A Novel Worm Propagation Model Considering the Connected State and Trap Mechanism of a Honeypot

Hanlun Li, Jianguo Ren*

Abstract—**Internet worms pose a serious threat to network security due to their automated scanning and propagation strategies. In this paper, we first propose a novel network worm attack model of SCEIRS based on worm behavior. A novel connected state is established to represent machines completely compromised by a connection scan but not infected. The connection rate and code delivery rate combined with this new state are introduced. Compared with the previous SEIRS model, the SCEIRS model can reflect the accurate transition behaviors of machines in the target acquisition (scanning a target machine and attempting to establish a connection before actual infection behaviors occur) and code delivery stages (attempting to infect a target machine through the delivery of worm code). Moreover, it provides us with a new perspective for analyzing the key parameters that affect worm propagation and the effect of countermeasures on worm containment. Then, we propose the M-SCEIRS model, which combines both the trap and feedback mechanisms of honeypots on the baseline of the proposed new state. In the M-SCEIRS model, the basic reproduction number, equilibria, and stability are also obtained. Numerical results suggest that countermeasures during the target acquisition and code delivery stages are crucial for mitigating worm propagation in the early stages. Furthermore, the two mechanisms of honeypots combined in the M-SCEIRS model are effective for worm control.**

Index **Terms—internet worm, propagation model, honeypot, basic reproduction number, state transition.**

I. Introduction

INTERNET worms are devious malware that can self-replicate and quickly spread across several machines on the internet or within a corporate network [1], [2]. The worm monitors the internet or an internal network for other vulnerable targets from its perch on an infected machine and then spreads to other machines by scanning the entire IPv4 address. In a matter of days, any vulnerable machines connected to the internet would be infected. On July 19, 2001, the Code Red worm began to spread by uniformly scanning the IP address space. Due to infection and the resulting increase in bandwidth consumption, this worm caused substantial disruptions, costing approximately $2 billion in financial losses [3]–[5].

In the fight against internet worms, honeypot technologies are intriguing weapons [6]–[8]. They can be used to reroute malicious worm traffic to dedicated spoof services, catch and analyze worms safely, and finally contain worm propagation across networks. Unfortunately, they are still lacking test experience when used over large networks.

An essential objective has been built to use a mathematical model to examine worm behavior and the effects of safety countermeasures on worm proliferation. Additionally, simulating the influence of honeypots on worm attacks provides critical parameters for controlling worm propagation.

Due to the similarity between biological viruses and malware, researchers have proposed many different biological models to explore the dynamic behaviors and characteristics of malware, including worms in networks in which the machines are considered nodes and their joints represent corresponding communications. The classical SI model [9] is the most basic epidemic propagation model, which consists of only susceptible and infected states. Subsequent models, including SIR [10], [11], SIRS [12], [13], SIQR [14], [15], SEIR [16], and SEIRS [17]–[19], consider more factors, such as the malware characteristics and human intervention (e.g., immunity and quarantine measures). Toutonji *et al.* [20] proposed an SEIRS model that considered accurate locations for dysfunctional nodes and their replacements in state transition. Based on Toutonji's work, J.D. Hernández Guillén *et al.* [21] proposed an improved SEIRS model that considered more realistic parameters related to worm propagation.

All the models assume that nodes from susceptible nodes turn into exposed or infected nodes in an instant. According to Glenn Gebhart [22], a complete worm attack process includes target acquisition (once completed, the nodes will transition from a susceptible to a connected state), delivery of hostile code (once completed, the nodes will transition from a connected to an exposed state), and execution of hostile code (once accomplished, the nodes will transition from an exposed state to an infected state). As a result, the transition from susceptible to exposed nodes is not immediate.

To tackle the shortcomings of the aforementioned models, we propose an accurate worm attack propagation model of SCEIRS, which focuses on the disruption approach in the worm propagation chain. A new "connected state" is introduced in this model, which depicts nodes that are

Manuscript received March 30, 2022; revised October 27, 2022.

This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20201462, the Natural Science Foundation of Xuzhou City under Grant KC21018, and the Postgraduate Research & Practice Innovation Program of Jiangsu Normal University under Grant 2021XKT1400.

Hanlun Li is a postgraduate student of the College of Computer Science, Jiangsu Normal University, Xuzhou, China (e-mail: 654818307@qq.com).

Jianguo Ren is an associate professor of the Research Center for Complex Networks and Swarm Intelligence, Jiangsu Normal University, Xuzhou, China (phone: +8613775848013; e-mail: rjgrjgrjgrjg@126.com).

connected to infected nodes but are not infected. Unlike other worm propagation models, we explore corresponding countermeasures for disrupting worm propagation in the connected state.

We integrate honeypots as state variables and combine both the trap and feedback mechanisms into a new model of M-SCEIRS to explore the influences of each mechanism on worm propagation when the connected state is introduced. The results show that honeypots, when used in combination with a complete worm attack propagation model, have the overall effect on containing worms.

This paper is organized as follows. Section 2 introduces the SCEIRS and M-SCEIRS models. The basic reproduction number and the global stability of the worm-free equilibrium are investigated in Section 3. In Section 4, numerical simulations and suggestions are presented. Finally, Section 5 concludes the paper.

II. Model Formulation

A. SCEIRS Model Considering the Connected State

Safety countermeasures are an essential component that affects worm propagation [23]. Thus, measuring their effects is of great importance for conquering worms and preventing their outbreaks in an early phase. To complete this task, an accurate model that can depict the process of a worm attack is needed. Algorithm 1 illustrates a typical scan-based worm attack and propagation algorithm.

Algorithm 1 worm attack and propagation process

1. Generate an IP address
2. Sending a TCP/SYN packet to a machine randomly
   If a TCP SYN-ACK packet is received, then
   Accomplish the three-way handshake and establish a connection with target machine
   Else
   Return to 2
3. Deliver hostile code to connected machine
4. Induce connected machine to execute hostile code
5. Newly infected machine starts from 1

Our proposed model is based on the SEIRS models proposed by Toutonji *et al.* [20] and J.D. Hernández Guillén *et al.* [21], in which the concept of network node dysfunction occurs in an infectious state, and the replacement of dysfunctional nodes occurs in a recovered state, which is consistent with a real network environment. The SEIRS models presented by Toutonji *et al* and J.D. Hernández Guillén *et al*., similar to other worm attack propagation models, have two main issues. First, the transition of nodes from a susceptible state to an exposed state is not immediate since a connection scan needs to be established before any infections. Furthermore, the impact of countermeasures on the target acquisition phase and code delivery phase should be considered. Second, the use of an abstract infection coefficient to illustrate the worm attack mechanism is overly simplistic.

The arguments above motivate us to develop a more suitable SCEIRS model for worm attacks that considers a new "connected state." The SCEIRS model classifies nodes as being in one of five different stages, and any node can potentially be in any of these stages at any time. The states of nodes in our model are defined as follows:

(1) Susceptible (*S*): including nodes that are vulnerable to worm attack.

(2) Connected (*C*): including nodes that have established a connection with infected nodes but have not yet been infected by worms.

(3) Exposed (*E*): including nodes that have been infected but have not yet executed hostile codes and thus are not actively infectious.

(4) Infected (*I*): including nodes that are actively scanning or infecting new victims.

(5) Recovered (*R*): including nodes that have been patched and thus immune to a worm attack temporarily.

The state-transition rules of the proposed SCEIRS model are shown in Figure 1, and the respective definitions of the variables and parameters involved are shown in Table 1.

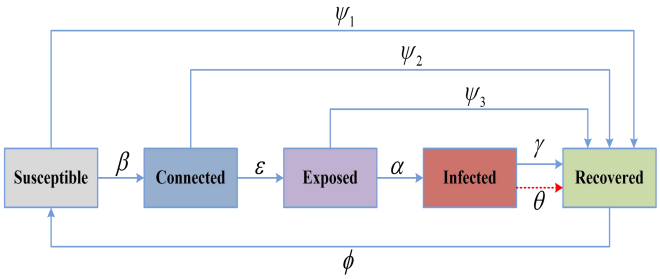


Fig. 1. The state-transition rules of the SCEIRS model.

TABLE I
Notation and Explanation for Proposed Models

Notation	Explanation
$S(t)$	number of susceptible nodes at time t
$C(t)$	number of connected nodes at time t
$E(t)$	number of exposed nodes at time t
$I(t)$	number of infected nodes at time t
$R(t)$	number of recovered nodes at time t
$M_A(t)$	number of susceptible honeypots at time t
$M_B(t)$	number of infected honeypots at time t
N	total number of nodes
M	total number of honeypots
β	connection rate between susceptible and infected nodes
β_1	connection rate under the effect of honeypots between susceptible and infected nodes
β_2	connection rate between infected nodes and honeypots
ε	code delivery rate
α	code execution rate
γ	transition rate from infected to recovered nodes
θ	dysfunctional rate
ψ_1	transition rate from susceptible to recovered nodes
ψ_2	transition rate from connected to recovered nodes
ψ_3	transition rate from exposed to recovered nodes
ϕ	transition rate from recovered to susceptible nodes
μ	feedback rate
λ	replacement rate

To model Code Red, [24] define χ as the average probability that an infected node hits a specific IP address in the scanning space per second, and χ is characterized as

$$\chi=\frac{\eta}{\Omega}, \tag{1}$$

where η denotes the average number of scans an infected node sends out per unit time, and Ω denotes the scanning space, i.e., the entire IPv4 address space ($\Omega=2^{32}$). Considering that the transition from a susceptible state to a connected state can be affected by security countermeasures, we define the connection rate as

$$\beta = \chi p = \frac{\eta p}{\Omega}, \tag{2}$$

where p denotes an adjustable constant governed by the network security status, such as some countermeasures that have been deployed. The SCEIRS model for the propagation dynamics of worms is given by the following system of ordinary differential equations:

$$\begin{cases} \frac{\mathrm{d}S}{\mathrm{d}t} = \phi R - \beta SI - \psi_1 S \\ \frac{\mathrm{d}C}{\mathrm{d}t} = \beta SI - (\varepsilon + \psi_2) C \\ \frac{\mathrm{d}E}{\mathrm{d}t} = \varepsilon C - (\alpha + \psi_3) E \\ \frac{\mathrm{d}I}{\mathrm{d}t} = \alpha E - (\theta + \gamma) I \\ \frac{\mathrm{d}R}{\mathrm{d}t} = \psi_1 S + \psi_2 C + \psi_3 E + (\theta + \gamma) I - \phi R \end{cases} \tag{3}$$

The SCEIRS model focuses on the connected state, and its transition and the impact of countermeasures to disrupt the transfer from the susceptible to the connected state, as well as from the connected state to the exposed state, are taken into account.

B. M-SCEIRS Model with the Trap and Feedback Mechanisms of a Honeypot

A honeypot is an ideal network decay to attract attacks by receiving probes from worms. It can be deployed in a variety of locations on a network. Figure 2 shows the deployment blueprint of the honeypot. A honeypot outside the firewall can trap more probes from an external network than from the internal network and does not increase any risk to the internal network. The downside of external honeypots is that they are incapable of capturing intrusions from internal networks. For a honeypot with effective trap mechanisms, it is vital that they are properly deployed.

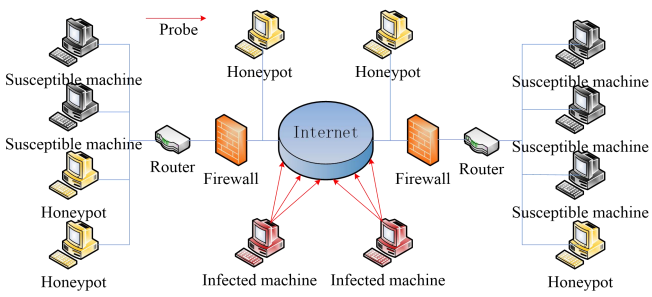


Fig. 2. Honeypot framework under a worm scan

Let κ be a constant determined by the deployment locations of honeypots in the network; then, the connection rate under the effect of honeypots between susceptible nodes and infected nodes is

$$\beta_1 = \frac{[\eta - \kappa M]^+ p}{\Omega}, \tag{4}$$

where $[\eta - \kappa M]^+ = \begin{cases} \eta - \kappa M, & \eta > M \\ 0, & \eta \le M \end{cases}$.

The ideal condition is that all probes are intercepted, and then infected nodes try to infect each of these honeypots. Thus, the attraction mechanism of honeypots not only consumes the attack resources of worms, interrupting the transition from a susceptible state to a connected state, but also serves as the premise to realize the feedback mechanism, which aims to actively send immunization codes to nodes that have not been immune.

Additionally, the connection rate between a honeypot and an infected node is

$$\beta_2 = \frac{\eta \kappa}{\Omega}. \tag{5}$$

Algorithm 2 trap and feedback mechanisms of honeypot

1. Choose a location to deploy and configure the honeypot
2. Keep the honeypot monitor at the network
   If a TCP/SYN packet is received, then
   Accomplish the three-way handshake and establish a connection with the infected machine
   Else
   Continue to perform step 2
3. Analyze the infection behavior of the worm and prepare immunization code
4. Scan the network and remotely patch vulnerable machines

Based on the SCEIRS model, the M-SCEIRS model, which incorporates both a trap mechanism and a feedback mechanism, is proposed. According to the feedback mechanism [25], [26], a honeypot can contain worms on a network scale by transmitting immunizing information to other nodes once it identifies a worm. To our knowledge, however, current models have not accounted for the trap mechanism. Worm probes can be intercepted dramatically, reducing the probability of worms infecting other nodes and preventing the early spread of worms. Using the previously discussed connection rate, we can add the trap mechanism to our model.

The M-SCEIRS model enables us to assess honeypot implementation from a new perspective and offers us a comprehensive strategy for efficiently conquering worms. Honeypots are included as state variables based on the five previously outlined states:

(1) Susceptible honeypots (M_A): including honeypots that have not yet detected a new worm.

(2) Infected honeypots (M_B): including honeypots that have established a connection with infected nodes and received hostile code from them.

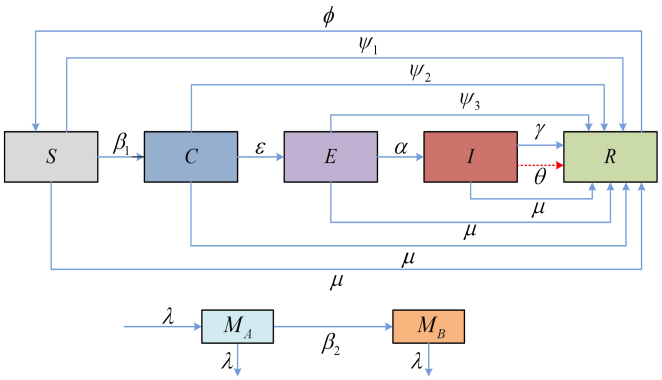


Fig. 3. The state-transition rules of the M-SCEIRS model.

The set of differential equations for the M-SCEIRS model is the following system:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = \phi R - \beta_1 SI - \psi_1 S - \mu S M_B \\ \dfrac{\mathrm{d}C}{\mathrm{d}t} = \beta_1 SI - (\varepsilon + \psi_2) C - \mu C M_B \\ \dfrac{\mathrm{d}E}{\mathrm{d}t} = \varepsilon C - (\alpha + \psi_3) E - \mu E M_B \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \alpha E - (\theta + \gamma) I - \mu I M_B \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = \psi_1 S + \psi_2 C + \psi_3 E + (\theta + \gamma) I + \mu (S + C + E + I) M_B - \phi R \\ \dfrac{\mathrm{d}M_A}{\mathrm{d}t} = \lambda V - \beta_2 I M_A - \lambda M_A \\ \dfrac{\mathrm{d}M_B}{\mathrm{d}t} = \beta_2 I M_A - \lambda M_B \end{cases} \tag{6}$$

III. Model Analysis

A. The Basic Reproduction Number

Using the reduction method, system (6) can be reduced as:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = \phi(N - S - C - E - I) - \beta_1 SI - \psi_1 S - \mu S M_B \\ \dfrac{\mathrm{d}C}{\mathrm{d}t} = \beta_1 SI - (\varepsilon + \psi_2) C - \mu C M_B \\ \dfrac{\mathrm{d}E}{\mathrm{d}t} = \varepsilon C - (\alpha + \psi_3) E - \mu E M_B \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \alpha E - (\theta + \gamma) I - \mu I M_B \\ \dfrac{\mathrm{d}M_B}{\mathrm{d}t} = \beta_2 I (M - M_B) - \lambda M_B \end{cases} \tag{7}$$

Let N and M denote the total number of nodes and honeypots in the network, respectively, which satisfy $N=S+C+E+I+R$ and $M = M_A + M_B$. Then, the compact feasible region of the system (7) is denoted by

$$\Omega=\left\{S, C, E, I, M_B \in \Re_+^5 \,\middle|\, S \geq 0, C \geq 0, E \geq 0, I \geq 0, 0 \leq S + C + E + I \leq N, 0 \leq M_B \leq M\right\},$$

which is a positively invariant set. The worm-free equilibrium of system (7) is

$$E_0 = (S_0, C_0, E_0, I_0, M_{B0}) = \left(\frac{\phi N}{\phi + \psi_1}, 0, 0, 0, 0\right).$$

We can compute the basic reproduction number through the spectral radius of a matrix called the next-generation matrix [27, 28]. Let

$$F = \begin{pmatrix} \beta_1 SI \\ 0 \\ 0 \\ 0 \end{pmatrix}, \; Z = \begin{pmatrix} (\varepsilon + \psi_2) C \\ (\alpha + \psi_3) E - \varepsilon C \\ (\theta + \gamma) I + \mu I M_B - \alpha E \\ \lambda M_B - \beta_2 I (M - M_B) \end{pmatrix},$$

Then

$$\xi = DF\big|_{E_0} = \begin{pmatrix} 0 & 0 & \beta_1 S_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\varsigma = DZ\big|_{E_0} = \begin{pmatrix} \varepsilon + \psi_2 & 0 & 0 & 0 \\ -\varepsilon & \alpha + \psi_3 & 0 & 0 \\ 0 & -\alpha & \theta + \gamma & 0 \\ 0 & 0 & \beta_2 M & \lambda \end{pmatrix}.$$

Then, we have

$$\Re_0 = \rho(\xi \varsigma^{-1}) = \frac{\phi \varepsilon \alpha \beta_1 N}{(\phi + \psi_1)(\varepsilon + \psi_2)(\alpha + \psi_3)(\theta + \gamma)}. \tag{8}$$

B. Global Stability of the Worm-Free Equilibrium

Theorem 1. Suppose that $\Re_0 \leq 1$, then, the worm-free equilibrium E_0 of the system (7) is globally asymptotically stable in Ω.

Proof. Set a Lyapunov function

$$\varpi = \alpha \varepsilon C + \alpha(\varepsilon + \psi_2) E + (\varepsilon + \psi_2)(\alpha + \psi_3) I, \tag{9}$$

differentiation gives

$$\begin{aligned} \dot{\varpi} &= \alpha \varepsilon \dot{C} + \alpha(\varepsilon + \psi_2) \dot{E} + (\varepsilon + \psi_2)(\alpha + \psi_3) \dot{I} \\ &= \alpha \varepsilon \beta_1 SI - (\varepsilon + \psi_2)(\alpha + \psi_3)(\theta + \gamma) I \\ &\leq \left[\alpha \varepsilon \beta_1 \frac{\phi N}{\phi + \psi_1} - (\varepsilon + \psi_2)(\alpha + \psi_3)(\theta + \gamma) \right] I \\ &= (\Re_0 - 1)(\varepsilon + \psi_2)(\alpha + \psi_3)(\theta + \gamma) I. \end{aligned}$$

Hence, if

$\Re_0 = \dfrac{\phi \varepsilon \alpha \beta_1 N}{(\phi + \psi_1)(\varepsilon + \psi_2)(\alpha + \psi_3)(\theta + \gamma)} < 0$, it implies $\dot{\varpi} < 0$ in Ω. By the LaSalle invariance principle [29], we complete the

proof of Theorem 1.

C. Sensitivity Evaluation

In order to eradicate the worms or inhibit their epidemics below a certain level, it is critical to have an overall knowledge of the effects of parameters that determine the basic reproduction number $\Re_0$ on worm propagation. We are therefore interested in studying the rate of change of $\Re_0$ as the parameter values are changed.

From equation (8), the basic reproduction number $\Re_0$ of system (7) depends on the following parameters: $\beta_1, \varepsilon, \alpha, \theta, \gamma, \psi_1, \psi_2, \psi_3$ and ϕ. Using a normalized sensitivity index (NSI), one may estimate the rate of change of $\Re_0$ given a change in the parameter value. $NSI[parm]$ is defined as:

$$NSI[parm] = \frac{parm}{\Re_0} \cdot \frac{\partial \Re_0}{\partial parm}. \tag{10}$$

Then, the NSI of $\Re_0$ with respect to $\beta_1, \varepsilon, \alpha, \theta, \gamma, \psi_1, \psi_1, \psi_3$ and ϕ are

$$NSI[\beta_1] = 1,\ NSI[\varepsilon] = \frac{\psi_2}{\varepsilon+\psi_2},\ NSI[\alpha] = \frac{\psi_3}{\alpha+\psi_3},$$

$$NSI[\theta] = -\frac{\theta}{\theta+\gamma},\ NSI[\gamma] = -\frac{\gamma}{\theta+\gamma},$$

$$NSI[\psi_1] = -\frac{\psi_1}{\phi+\psi_1},\ NSI[\psi_2] = -\frac{\psi_2}{\varepsilon+\psi_2},$$

$$NSI[\psi_3] = -\frac{\psi_3}{\alpha+\psi_3},\ NSI[\phi] = \frac{\psi_1}{\phi+\psi_1}.$$

From these results, we can see that $\Re_0$ decreases as β_1, ε, α, and ϕ decrease or θ, γ, ψ_1, ψ_2, and ψ_3 increase. For illustration, we have computed the NSI for the special case of parameter values listed in Table II. The NSI and corresponding % values in Table III represent the changes in parameter values needed for a 1% reduction in $\Re_0$.

TABLE II
VALUES OF THE PARAMETERS THAT CORRESPOND TO THE WORM-FREE EQUILIBRIUM

Parameter	Value	Unit
β_1	0.00000032783	second^{-1}
ε	0.08	second^{-1}
α	0.03	second^{-1}
θ	0.033	second^{-1}
γ	0.005	second^{-1}
ψ_1	0.0005	second^{-1}
ψ_2	0.0007	second^{-1}
ψ_3	0.00035	second^{-1}
ϕ	0.000005	second^{-1}

From Table III, to get a 1% decrease in the value of $\Re_0$, it is necessary to decrease the values of β_1 and ϕ by 1% and 1.010%. Besides, a 1.010% increase in the values of ψ_1 is required to achieve a 1% reduction in the value of $\Re_0$. Consequently, from the NSI, the optimum approaches of reducing the value of $\Re_0$ are to decrease the connection rate between susceptible and connected nodes (β_1), decrease the transition rate from recovered to susceptible nodes (ϕ), and increase the transition rate from susceptible to recovered nodes (ψ_1), respectively.

TABLE III
NSI OF $\Re_0$ AND CHANGE IN PARAMETER FOR 1% CHANGE IN $\Re_0$

Parameter	$NSI[parm]$	Corresponding % changes
β_1	1	1
ε	0.008674	115.287
α	0.0115	86.730
θ	-0.8684	-1.151
γ	-0.1315	-7.604
ψ_1	-0.9901	-1.010
ψ_2	-0.008674	-115.287
ψ_3	-0.0115	-86.730
ϕ	0.9901	1.010

IV. NUMERICAL EVALUATION

In this section, we first investigate the behavior of the existing SEIRS and propose the SCEIRS model at the worm-free equilibrium point. Then, the effects of worm containment strategies determined by the values of coefficients on the target acquisition and code delivery stages will be evaluated. Next, we simulate the effects of the trap and feedback mechanisms of honeypots on the proposed M-SCEIRS model.

A. The SCEIRS model

The modified SEIRS model for worm propagation proposed by J.D. Hernández Guillén *et al.* [21] is given by the following system:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = \phi R - \dfrac{\omega}{N} SI - \psi_1 S \\ \dfrac{\mathrm{d}E}{\mathrm{d}t} = \dfrac{\omega}{N} SI - (\alpha + \psi_3) E \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \alpha E - (\theta + \gamma) I \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = \psi_1 S + \psi_3 E + (\theta + \gamma) I - \phi R \end{cases} \tag{11}$$

In system (11), the incidence is defined by $\frac{\omega}{N} SI$ (ω is the infection rate), while in the system (3), we define the incidence as:

$$\beta SI = \chi pSI = \frac{\eta p}{\Omega} SI.$$

In the SCEIRS model, the incidence is defined by a precise physical definition that considers the process of establishing connections between worms and target nodes. Using the fourth-order Runge–Kutta (RK4) method, we illustrate the dynamics of the two systems by performing numerical simulations. The experiments focus on actively scanning worms such as Code Red, which uses the scanning strategy to compromise susceptible nodes in the network that have a 2^{32} address space. In the simulations, each infected node launches 4,000 probes per second. The total number of nodes is N = 360,000, with initial values: S (0) = 359,950, C (0) = 0, E(0) = 0, I (0) = 50, and R (0) = 0. The remaining

parameters are as follows: κ = 8, p = 0.88, ω = 1.27, β = 0.0000075, ε = 0.08, α = 0.03, ψ_1 = 0.0005, ψ_2 = 0.0007, ψ_3 = 0.00035, ϕ = 0.000005, θ = 0.033, γ = 0.005, λ = 0.25, and μ = 0.00001. Consequently, we obtain the following basic reproduction number of the system (11):

$$\Re_0 = \frac{\alpha\omega\phi}{(\phi+\psi_1)(\alpha+\psi_3)(\theta+\gamma)} = 0.33 < 1.$$

The reproduction number of system (3) is as follows:

$$\Re_0 = \frac{\phi\varepsilon\alpha\beta N}{(\phi+\psi_1)(\varepsilon+\psi_2)(\alpha+\psi_3)(\theta+\gamma)} = 0.33 < 1.$$

Fig. 4 and Fig. 5 show the dynamics of the system (11) and system (3), respectively. The results show, as expected, that the exposed and infected nodes in Fig. 5 reach their peak more slowly than those in Fig. 4 due to the time it takes for worms to scan and connect with the target nodes. Furthermore, since the chain of worm propagation can be disrupted at the target acquisition and code delivery stages, there is a decline in the scale of infection in Fig. 5 compared with Fig. 4.

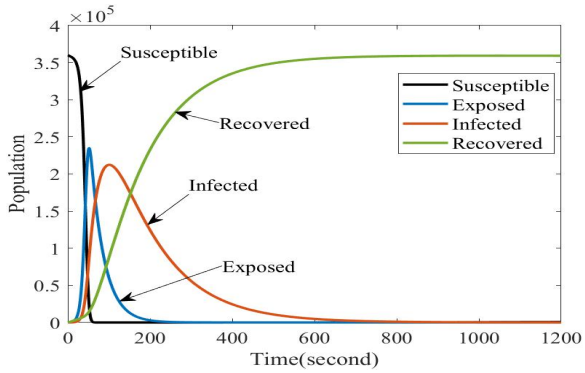


Fig. 4. Worm-free behaviors of the SEIRS model.

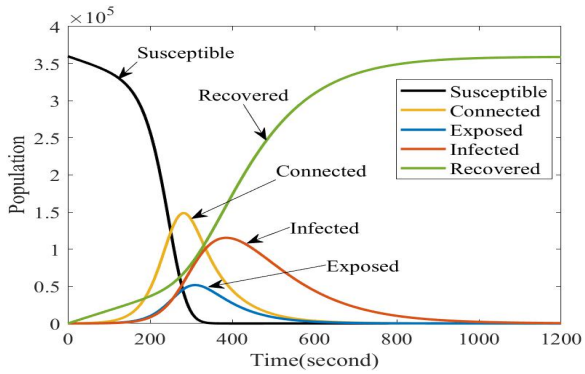


Fig. 5. Worm-free behaviors of the SCEIRS model.

To illustrate the impact of the connected state on the infection outbreaks, we use varying connection rates and code delivery rates in the following time history and three-dimensional phase plots.

Case I: Connection rate

Highlighting the connection rate is our distinction. Worm propagation begins with targeting a system and establishing a connection with it. However, no existing models consider this early factor that influences worm spread at the target acquisition stage. From equation (2), p is an important parameter that affects the connection rate. We denote $(1-p)$ as the interception rate of probes launched by worms. As seen in Fig. 6 and Fig. 7, a lower value of p is beneficial for containing the spread of the worm in terms of the reduction of the speed and number of newly increased connected and infected nodes. To decrease p, countermeasures are needed, such as turning off unneeded services or deploying a firewall on the path between susceptible nodes and infected nodes. Fig. 8 and Fig. 9 show the phase plots for (C, E, I) and (S, I, R) states, respectively, with different values of p.

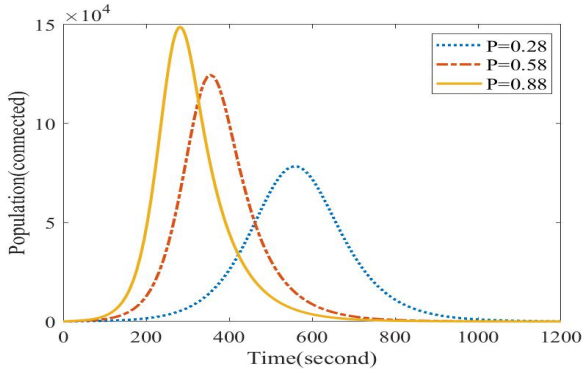


Fig. 6. Effect of interception rate on the connected state.

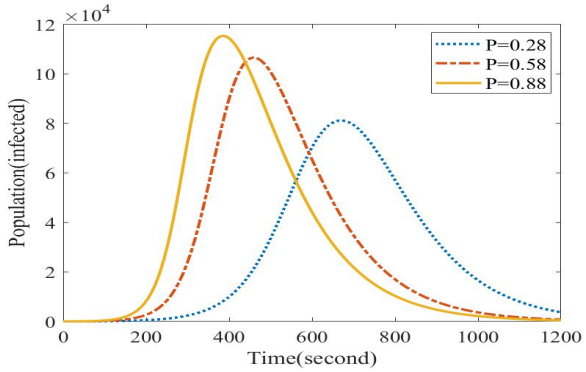


Fig. 7. Effect of the interception rate on the infected state.

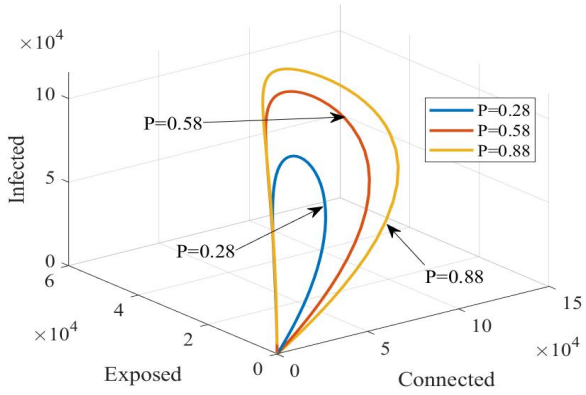


Fig. 8. Phase plot of the SEIRS model (C, E, I).

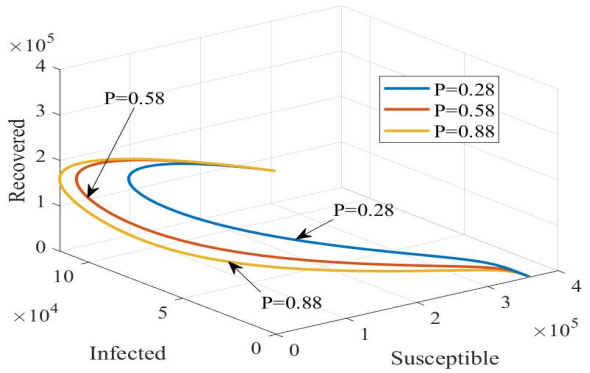


Fig. 9. Phase plot of the SEIRS model (*S*, *I*, *R*).

Case II: Delivery rate of hostile code

After the connection has been established, the next step for the infected node is finding ways to transfer its hostile code to the target system. Fig. 10 and Fig. 11 show the effects of different code delivery rates on connected and infected states. From Fig. 10, a lower code delivery rate results in a longer time for the connected population to reach its peak, and the longer the time, the larger the peak is. In Fig. 11, a lower code delivery rate results in a lower overall number of infected nodes. Furthermore, a spot of decrease in ε results in a considerable reduction in the total number of infected nodes. From this point, it is essential for administrators and users to pay attention to decreasing ε to disrupt the transmission of hostile code, thus containing worm propagation at an early stage. The most common method for this step is to maintain an Intrusion Detection System (IDS). For an unknown worm, an effective IDS can provide a real-time warning mechanism against newly released worms. Fig. 12 and Fig. 13 show the phase plots for the (*C*, *E*, *I*) and (*S*, *I*, *R*) states, respectively, with different values of ε.

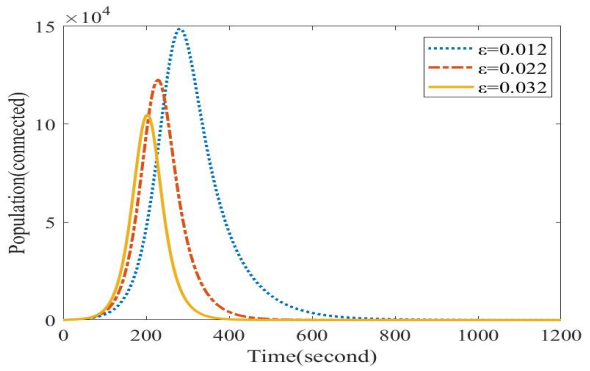


Fig. 10. Effect of the code delivery rate on the connected state.

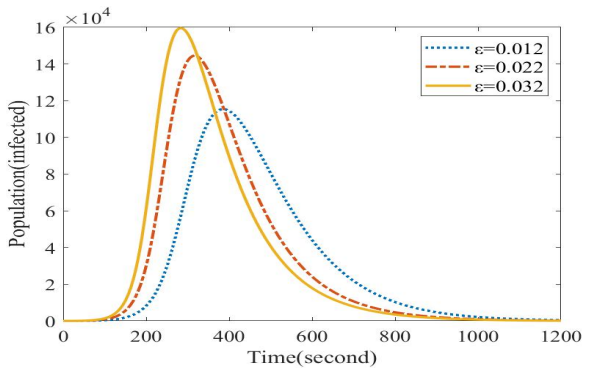


Fig. 11. Effect of the code delivery rate on the infected state.

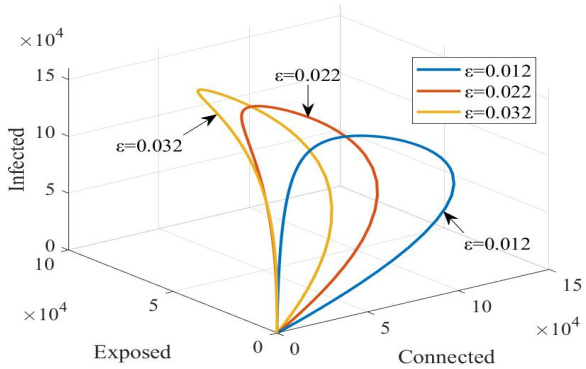


Fig. 12. Phase plot of the SCEIRS model (*C*, *E*, *I*).

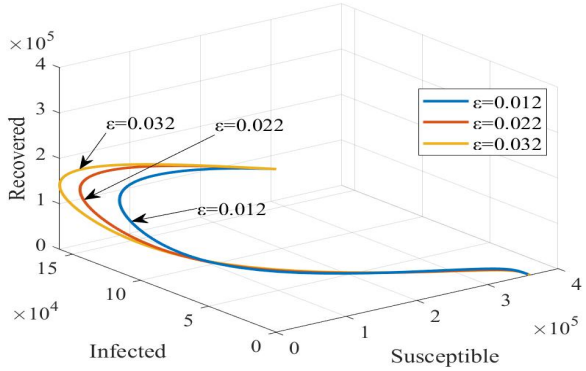


Fig. 13. Phase plot of the SCEIRS model (*S*, *I*, *R*).

The two cases discussed above cover the impact of the connection rate and code delivery rate on the target acquisition and code delivery stages. The simulation results show that worm propagation can be slowed and contained if we consider the corresponding safety measures to disrupt these two stages.

B. Control strategies

Following the target acquisition and code delivery stages, assuming the worm successfully delivered hostile code to the target system, a discussion of containment strategies for continued worm dispersion is presented. From the explicit expression of the basic reproductive number $\Re_0$ of the system (3) and taking into account that $0 < \beta, \varepsilon, \alpha, \theta, \gamma, \psi_1, \psi_2, \psi_3, \phi \leq 1$, the following can be determined:

$$\frac{\partial \Re_0}{\partial \alpha} = \frac{\phi \varepsilon \beta \psi_3 N}{(\phi+\psi_1)(\varepsilon+\psi_2)(\alpha+\psi_3)^2(\theta+\gamma)} > 0.$$

According to Theorem 1, the prevalence of worms in networks is entirely governed by $\Re_0$. In addition to the connection rate and code delivery rate stated above, it should be noted that the code execution rate is a crucial parameter that can reduce the value of $\Re_0$. Consequently, we evaluate the effect of α on the dynamics of worm spread. On the other hand, even if malicious code has resided on the target system, it is still possible to contain the epidemic by taking safety precautions. To obtain effective worm containment solutions, the relationship between α and the maximum number of infected nodes is determined by repeated numerical simulations using the same parameter values in part A. From

the simulation results depicted in Figure. 14, the following conclusions and suggestions can be drawn:

(i) There exists a threshold for α, below which the maximum population of infected nodes sharply increases. Otherwise, the variation in the maximum population of infected nodes is minimal.

(ii) Effective countermeasures are necessary to keep the value of α as low as possible. Since worms use various means to induce a target system to execute malicious code, system administrators must adopt the appropriate methodologies to remain ahead of the attackers. Another strategy for users to decrease the value of α is to refrain from running programs they have no reason to trust, which requires system administrators to empower users with adequate training and reminders.

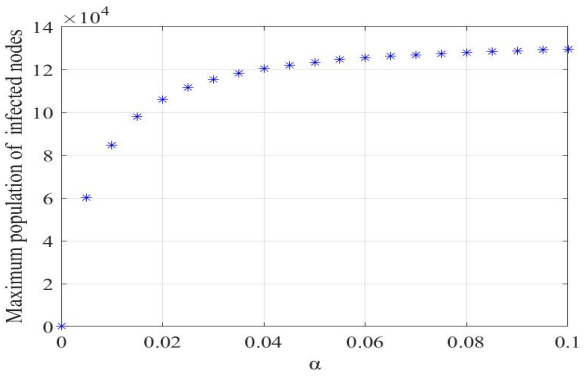
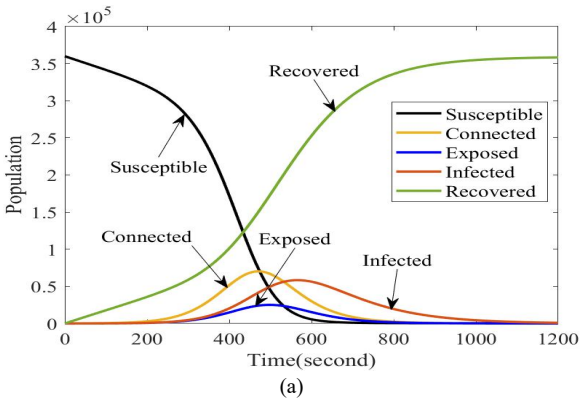


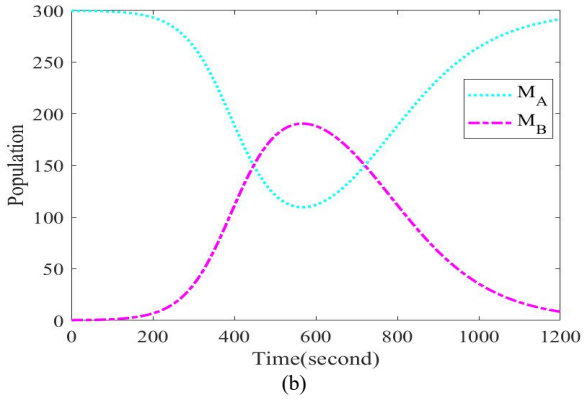
Fig. 14. The relationship between α and maximum population of infected nodes.

C. The M-SCEIRS model

To investigate the impact of honeypots on suppressing worm propagation, we implement the M-SCEIRS model with several simulations. In the M-SCEIRS model, we set $M = 300$, and from (4) and (5), we have the following values: $\beta_1 \approx 3.28 \times 10^{-7}$ and $\beta_2 \approx 7.45 \times 10^{-6}$. The rest of the initial values and parameters are the same as those of the SCEIRS model. Then, we can obtain that $\Re_0 = 0.26 < 1$, which implies that the worm eventually disappears according to Theorem 1. Fig. 15 shows the dynamics of the system (6). Comparing Fig. 15 with Fig. 5, a noticeable reduction occurs in both population and propagation speed for connected, exposed and infected nodes, which confirms the effectiveness of the honeypot for conquering worms. Fig. 16 and Fig. 17 show the phase plots for the (C, I, M_A) and (C, I, M_B) states, respectively, with different initial values of infected nodes.



(a)



(b)

Fig. 15. Worm-free behaviors of the M-SEIRS model.

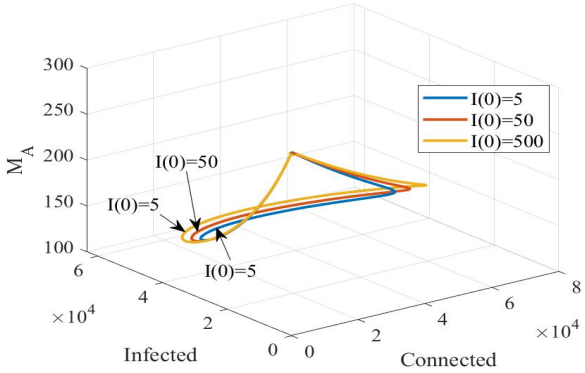


Fig. 16. Phase plot of the M-SCEIRS model (C, I, M_A).

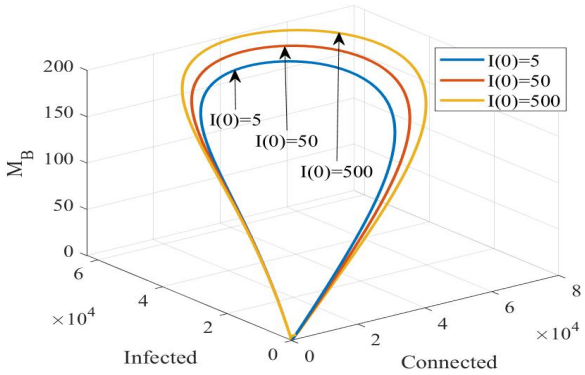


Fig. 17. Phase plot of the M-SCEIRS model (C, I, M_B).

With the baseline of the M-SCEIRS model, we examine the trap mechanism of honeypots whose implementation is subject to the number and location of deployed honeypots in case I and the feedback mechanism of honeypots in case II.

Case I: Number and location of deployed honeypots

Fig. 18 and Fig. 19 show the effect of the number of deployed honeypots on connected and infected nodes, respectively. It can be noted that when we set the number of honeypots to zero, the M-SCEIRS model turns into the SCEIRS model. From Fig. 19, compared with the network without honeypots, when deployed with 300 honeypots, the maximum number of infected nodes is reduced by nearly fifty percent. Fig. 22 and Fig. 23 show the effect of different values of parameter κ on connected and infected nodes, respectively. The result indicates that a reasonable value of κ also has a considerable effect on containing worm spread. Combining Fig. 18 and Fig. 23 with Formula (4), we can conclude that the trap mechanism of the honeypot is effective, and we can improve its efficiency either by increasing the

number of deployed honeypots or choosing a reasonable deployment location. Normally, numerous honeypots can increase the maintenance cost, whereas we can increase κ as much as we can by choosing appropriate deployment locations. Because the value of κ has an overall effect on containing the speed of worm propagation by intercepting the probes sent by worms to increase κ, network administrators must investigate the topology of the entire network and then deploy honeypots at key locations. Fig. 20, Fig. 21, Fig. 24 and Fig. 25 further reveal the effect of the trap mechanism of the honeypot on the dynamics of the M-SCEIRS model by phase plots.

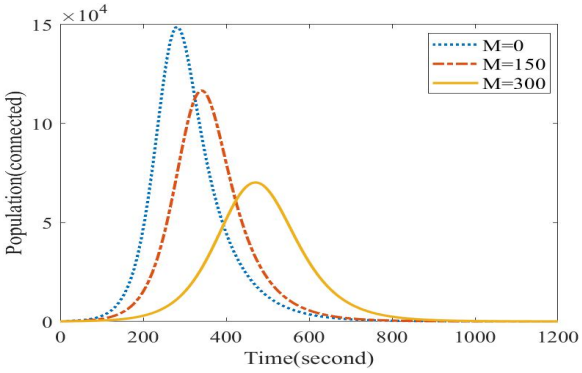


Fig. 18. Effect of the number of deployed honeypots on the connected state.

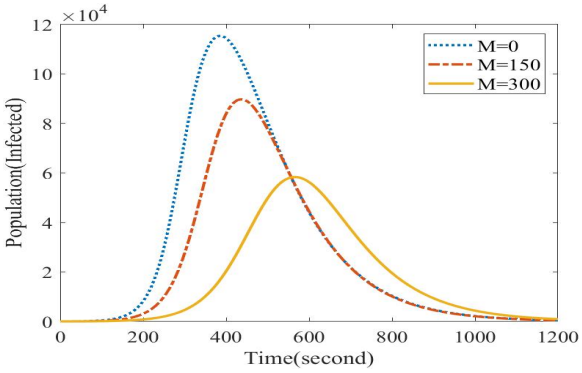


Fig. 19. Effect of the number of deployed honeypots on the infected state.

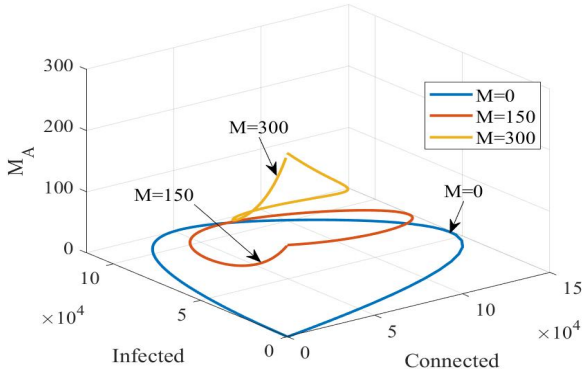


Fig. 20. Phase plot of the M-SCEIRS model (C, I, M_A).

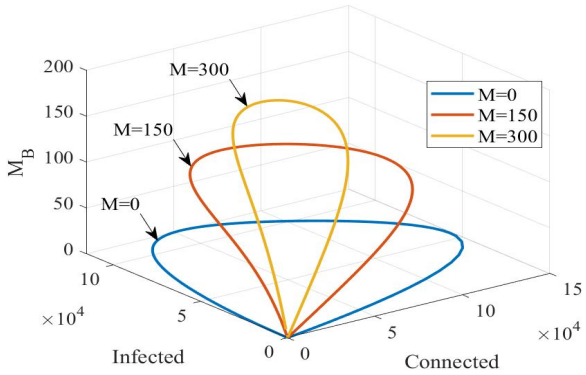


Fig. 21. Phase plot of the M-SCEIRS model (C, I, M_B).

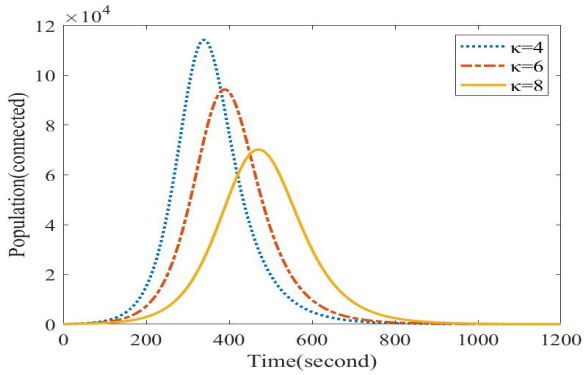


Fig. 22. Effect of the location of the deployed honeypot on the connected state.

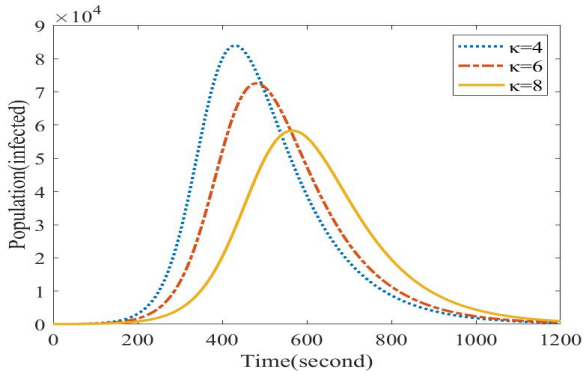


Fig. 23. Effect of the location of the deployed honeypot on the infected state.

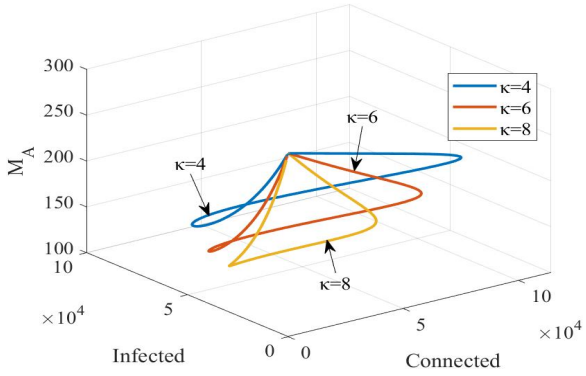


Fig. 24. Phase plot of the M-SCEIRS model (C, I, M_A).

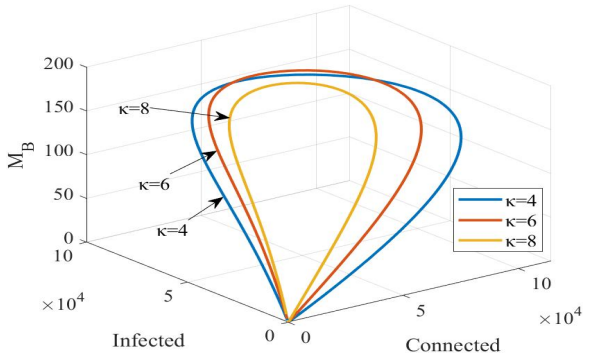


Fig. 25. Phase plot of the M-SCEIRS model (C, I, M_B).

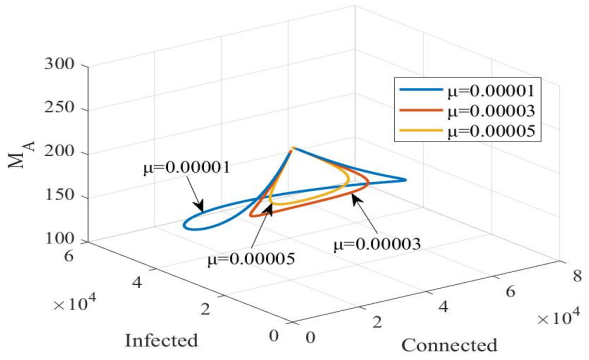


Fig. 28. Phase plot of the M-SCEIRS model (C, I, M_A).

Case II: Feedback rate of honeypots

Fig. 26 and Fig. 27 show the effect of the feedback rate on connected and infected nodes, respectively. In line with expectations, as a result of the higher feedback rate, the maximum number of connected and infected nodes and the speed with which they become connected and infected are both reduced. This result means that after successfully capturing the new worm sample, real-time feedback should be made to ensure that worm diffusion is inhibited. Based on the above analysis, the feedback mechanism of honeypots is crucial for worm control, especially when dealing with a new type of worm. Fig. 28 and Fig. 29 show the phase plots for the (C, I, M_A) and (C, I, M_B) states, respectively, with different feedback rates.

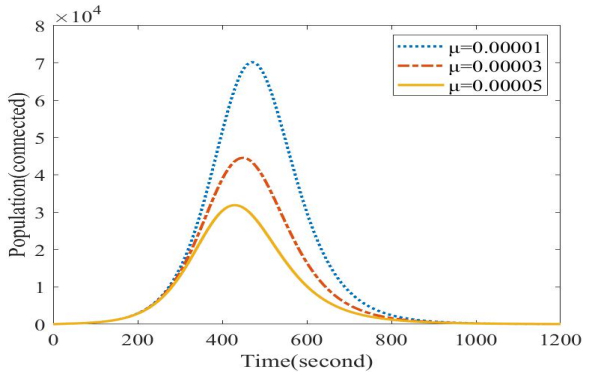


Fig. 26. Effect of the feedback rate on the connected state.

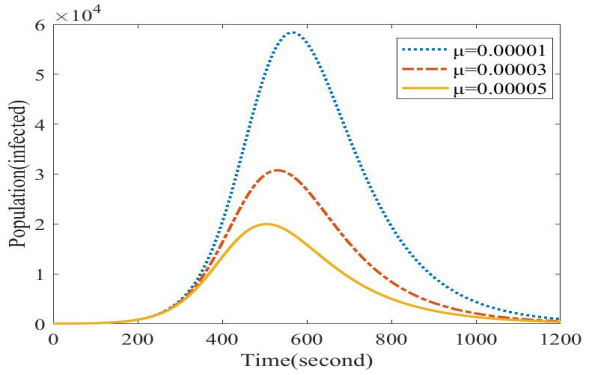


Fig. 27. Effect of the feedback rate on the infected state.

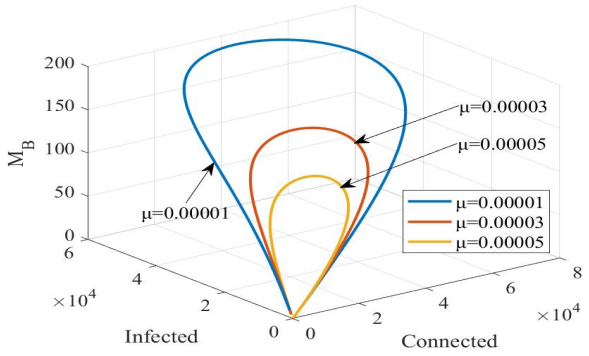


Fig. 29. Phase plot of the M-SCEIRS model (C, I, M_B).

D. Further discussion

The relationship between the two mechanisms of honeypots and the dynamics of the M-SCEIRS model will be investigated. Case I in part C relates the trap mechanism to the number and location of deployed honeypots. By combining equations (4) and (8), we can determine that the value of $\Re_0$ is effected by the parameters κ and M. Denote $\tau(M) = \kappa M$, then take partial derivatives of $\Re_0$ with respect to β_1 and $\tau(M)$, we can obtain:

$$\frac{\partial \Re_0}{\partial \beta_1} = -\frac{\phi \varepsilon \alpha N}{(\phi+\psi_1)(\varepsilon+\psi_2)(\alpha+\psi_3)(\theta+\gamma)} > 0.$$

$$\frac{\partial \Re_0}{\partial \tau(M)} = \frac{\partial \Re_0}{\partial \beta_1}\frac{\partial \beta_1}{\partial \tau(M)} = -\frac{\phi \varepsilon \alpha p N}{(\phi+\psi_1)(\varepsilon+\psi_2)(\alpha+\psi_3)(\theta+\gamma)} < 0.$$

Since $\Re_0$ determines the full dynamics of the M-SCEIRS model, we will discuss how the trap mechanism influences the epidemic that is governed by $\Re_0$ over time. Taking the following set of parameters: p = 5, ε = 0.08, α = 0.03, ψ_1 = 0.0005, ψ_2 = 0.0007, ψ_3 = 0.00035, ϕ = 0.000005, θ = 0.033, γ = 0.005, N = 360,000, and η = 4,000, Figure. 30 demonstrates that $\Re_0$ and $\tau(M)$ have a negative linear relationship. Considering that $\Re_0$ decreases as $\tau(M)$ increases, we can prevent the spread of worms in networks effectively by

(i) deploying honeypots intentionally as opposed to blindly. Investigating the boundary of the network and placing honeypots in strategic areas is required.

(ii) deploying an adequate quantity of honeypots. In Fig. 30, the relationship between threshold value $\Re_0$ and $\tau(M)$ addresses the question of how many honeypots need to be financially provided to control the epidemic.

The expression of $\Re_0$ of the system (7) shows that the feedback mechanism does not affect this threshold value. The

analysis results from Case II in part C reveal, however, that it does inhibit worm diffusion.

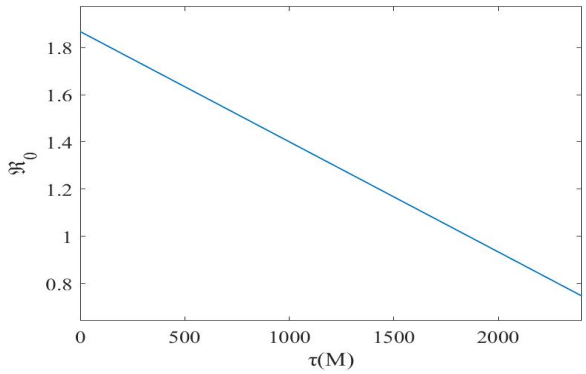


Fig. 30. The effect of $\tau(M)$ on $\Re_0$.

V. Conclusion

To accurately predict worm behavior and the effect of countermeasures on the target acquisition and code delivery phases of worm attacks, the SCEIRS model is established by introducing a new connected state. Based on the SCEIRS model, we defined the connection rate and code delivery rate in the target acquisition and code delivery phases, respectively. Numerical results show that we can mitigate the damage caused by worms by both increasing the interception rate and decreasing the code delivery rate. Moreover, multiple countermeasures are suggested to disrupt the transition both from the susceptible state to the connected state and from the connected state to the exposed state. In the absence of proper worm propagation models and testing experience of honeypots in large-scale networks, we proposed the M-SCEIRS model, which combines both trap and feedback mechanisms to investigate the function of honeypots in worm containment. Mathematically, we obtained the basic reproduction number $\Re_0$ that governs the full dynamics of the M-SCEIRS model. The parameter β_1, which determines the trap mechanism of the honeypot, can greatly affect the value of R_0. In contrast, the feedback rate that determines the feedback mechanism of the honeypot does not affect $\Re_0$. Evaluations indicate that the effect of the trap mechanism, which is determined by the number and location of deployed honeypots, is essential for the early control of worm spread. On the other hand, increasing the feedback rate also has an overall effect on worm containment. In the future, we will combine the honeypot technique with the quarantine method to construct a more efficient worm defense system.

References

[1] Y. Tang, J. Q. Luo, B. Xiao and G. Y. Wei, "Concept, characteristics and defending mechanism of worms," *IEICE Transactions on Information and Systems*, vol. 92, no. 5, pp. 799-809, 2009.

[2] S. Qing and W. Wen, "A survey and trends on Internet worms," *Computers and Security*, vol. 24, no. 4, pp. 334-346, 2005.

[3] H. Berghel, "The code red worm," *Communications of the ACM*, vol. 44, no. 12, pp. 15-19, 2001.

[4] J. Cowie, A. Ogielski, B. Premore and Y. Yuan, "Global routing instabilities during code red ii and nimda worm propagation," *Renesys Corporation*, 2001.

[5] D. Moore, C. Shannon and K. Claffy, "Code-red: a case study on the spread and victims of an Internet worm," *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement*, pp. 273–284, 2002.

[6] F. Raynal, Y. Berthier, P. Biondi and D. Kaminsky, "Honeypot forensics part 1: analyzing the network," *IEEE Security and Privacy Magazine*, vol. 2, no. 4, pp. 72-78, 2004.

[7] D. Graham-Rowe, "Honeypot for hackers," *New Scientist*, vol. 171, no. 2303, pp. 15-15, 2001.

[8] N. Provos, "A virtual honeypot framework," *USENIX Security Symposium*, vol. 173, no. 2004, pp. 1-14, 2004.

[9] G. Streftaris and G. J. Gibson, "Statistical inference for stochastic epidemic models," *Proc. 17th International Workshop on Statistical Modeling*, pp. 609-616, 2002.

[10] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences*, vol. 115, no. 772, pp. 700–721, 1927.

[11] J. Satsuma, R. Willox, A. Ramani, B. Grammaticos and A. S. Carstea, "Extending the SIR epidemic model," *Physica A: Statistical Mechanics and its Applications*, vol. 336, no. 3, pp. 369-375, 2004.

[12] L. Acedo, G. González-Parra and A. J. Arenas, "Modal series solution for an epidemic model," *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 5, pp. 1151-1157, 2010.

[13] B. K. Mishra, and S. K. Pandey, "Fuzzy epidemic model for the transmission of worms in computer network," *Nonlinear Analysis: Real World Applications*, vol. 11, no. 5, pp. 4335–4341, 2010.

[14] Q. Liu, D. Jiang and T. Hayat, "Dynamics of a stochastic multigroup SIQR epidemic model with standard incidence rates," *Journal of the Franklin Institute*, vol. 356, no. 5, pp. 2960-2993, 2019.

[15] H. Hethcote, Z. Ma and S. Liao, "Effects of quarantine in six endemic models for infectious diseases," *Mathematical Biosciences*, vol. 180, no. 1, pp. 141-160, 2002.

[16] Y. Hua and G. Chen, "Network virus-epidemic model with the point-to-group information propagation," *Applied Mathematics and Computation*, vol. 206, no. 1, pp. 357-367, 2008.

[17] B. K. Mishra and S. K. Pandey, "Dynamic model of worms with vertical transmission in computer network," *Applied Mathematics and Computation*, vol. 217, no. 21, pp. 8438–8445, 2011.

[18] C. Wang and S. Chai, "Hopf bifurcation of an SEIRS epidemic model with delays and vertical transmission in the network," *Advances in Difference Equations*, vol. 2016, no. 1, pp. 1-19, 2016.

[19] B. K. Mishra and D. K. Saini, "SEIRS epidemic model with delay for transmission of malicious objects in computer network," *Applied Mathematics and Computation*, vol. 188, no. 2, pp. 1476–1482, 2007.

[20] O. A. Toutonji, S. M. Yoo and M. Park, "Stability analysis of VEISV propagation modeling for network worm attack," *Applied Mathematical Modelling*, vol. 36, no. 6, pp. 2751-2761, 2012.

[21] J. H. Guillén, A. M. del Rey and L. H. Encinas, "Study of the stability of a SEIRS model for computer worm propagation," *Physica A: Statistical Mechanics and its Applications*, vol. 479, pp. 411 – 421, 2017.

[22] G. Gebhart, "Worm Propagation and Countermeasures," *SANS Institute InfoSec Reading Room*, 2004.

[23] C. C. Zou, W. Gong and D. Towsley, "Code red worm propagation modeling and analysis," *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pp. 138-147, 2002.

[24] C. C. Zou, D. Towsley and W. Gong, "On the performance of Internet worm scanning strategies," *Performance Evaluation*, vol. 63, no. 7, pp. 700-723, 2006.

[25] J. G. Ren and Y. H. Xu, "a compartmental model to explore the interplay between virus epidemics and honeynet potency," *Applied Mathematical Modelling*, vol. 59, pp. 86-99, 2018.

[26] Q. Fu, Y. Yao, C. Sheng and W. Yang, "Interplay between malware epidemics and honeynet potency in industrial control system network," *IEEE Access*, vol. 8, pp. 81582-81593, 2020.

[27] M. G. Roberts and J. Heesterbeek, "Characterizing the next-generation matrix and basic reproduction number in ecological epidemiology," *Journal of Mathematical Biology*, vol. 66, no.4, pp. 1045-1064, 2013.

[28] C. Castillo-Chavez, Z. Feng and W. Huang, "On the computation of $\Re_0$ and its role on global stability," *Mathematical Approaches for Emerging and Re-emerging Infection Diseases: An Introduction*, vol. 125, pp. 31-65, 2002.

[29] J. P. LaSalle, "The stability of dynamical systems," *Society for Industrial and Applied Mathematics*, 1976.