this information to their neighbors [2]. Consequently, this information is communicated with the IoT enabled devices like routers, enabling access to the broader infrastructure for data processing and retrieval. IoT network infrastructure usually exploits smart sensor nodes, which are low power and generally equipped with one or more sensors, processor, memory, a power supply, a radio and an actuator [15]. The sensor gathered environmental data, the radio enables communication to transfer the data, the processor converts the data into transmission signals, power supply utilizes battery pack or secondary source of energy, depending on the deployment area. Actuators are incorporated based on the application needs and memory is mainly RAM for processing capability.

Unlike in traditional networks, WSN has its own design and resource constraints. Resources constraint includes a limited among of energy, short communication range, low bandwidth and limited processing and storage in each node. By exploiting the aforementioned constraints, an adversary can compromise a sensor node, modified the data integrity, eavesdrop messages on transit, inject fake messages and waste network resources [15].

However, incorporating security into WSN suffered from limited storage, communication, computation, and processing capabilities [15]. For the botmaster, such weaknesses can be exploits to advance its attack. Limited processing, storage and communication capabilities can result in message broadcast along with the medium that can be captured and modified by the botmaster. Low energy associated with the sensor nodes can similarly give an insight into the node's lifespan.

3.3. IoT Botnet Formation

The bitterly experienced of Mirai attack of late 2016 affected 65,000 devices in 20 hours as well its botnets attainment of peak size of 600,000 nodes [3] necessitate the need for studying botnet formation and propagation in IoT network. Mirai stylistically mimics the worm-propagation style in attacking the huge number of devices. This characterized by a period of scanning targeting vulnerable network-enabled cameras, residential access, routers, baby monitors and small profile network-enabled devices [4]. Silva [16] highlighted that IoT Botnet has three operational components: bots, command, and Control(C&C) infrastructure and a botmaster. Bots are the compromised devices spread over the internet to launch a series of attacks. The C&C provides new code, directs and instructs the bots by means of communication from the botmaster for an attack to be launch. Therefore, the botmaster is the most influential component in controlling both C&C and the bots in IoT botnet formations. On the other hand, during the rapid scanning phase, the malware forwards TCP SYN messages to random IPv4 addresses on ports 23 and 2323. If there are successful connections, the botmaster will then make an attempt to access the device using a dictionary attack based on 62 which is commonly used default credentials [1]. By getting successful access to the device, the logging and the device IP are recorded on the server which will turn trigger a loader to download the malware on the target devices [3].

According to [1], up to date, victims of IoT botnet are affected through network scanning, targeting similar ports and exploits weak logging credentials for penetration. Consequently, it is observed that IoT is not well conversant by the users and hence lacking the required security measures [1]. In this regard, Gardner [4] emphasized entity-relationship among users, botnets, ISP and authorities engaging in a relationship with one another when a botnet is deployed to the user's network.

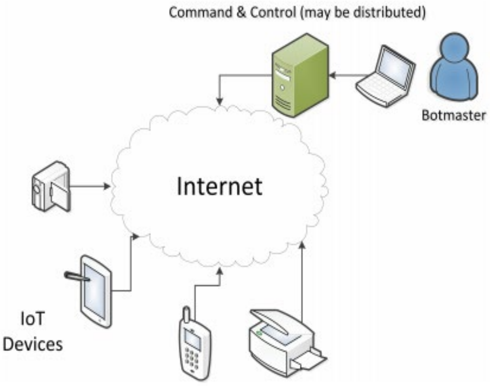


Figure 1. Typical Botnet Using Internet of Thing [4]

Once the botnet is detected by the ISP, the ISP will inform the users to increase the security of their devices and the authority for prompt action [4]. However, based on the user's perspective, IoT Botnet with Attack Information (IoT-BAI) was proposed by [4] to capture transient behavior that could exist when threat information is passed on to the users. In contrast to our proposed IoT-SEIF, IoT-SEIF aimed at finding the transient behavior based on command and control behavior and the action by the Law Enforcement Agents with respective to the influential factors that can enhance the mitigation strategy of the botnet propagation.

3.4. IoT Botnet Propagation

Botmaster propagates bots to other parts of the network by infecting other IoT devices via command and control as demonstrated in Figure 1 above. According to [17], the life cycle of a botnet starts with the infection process, where the botmaster uses different scanning methods to capture a large number of targets to be infected and converts to bots. As demonstrated in Figure 2 below, the botmaster takes advantage of users with low or lack of knowledge of network security to gain unauthorized access by exploiting weak or default passwords commonly associated with IoT devices while keeping their bots alive without being detected. When users connect the devices to the network, they hardly change their default password in a timely manner making them vulnerable to malware attacks [4].

By converting the vulnerable IoT devices to bots, in the rallying stage of the botnet life cycle, the bots connect to the C&C server to show to the botmaster that it has already become a successful zombie [18]. Thereafter, is the command and report stage in which the bot listen to the C&C servers or connect to them periodically to get new commands from the botmaster. The botmaster with total control of the bots can make several choices among the bots. In the abandoned stage of the botnet life cycle, when a bot is no more useful to the botmaster due to certain conditions, the bot may be abandon by the botmaster. In any situation, even if the bot is disabled the bot is still available. In this case, the botnet can only be destroyed if and if its entire bots are detected or abandoned or when the C&C servers are detected and closed [19]. Otherwise, the same devices that turn into bots can be used in multiple attacks if they are not detected and the malware is removed from them [4].

doi:10.20944/preprints201912.0097.v1

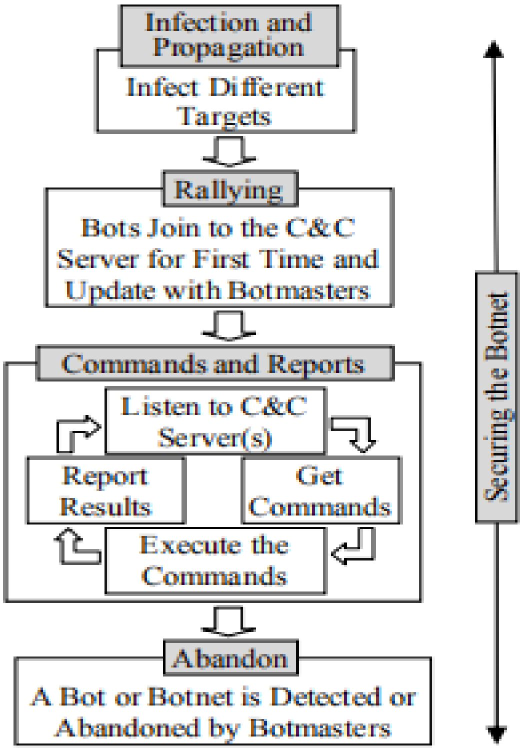

Infection and Propagation
Infect Different Targets
Rallying
Bots Join to the C&C Server for First Time and Update with Botmasters
Commands and Reports
Listen to C&C Server(s)
Report Results
Get Commands
Execute the Commands
Abandon
A Bot or Botnet is Detected or Abandoned by Botmasters
Securing the Botnet

Figure 2. Botnet Life Cycle Diagram [17]

Although botnets vary in sizes, from large-scale botnets that comprise of up to thousands of bots to a small scale botnets that comprise of hundreds of bots. Irrespective of its size, botnets are purposely created to carry out malicious activities in computer networks [20]. Considering the emergence of IoT botnet that utilizes Mirai malware, it has successfully shown that it is capable of powering some of the most powerful DDoS attacks that have been seen so far on the internet [4]. In late 2016 to early 2017, Mirai botnet was able to gain traction that resulted to attract public attention with a series of high-profile, large-scale DDoS attacks [21]. In a quest to propagate to other devices, Mirai was able to attack numerous devices with approximate 600,000 nodes at peak [3]. The events around Mirai according to Acrali et al. [1] demonstrate the spectacular threat of botnets in the IoT platform. In this regard, to fight against cybercriminals, it is important for defenders to understand malware behavior, such as propagation or membership recruitment patterns, the size of botnets, and the distribution of bots.

4. Proposed Model

We built an IoT-SIEF, a novel model that will consider the node's processing capabilities and command and control behavior for bots membership recruitment in the botnet formation. The model takes an assumption of the existence of bot in an IoT wireless environment. Then the model will measure the transient behavior of the infection with respect to the bots' condition. Our focus is particularly on memory availability that determine the abandon rate of bots based on its suitability from the botmaster command and control behavior. To understand the propagation and mitigation of IoT botnets respectively. Clear assumptions are stated as follows:

1. Dynamic network with mobility of nodes, that is, nodes can be added/ remove from the network.
2. A random network deployment.
3. Homogeneous IoT nodes with different memory status at a particular time.

IoT-SEIF has similar emerges from the epidemic model with an add-on class F to stand for the object of forensic interest. In this work, the object of forensic interest is defined as infectious nodes with high memory availability that cannot be abandon by the botmaster. Such nodes will be transferred at a probability complement of the abandon nodes to the forensic class F. Similarly, due to interaction with other IoT services, forensic nodes can lose their forensic data upon in contact with certain operations associating them with neighboring nodes. Nodes in the forensic class can be removed upon losing their forensic data and return to the susceptible class S.

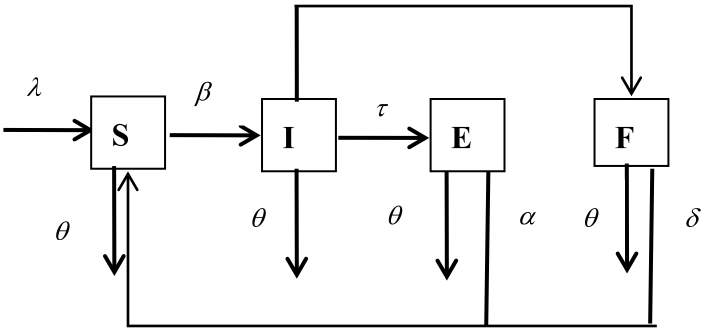


Figure 3. Flow diagram of the Proposed IoT-SEIF and Key transmission parameters

In this regard, given a population of nodes as N over a series of time interval t, a set of compartments represent possible node state will emerge. The rate of change of node from one state to another is dynamically represented using a system of differential equations. The susceptible state S is defined as a state in which nodes are not infected but vulnerable to infections at a given time t. By getting contact with the malware packets, the susceptible nodes get infected and can be converted to a bot at infectious states I. The bots are carrier nodes affected by the malware and capable of transmitting the infections to the susceptible nodes in contact. The botmaster with total control of the bots at the infectious state, can make several choices among the bots. When a bot is no more useful to the botmaster due to certain conditions at the infection state, the bot may be abandon by the botmaster. Such abandon bots can be inactive and transfer to the latent state E. Finally, the selected bots by the botmaster are considered to be the object of forensic interest that is capable of holding data of forensic interest at a given time t. Then the fractions of the susceptible $S(t)$, infectious $I(t)$, Expose nodes $E(t)$ and forensic nodes $F(t)$ make of the total number of nodes in the population at the time t, this can be expressed mathematically in equation (1) below

$$N = S(t) + I(t) + E(t) + F(t) = \frac{S(t)}{N} + \frac{I(t)}{N} + \frac{E(t)}{N} + \frac{F(t)}{N} \tag{1}$$

However, for the successful transmission of the attack, there should be a successful infectious contact β with successful transmission events that convert IoT nodes into a bot. The model is demonstrated in Figure 3 and defined mathematically using a differential equation as can be expressed equation (2) below

$$\begin{cases} \frac{dS}{dt} = \lambda - \beta S(t)I(t) + \delta F(t) + \alpha E(t) - \theta S(t) \\ \frac{dI}{dt} = \beta S(t)I(t) - (\tau + \gamma + \theta)I(t) \\ \frac{dE}{dt} = \tau I(t) - (\theta + \alpha)E(t) \\ \frac{dF}{dt} = \gamma I(t) - (\delta + \theta)F(t) \end{cases} \tag{2}$$

Where

$S(t)$ the total number of susceptible in the network

$I(t)$ the total number of bots in the network

$E(t)$ the total number of abandon nodes in the network

$F(t)$ the total number of forensic nodes

λ number of new nodes added to the network system

β conversion rate of nodes to bots

τ probability of abandoning bots driving by the command and control propagation behavior.

α rate of returning abandon bots back to the susceptible class by removing the malware packet.

γ probability of selecting forensic nodes drive by the activities of law enforcement

δ rate of transferring forensic nodes back to the susceptible class as a result of data loss due to interaction with other nodes.

θ nodes leaving the system due to mobility or other activities.

By attacking a particular node and successful conversion to a bot. The botmaster will subsequently expand its attack surface to the susceptible IoT sensor nodes. Nodes in the susceptible class S can be lost due to coming in contact with the malware packet based on the botnet conversion rate. The conversion of nodes can be achieved via random scanning of all the sensor nodes in the IoT network. Consequently, the susceptible nodes will transit to the infectious state I. The botmaster's command and control component. Similarly, in the abandoned phase of the botnet life cycle, selected bots can be denied from propagating the attack further to the neighboring nodes as such there will be inactive and enter the latent infectious state E at an abandon rate τ . However, upon identifying the latent infectious nodes, the malware packet can be removed and the nodes transit to the back to the susceptible class. Similarly, for the law enforcement agents, bots selected by the botmaster to further propagate the attack can be of forensic interest. Therefore, infected nodes can be identified from the I state and transits to the forensic state F at a probability complement of abandon rate $\gamma = (1 - \tau)$. At the forensic state, nodes can similarly lose their forensic value due to interaction with other IoT nodes and return back to the susceptible class S at the rate of δ . Finally, at every stage of the model, there is probability θ of removing a node (s) due to mobility and other influential factors that affect WSN.

4.1. Nodes Population & Scoping

IoT network is usually attributed to WSN. WSN consists of small sensing devices with constricted bandwidth, power, and computational capabilities. In this regard, the sensing coverage per sensor node to determine the communication range based on the radius of the sensor. The nodes population is initially deployed in a small area with all the sensor nodes considered to be susceptible. Contrary, to Acrali et al. [1], we consider the infection to randomly scan all the susceptible nodes within the scope of the coverage of the infected node (bot).

We similarly assume that the population consists of a dynamic number of nodes with removing and addition of nodes back to the network.

4.2. Infection rate

Based on the life cycle of a botnet, it starts with the infection process, where the botmaster uses different scanning methods to capture a large number of targets to be infected and converts to bots. In this regard, the bots can be used to propagate the attack to the remaining susceptible nodes. Hence, the infection rate can be proportional to the number of contact nodes per bots per time. Although botnet usually needs to collect as many bots as possible to have a sufficient attack force, a bot can be abandon when it is found that it is no more useful to the botmaster due to certain conditions.

4.2. *Abandon rate*

Hejazi and Ferrari [11] denote ζm ∈ [0, 1] as the fraction of the remaining free memory in sensor nodes during propagation of sensed data; This coefficient can thus be expressed in equation (3) below

$$\zeta_m = \frac{r_m}{i_m} \tag{3}$$

Where r_m is the remaining free memory of each node and i_m is the initially available memory, Hence, the probability of abandoning node τ depends on the size of the malware packet installed on the infected node by the botmaster and the fraction of the remaining free memory of the infected node (bot) during propagation. The processing of malware packet and the abandon probability rate can be mathematically expressed in equation(4) and (5) respectively below:

$$Malware_{process} = Malware \ \ packet \ \ Size \times \zeta_m \tag{4}$$

$$\text{Abandon rate} \quad \tau = \frac{1}{Malware_{\Pr ocess}} \tag{5}$$

4.3. *Forensic rate*

The rate of identifying the object of forensic interest depends largely on the sensor capability to hold data. In the event of malware attack, the malware can spread together with the normal data to the neighboring nodes. Therefore, the nodes selected by the botmaster to propagate the malware due to their memory capabilities can similarly be classified as the object of forensic interest. Therefore, the rate of identifying forensic nodes γ by the law enforcement agents from an infectious class I is the probability complement of the abandon rate τ of the botmaster and is given using the equation (6) below.

$$\text{Forensic rate} \quad \gamma = 1 - \tau \tag{6}$$

4.3. *Data Loss rate*

Certain operations in IoT network associating one node to another which can result in data consumption among neighboring nodes. Hence, nodes can be identified as forensic nodes but due to its associated services can lose its forensic data and return back to its infectious class I without any forensic value. We defined the data loss rate using equation (7) below:

$$\delta = number \ \ of \ \ packet \ \ sent \ \times \ Contact \ \ rate \tag{7}$$

Consequently, if there exists a large number of packets sent by a forensic node with a high contact rate with other devices, the significant number of forensic nodes will lose their forensic value.

5. Stability Analysis

The analysis techniques utilized in most of the malware propagation works are based on stability analysis of the proposed model[4]. The concept generally is to understand the steady-state effects of different parameters in the models using stability analysis. To achieve this, the basic reproduction number R_0 which determines the number of secondary bots produced by a single (typical) infection in a completely susceptible population can first be obtained.

R_0 often serves as a threshold parameter that predicts whether a botnet will spread in an IoT platform, to achieve this we check the stability of the model based on botnet-free and botnet-endemic stability states.

5.1. Determination of basic reproductive number R_0

To generate R_0 from the mathematical model of equation(2), we considered states consists of infectious parameter, which include the following equations.

$$\begin{cases} \dfrac{dI}{dt} = \beta S(t)I(t) - (\tau + \gamma + \theta)I(t) \\ \dfrac{dE}{dt} = \tau I(t) - (\theta + \alpha)E(t) \end{cases}$$

Next is to determine the infectious and transition parameters in matrix form. We denote F and V as matrices for the infectious and transition parameters respectively, defined in equation (8) and (9) respectively.

$$F = [\beta SI] \tag{8}$$

$$v = \begin{bmatrix} (\tau + \gamma + \theta)I \\ -\tau I + (\theta + \alpha)E \end{bmatrix} \tag{9}$$

Taking derivatives of F and V with respect to I and E, F and V can be respectively transformed in equation (10) and (11)

$$F = \begin{bmatrix} \beta S & 0 \\ 0 & 0 \end{bmatrix} \tag{10}$$

$$V = \left[\begin{bmatrix} (\tau + \gamma + \theta) & 0 \\ -\tau & (\theta + \alpha) \end{bmatrix}\right] \tag{11}$$

Obtaining V^{-1} as in equation (12) and (13)

$$V^{-1} = \frac{1}{(\theta + \alpha)(\tau + \gamma + \theta)} \begin{bmatrix} (\theta + \alpha) & 0 \\ \tau & (\tau + \gamma + \theta) \end{bmatrix} \tag{12}$$

$$V^{-1} = \begin{bmatrix} \dfrac{1}{(\tau + \gamma + \theta)} & 0 \\ \dfrac{\tau}{(\theta + \alpha)(\tau + \gamma + \theta)} & \dfrac{1}{(\theta + \gamma)} \end{bmatrix} \tag{13}$$

Multiplying F and V^{-1} and obtained equation (14) below

$$FV^{-1} = \begin{bmatrix} \dfrac{\beta S}{(\tau + \gamma + \theta)} & 0 \\ 0 & 0 \end{bmatrix} \tag{14}$$

Then the basic reproduction number R0 is the largest eigenvalue of equation (14) which can be expressed in equation (15)

$$R_0 = \frac{\beta S}{(\tau + \gamma + \theta)} \tag{15}$$

Considering botnet free equilibrium state when I(0)=0, E(0)=0, and F(0)=0, then S(0) $=\frac{\lambda}{\theta}$, then the basic reproductive number is given in equation (16) below.

$$R_0 = \frac{\beta \lambda}{\theta(\tau + \gamma + \theta)} \tag{16}$$

If the value of R0 < 1 the botnet propagation will be eliminated within the IoT wireless network and the proposed model will stabilize at botnet-free equilibrium, else if R0 > 1, the botnet will propagate consistently within the IoT network and the proposed model will stabilize at botnet-endemic equilibrium.

5.2. Botnet-free equilibrium stability state

To determine the system stability at botnet-free equilibrium state, we will assume $I(t) = 0;$ meaning no malware attack exist on the network and $S(0) = \frac{\lambda}{\theta}$ meaning the entire nodes on the network are susceptible. And all other states $E(0) = 0;$ and $F(0) = 0$ are considered to be zero. Then, the system can be express as in equation (17) below:

$$\begin{bmatrix} -(\beta I + \theta) & -\beta S & \alpha & \delta \\ \beta I & \beta S - (\tau + \gamma + \theta) I & 0 & 0 \\ 0 & \tau & -(\theta + \alpha) & 0 \\ 0 & \gamma & 0 & -(\delta + \theta) \end{bmatrix} \tag{17}$$

Subs

tituting $\left(\bar{S}\,\bar{I}\,\bar{E}\,\bar{F}\right) = \left(\frac{\lambda}{\theta}, 0, 0, 0\right)$ into (17) we will have the matrix solution of equation (18)

$$\begin{bmatrix} -\theta & -\beta\frac{\lambda}{\theta} & \alpha & \delta \\ 0 & \beta\frac{\lambda}{\theta} & 0 & 0 \\ 0 & \tau & -(\theta + \alpha) & 0 \\ 0 & \gamma & 0 & -(\delta + \theta) \end{bmatrix} \tag{18}$$

Next is to determine the jacobian matrix of equation (18), and expressed in equation (19) below.

$$\det(J(\frac{\lambda}{\theta}, 0, 0.0) - \eta I) = \begin{bmatrix} -\theta - \eta & -\beta\frac{\lambda}{\theta} & \alpha & \delta \\ 0 & \frac{\beta\lambda}{\theta} - \eta & 0 & 0 \\ 0 & \tau & -(\theta + \alpha) - \eta & 0 \\ 0 & \gamma & 0 & -(\delta + \theta) - \eta \end{bmatrix} \tag{19}$$

From equation (19) we have $\eta_1 = -\theta$, $\eta_2 = -(\theta + \alpha)$, $\eta_3 = -(\delta + \theta)$ and $\eta_4 = \frac{\beta\lambda}{\theta}$. hence for the eigenvalue, η_1, η_2, η_3 are all negative values indicate that the botnet free-equilibrium is locally asymptotically stable at $S(0) = \frac{\lambda}{\theta}$. However, for the eigenvalue η_4, the system can only be stable if $R_0(\tau + \gamma + \theta) < 1$, otherwise $S(0) = \frac{\lambda}{\theta}$ remain unstable.

5.3. Botnet-endemic equilibrium stability state

In the botnet endemic equilibrium, we assume $I(t) = I$, meaning there are existing bots in the network. In this case

$$\left(\bar{S}\,\bar{I}\,\bar{E}\,\bar{F}\right) = \left(0, \bar{I}, 0, 0\right)$$

next is to substitute the values into equation (17) and determine the Jaccobian of the matrix which can be expressed in equation (20).

$$\det(J(0,\bar{I},0,0) - \eta I) = \begin{bmatrix} -(\beta\bar{I} + \theta) - \eta & 0 & \alpha & \delta \\ \beta\bar{I} & -(\tau + \gamma + \theta)I - \eta & 0 & 0 \\ 0 & \tau & -(\theta + \alpha) - \eta & 0 \\ 0 & \gamma & 0 & -(\delta + \theta) - \eta \end{bmatrix} \quad (20)$$

Similarly, from the matrix of equation (20), all the eigenvalues have negative value, it follows that the botnet-endemic equilibrium is locally asymptotically stable for all the value of I .

6. Results

6.1. Numerical Simulation

To understand the impact of free memory efficiency in IoT botnet propagation based on botmaster command and control behaviour, as well as the forensic capability in mitigating the attack, we outlined the values of our model parameters as shown in Table1 and Table2 based on assumed scenario A and B

Table 1. Scenario A

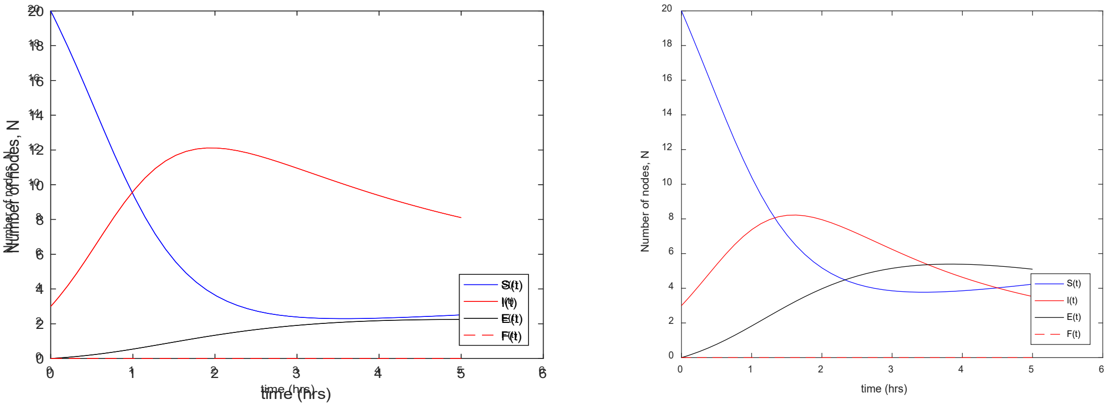
Parameter	**Value**
S(t)	20
I(t)	3
E(t)	0
F(t)	0
Av(r_m)	20mb
i_m	100mb
Malware Packet size	50mb
τ	0.1
γ	0

Table 2. Scenario B

Parameter	Value
S(t)	20
I(t)	3
E(t)	0
F(t)	0
Av(r_m)	5mb
i_m	100mb
Malware Packet size	50mb
τ	0.4
γ	0

doi:10.20944/preprints201912.0097.v1

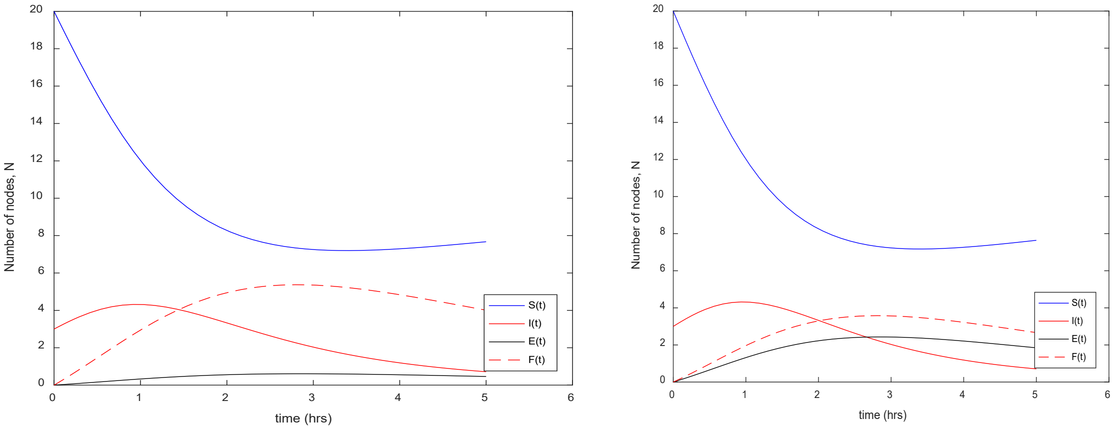
While other model parameters are fixed with values(β =0.106, λ =3, δ =0.075, α =0.06, and θ =0.301) the resulting botnet attack scenarios and analysis are shown in Figure 4 below:



Scenario(**a**) Scenario (**b**)

Figure 4. typical botmaster command and control behavior based on memory efficiency without the involvement of forensic parameter as a mitigating factor: (**a**) with three bots having an average of 20mb free memory, the botmaster abandon less nodes and propagate the attack to more than 12 nodes(red)(I(t); (**b**)with three nodes having an average of 5mb free memory, the botmaster abandon more nodes and propagate the attack to only 8 nodes at peak (red)I(t).

In the previous scenario, we consider the propagation based on memory efficiency and the botmaster command and control behaviour without the impact of forensic parameter as a mitigating factor. Now, we consider forensic parameter with $\gamma = 1 - \tau$ and applied to scenario A and B. The resulting impact of forensic parameter as a mitigation techniques on IoT botnet propagation is shown in Figure 5 below:



Scenario(**a**) Scenario (**b**)

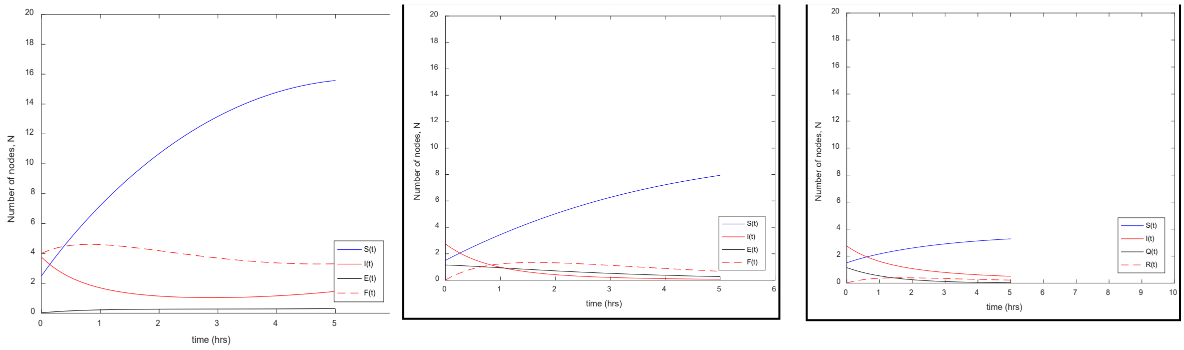
Figure 5. typical botmaster command and control behavior based on memory efficiency with the involvement of forensic parameter as a mitigating factor: (**a**) with three bots having an average of 20mb free memory, the botmaster abandon less nodes and propagate the attack to only 4 nodes at peak(red)(I(t) while most bots were absorbed to forensic compartment F(t)(dotted red line) that suppressed propagation further; (**b**)with three nodes having an average of 5mb free memory, the botmaster abandon more nodes and propagate the attack to only 4 nodes at a peak (red)I(t) at a time, this is as result of absorbing bots to forensic compartment that suppress the propagation further. Thanks to forensic parameter that limits the propagation of bots at standstill no matter how the botmaster command and control behaviour.

6.2. Comparison with other Model

In this section, we are to examine the impact of our model parameters on the peak value and the peak period of IoT botnet propagation based on the value of our R_0. The numerical results generated from the MatLab based on the initial value obtained from the work of Khanh [8] are represented in Figures 6 below.

Figure 6a shows the time series solution of the model with R_0>1. For $\lambda=7, \delta=0.075, \alpha=0.06, \tau=0.1$ $\gamma=1-\ \tau$ and $\theta=0.295$. Then the value of our R_0=1.8324>1. With the initial condition given as (S(0)=2.5, I(0)=3.75, E(0)=0.025, F(t)=4.0)[8]. In this case, the botnet-endemic equilibrium is locally asymptotically stable. Meaning that the number of bots generated tends to be positive value 1.888 as time tends $t \rightarrow \infty$ which shows that the botnet propagation will persist within the IoT network.

Figure 6(b) and 6(c) shows the time series solution of the proposed IoT SIEF model and SIQR model [8] with R_0<1, For $\lambda=3, \delta=0.075, \alpha=0.06,\ \tau=0.1,\ \gamma=1-\ \tau$ and $\theta=0.301$.. Then the value of our R_0=0.8133<1 and that of SIQR R0=0.8138<1[8]. With the initial condition (S(0)=1.5, I(0)=2.75, E(0)=1.15, F(t)= 0.01) [8]. In this regard, the botnet-free equilibrium is locally asymptotically stable. Meaning in both our proposed and SIQR models, the number of bots or infections nodes I(t) approaches 0 as time $t \rightarrow \infty$, this shows that the botnet propagation will be eliminated within the IoT network.

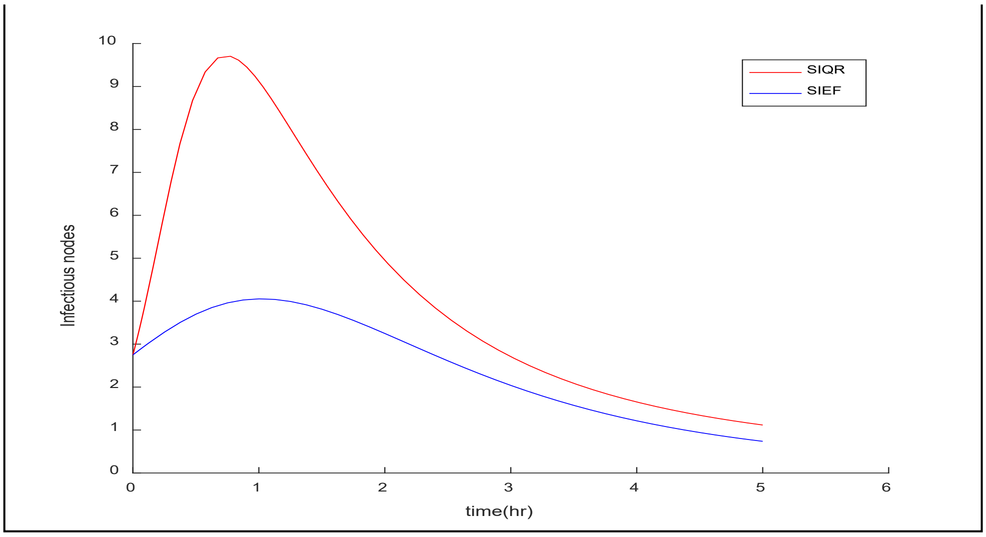


(**a**) (**b**) (**c**)

IoT SIEF Model with R_0>1.	IoT SIEF Model with R_0<1.	SIQR Model[8] with R_0<1.

Figure 6. IoT botnet propagation with respect to value of R_0 (**a**) with R_0>1, the number of bots declined from peak value 3.75 (red)(I(t) downward to zero and raises again to 1.5 bots as $t \rightarrow \infty$, which shows that the botnet propagation will persist.(**b**)with R_0<1, the number of bots declined from peak value 2.75 (red)(I(t) downward to zero as $t \rightarrow \infty$ which shows that the propagation will be eliminated(**c**) Similarly, with R_0<1, using SIQR model, the number of bots declined from peak value 2.75 (red)(I(t) downward to zero as $t \rightarrow \infty$ which shows that the propagation will be eliminated also.

However, in large-scale IoT deployment, our proposed IoT-SIEF model will efficiently mitigate IoT botnet propagation with R_0<1 than SIQR model. Figure 7 below shows the transient response of our proposed IoT-SIEF model and SIQR model with respect to the number of secondary infectious nodes I(t) for an increase number of susceptible nodes from 1.5 to 20 while other parameters remain constant. That is (S(0)=20, I(0)=2.75, E(0)=1.15, F(t)= 0.01)[8].



(**Figure** 7)

Figure 7. Comparison of the proposed IoT-SIEF model with SIQR model in mitigating botnet propagation. With the value of R_0=0.8133<1 of our proposed IoT-SIEF model and that of SIQR R_0=0.8138<1[8], our proposed model mitigate the number of secondary bots with a peak value of 4 bots and delayed the peak period of the propagation to 1hr before decaying the number of bots to zero. However, SIQR model mitigates the number of secondary bots at peak value of 10 nodes while delaying the propagation peak period to less than 1hr before decaying the number of bots to zero. Consequently, our model can mitigate the botnet propagation in large scale IoT network deployment with more than 50% efficiency in comparison with SIQR model.

7. Discussion

Based on our results, our findings have some implications both to the malicious and defensive actors. To maximize the number of target bots, the botmaster needs to know the nodes with high free memory space that can process and propagates malware packets faster to the neighboring nodes. Thereby abandoning nodes with low processing capabilities that will not be of significant in terms of processing and propagation of the botnet attack. For the defense actors, the model highlight the implication of the botmaster behavior and the effect of high powered free memory sensor nodes. By understanding the command and control behavior of the botmaster and the effect of high powered memory sensor nodes, the defense actors should determine the kind of sensor nodes to be deployed in small or large scale IoT WSN. Country to the existing approaches that minimize the frequency of contacts between infected and susceptible nodes in mitigating the number of secondary infections. However, we found that propagation can similarly be successful if the botmaster prioritizes sensor nodes with higher free memory available for processing of malware packets while abandoning those with low memory space. Sensor nodes with free memory space process malware codes faster, hence, instead of maximizing contact rate to the nodes with low memory space, it will be wise for the botmaster to choose bots with higher free memory for processing the malware packet faster and propagate to the neighboring nodes.

From investigation perceptions, defense actors can simply identify those infected nodes that can be of interest to the botmaster. As such those nodes should be identified and prioritize as an object of forensic interest and should be prevented from further exploitation by the botmaster. Therefore, the number of target nodes by the malicious actors can be curtail from infecting large-scale number of nodes by decreasing the infection peak values and limit the propagation by delaying the infectious peak period. Consequently, with the likelihood of large scale deployment of IoT WSN in the near future, the proposed IoT-SIEF model can play a significant role in dealing with the threats of botnet attack. Hence, in the future, we need to develop an artificially intelligent system that can predict the command and control behavior and instill into the IoT WSN to prioritize nodes that can be of forensic interest to the investigators.

8. Conclusions

Considering the memory efficiency of a node as an important factor in IoT wireless network performance, it will be of great importance for botnet propagation and defending strategy. However, existing IoT botnet propagation models and defense strategies have not given any consideration to node's memory efficiency. In this paper, we proposed an IoT-SIEF model with forensic capability that incorporates memory efficiency in addition to other influential factors to determine the botmaster propagation behavior in IoT botnet formation. The model shows that memory efficiency can efficiently affect botnet propagation by enabling the botmaster to select or abandon bots in propagating the attack further. Similarly, the forensic add-on of the model serves as a defense strategy that mitigates the propagation of secondary bots peak value and delayed the propagation peak period. Meanwhile, the simulation results show that the model will be applicable to large-scale IoT deployment networks by defending the network against the impact of IoT botnet propagation and as well generate the object of forensic interest.

Acknowledgments: Above work was supported partly by the Universiti Putra Malaysia (UPM)

Conflicts of Interest: The Authors declare no conflict of interest

References

1. Acarali, D.; Rajarajan, M.; Komninos, N.; Zarpelão, B.B. Modelling the Spread of Botnet Malware in IoT-Based Wireless Sensor Networks. Security and Communication Networks, **2019,**2019, 1-13.
2. Tang, S.; Mark. B.L. Analysis of virus spread in wireless sensor networks: An epidemic model. In 7th International Workshop on Design of Reliable Communication Networks, Washington DC, USA, 25-28 Oct. 2009; IEEE.
3. Antonakakis, M.; April, T.; Bailey, M.; Bernhard, M.; Bursztein, E.; Cochran, J.; Durumeric, Z.; Halderman, J.A.; Invernizzi, L.; Kallitsis, M. Understanding the mirai botnet. In 26th {USENIX} Security Symposium ({USENIX} Security 17), Vancouver, BC, Canada, August 16-18, 2017;.
4. Gardner, M.T., C. Beard, and D. Medhi. Using SEIRS epidemic models for IoT botnets attacks. In Proceeding of 13th International Conference on Design of Reliable Communication Networks, Munich, Germany, March 08-10; 2017; VDE: Berlin, Germany, 2017; pp.62-69.
5. Farooq, M.J.; Zhu, Q. Modeling, analysis, and mitigation of dynamic botnet formation in wireless iot networks. IEEE Transactions on Information Forensics and Security, **2019,** 14, 2412-2426.
6. Mishra, B.K.; Keshri, N. Mathematical model on the transmission of worms in wireless sensor network. Applied Mathematical Modelling, **2013,** 37, 4103-4111.
7. Feng, L.; Song, L.; Zhao, Q.; Wang, H. Modeling and stability analysis of worm propagation in wireless sensor network. Mathematical Problems in Engineering, **2015,** 2015,1-8.
8. Khanh, N.H. Dynamics of a Worm Propagation Model with Quarantine in Wireless Sensor Networks. Appl. Math, **2016,** 10, 1739-1746.
9. Wang, T.; Wu, Q.; Wen, S.; Cai, Y.; Tian, H.; Chen, Y.; Wang, B. Propagation modeling and defending of a mobile sensor worm in wireless sensor and actuator networks. Sensors, **2017.** 17, 1-17.
10. Ji, Y.; Yao, L.; Liu, S.; Yao, H.; Ye, Q.; Wang, R. The Study on the Botnet and its Prevention Policies in the Internet of Things. In Proceeding of 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD)), Nanjing, China, 9-11 May, 2018; IEEE.pp.837-842.
11. Hejazi, P.; Ferrari, G. Energy and Memory Efficient Data Loss Prevention in Wireless Sensor Networks. 2018.Sensors, **2017,**1-18.

12. Hethcote, H.W. An immunization model for a heterogeneous population. Theoretical population biology, **1978,**14, 338-349.
13. Kephart, J.O.; White, S.R. Directed-graph epidemiological models of computer viruses, in Computation: the micro and the macro view.World Scientific. **1992**, 71-102.
14. Kocakulak, M.; Butun I. An overview of Wireless Sensor Networks towards internet of things. In proceeding of 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 9-11 January, 2017; IEEE: 2017
15. Yick, J.; Mukherjee, B.; Ghosal, D. Wireless sensor network survey. Computer networks, **2008,** 52, 2292-2330.
16. Silva, S.S.; Silva, R.M.; Pinto, R.C.; Salles, R.M. Botnets: A survey. Computer Networks, **2013,** 57, 378-403.
17. Eslahi, M.; Salleh, R.; Anuar, N.B. Bots and botnets: An overview of characteristics, detection and challenges. In Procceeding of IEEE International Conference on Control System, Computing and Engineering, Penang, Malaysia, 23-25 November, 2012; IEEE: 2013, pp. 349-354.
18. Kok, J.; Kurz, B. Analysis of the botnet ecosystem. In Proceeding of 10th Conference of Telecommunication, Media and Internet Techno-Economics (CTTE), Berlin, Germany, 16-18 May, 2011; VDE.
19. Schiller, C.; Binkley, J.R. Botnets: The killer web applications, 1st ed.; Elsevier: www.sciencedirect.com/book/9781597491358/botnets, 2011; pp.1-480
20. Hachem, N.; Mustapha, Y.B.; Granadillo, G.G.; Debar, H. Botnets: lifecycle and taxonomy. In Proceeding of Conference on Network and Information Systems Security, La Rochelle, France, 18-21 May, 2011, IEEE.
21. Kolias, C., Kambourakis, G.; Stavrou, A.; Voas, J. DDoS in the IoT: Mirai and other botnets. Computer, **2017,** 50, 80-84.
22. Mansouri, M;, Sardouk, A.; Merghem-Boulahia, L.; Gaiti, D.; Snoussi, H.; Rahim-Amoud, R.; Richard, C. Factors that may influence the performance of wireless sensor networks. Smart Wireless Sensor Networks, **2010,** 29-51.