Malware propagation model for cluster-based wireless sensor networks using epidemiological theory

Xuejin Zhu¹ and Jie Huang¹,²

¹ School of Cyber Science and Engineering, Southeast University, Nanjing, Jiangsu, China
² Purple Mountain Laboratories, Nanjing, Jiangsu, China

ABSTRACT

Due to limited resources, wireless sensor network (WSN) nodes generally possess weak defense capabilities and are often the target of malware attacks. Attackers can capture or infect specific sensor nodes and propagate malware to other sensor nodes in WSNs through node communication. This can eventually infect an entire network system and even cause paralysis. Based on epidemiological theory, the present study proposes a malware propagation model suitable for cluster-based WSNs to analyze the propagation dynamic of malware. The model focuses on the data-transmission characteristics between different nodes in a cluster-based network and considers the actual application parameters of WSNs, such as node communication radius, node distributed density, and node death rate. In addition, an attack and defense game between malware and defending systems is also established, and the infection and recovery rates of malware propagation under the mixed strategy Nash equilibrium condition are given. In particular, the basic reproductive number, equilibrium point, and stability of the model are derived. These studies revealed that a basic reproductive number of less than 1 leads to eventual disappearance of malware, which provides significant insight into the design of defense strategies against malware threats. Numerical experiments were conducted to validate the theory proposed, and the influence of WSN parameters on malware propagation was examined.

INTRODUCTION

Wireless sensor networks (WSNs) are multi-hop self-organizing network systems formed by large-scale sensor nodes communicating with each other deployed in the monitoring area. With the rapid development of the Internet of Things and sensor technology, WSNs are widely used in various civilian and military fields, such as in smart cities, environmental monitoring, and data collection (Ahmad et al., 2017; García et al., 2020; Khalifeh et al., 2021; Cai & Mirrabbasi, 2021; Lazarescu & Poolad, 2021). However, due to the limited energy and computing resources of sensor nodes, it is difficult to construct a complex security protection system (Farjamnia, Gasimov & Kazimov, 2019). To reduce costs, manufacturers of sensor-node devices could neglect security. At the same time,
sensor nodes are usually deployed in an open environment, which provides convenience for attackers (Liu et al., 2020). WSNs face many security threats, including denial-of-service (DoS) attacks, node capture, malware infection, and others (Souissi, Ben Azzouna & Ben Said, 2019). WSN security issues have attracted wide attention in academic circles and industrial circles.

Malware is a computer program that threatens or harms a network or system. The most common malware includes viruses, worms and Trojan horses, among others. In a WSN application, data are transmitted between adjacent nodes through wireless links, which creates favorable conditions for the propagation of malware. Malware may cause node failure, data leakage, DoS, and other failures, and can infect surrounding nodes through wireless transmission (Ojha et al., 2021). An attacker can target certain nodes, inject malware, and use the propagation mechanism to propagate the malware to the entire network, resulting in damage to the entire network system. However, the WSN system generally deploys protective measures against malware propagation that can detect data sent by system nodes and repair the nodes infected with the malware. However, implementation of these protective measures will predominately occupy limited communication channel resources and consume node power, which will increase data-transmission delays and shorten the lifecycle of network nodes. Therefore, it is necessary to choose an appropriate defense strategy to suppress the propagation of malware and minimize the loss of the entire system (Zhou, Shen & Liu, 2020). Therefore, it is vital to study the principles and mechanisms of malware propagation under attack and defense scenarios in WSNs.

In traditional Internet scenarios, epidemiological models have been widely used to study malware propagation, and a large number of computer malware propagation models have been proposed (Chen & Ji, 2005; Wang et al., 2014). However, because WSNs have the unique characteristics of limited node energy, node communication radius, and high node density, the mechanism of WSN Internet malware propagation clearly differs from traditional Internet scenarios. Therefore, the computer malware propagation model cannot be directly applied to malware propagation in WSNs.

In this study, the malware propagation model for cluster-based WSN network topology is evaluated based on game theory and epidemiology. Its primary contributions are as follows:

(1) An attack-defense game is proposed between malware and a WSN defend system. The mixed strategy Nash equilibrium of the model is obtained, and the infection rate of malware and the recovery rate of the system is calculated based on the Nash equilibrium solution of both parties in the game.

(2) A malware propagation mathematical model for cluster-based WSNs (based on the classic SIR model) is proposed under the attack-defense game. The model considers the data-transmission characteristics of cluster head nodes and common nodes, as well as the actual application scenarios and characteristics of WSNs, including node communication radius, node density, and node death.

(3) The basic reproductive number $R_0$ and equilibrium point of the model are deduced, and the stability of the equilibrium point proved. When $R_0 < 1$, WSN malware eventually disappears; otherwise, WSN malware will exist consistently.
Numerical simulations are proposed for the proposed model; the experimental results support the correctness and effectiveness of the proposed model and elucidate the relationship between malware propagation and WSN parameters.

The remainder of the paper is organized as follows. In ‘Related Work’, related work on malware propagation is discussed. In ‘Proposed Model’, a novel malware propagation model for cluster-based WSNs under an attack-defense game is introduced. In ‘Existence and Stability of Equilibrium’, the equilibrium points and stability of the model system are deduced. The simulation and numerical analysis results for the malware propagation model are presented in ‘Simulation and Numerical Analysis’. Conclusions are drawn in ‘Conclusion’.

RELATED WORK

Research toward exploring the malware propagation behavior of WSNs has resulted in several achievements. A robust survey has summarized malware propagation models in networks (Queiruga-Dios et al., 2017). In a study evaluating the propagation dynamics of worms in time and space in WSNs, Khayam & Radha (2005) considered the physical, MAC, and network layers of actual sensor networks according to the topology characteristics of WSNs, and proposed the topologically aware worm propagation model (TWPM). Shen et al. (2016) proposed a malware propagation model based on the epidemiology theory, and solved the problem of how to evaluate the reliability of sensor nodes in the case of malware propagation, so as to ensure efficient, continuous, and reliable transmission of sensory data from the node to the sink. Mishra & Keshri (2013) proposed an infectious disease model that includes a vaccination room. Their model not only reflects the temporal and spatial dynamics of the worm propagation process, but also performs mathematical analysis and numerical simulation on the worm propagation process. Nwokoye & Umeh (2018) developed the analytic-agent cyber dynamical systems analysis and design method (A²CDSADM) by combining the prevalent analytical and agent methods. This method can not only reflect the time dynamics of malware spreading, but can also observe the spatial dynamic changes of sensor nodes of different groups. Qiao et al. (2014) found that WSNs have the characteristics of small-world networks. They studied the epidemics in the small world of tree-based networks and calculated the epidemic threshold for epidemic outbreaks.

Many studies based on classic SIR epidemiology have also been undertaken. Wang & Li (2009) considered the energy of the sensor node to be limited; thus, based on the SIR model, the node death state was introduced to obtain a new model, i.e., iSIRS. The authors further proposed the EiSIRS virus propagation model (Wang, Li & Li, 2010), which describes the process of worm propagation in WSNs with sleep mechanisms. Their experimental results demonstrate that the remaining energy of nodes and the sleep scheduling mechanism are effective against worms. The propagation of viruses in WSNs has a certain impact. Liping et al. (2015) proposed an improved SIRS worm propagation model that considers the communication radius and distribution density of WSN nodes, determines the model’s equilibrium and basic reproductive number, and obtains large-scale worm propagation.
conditions. Zhu, Zhao & Wang (2015) proposed a SIRS malware propagation model with a state-feedback controller. Through the analysis of model stability and Hopf bifurcation, the state-feedback method was successful for unstable stable states and periodic oscillation control. Tang et al. proposed an SI model based on node dormancy maintenance that provides a repair function when an infected node enters the dormant state (Tang & Mark, 2009; Tang, 2011). The improved SI model can effectively prevent virus propagation in the network without adding any additional hardware workload or computational overhead.

The attack and defense parties of malware in WSNs can be regarded as having a game relationship. Therefore, game theory is also widely used in the security of WSNs, especially in malware-related fields. Abdalzaher et al. (2016) proposed a node protection model based on a Stackelberg game, which can be adapted to two different malicious node attack scenarios. In the first scenario, the attacker selects a group of nodes for which the protection degree is lower than a certain threshold to attack. In the second scenario, the attacker’s goal is to defend the weakest node in the previous round of attack. Wang, Li & Dong (2018) proposed an improved two-dimensional (2D) cellular automata model and a multi-role evolutionary game model to describe the process of malware propagation. Based on the existing 2D cellular automata malware model, the epidemiological propagation mechanism is improved, and the dynamic equation of strategy evolution is given. Shen et al. (2017) proposed a non-cooperative non-zero-sum game to describe the interaction between heterogeneous WSNs (HWSNs) system and malware. The game model can predict infection behavior of malware. Further, the author has established a node reliability evaluation mechanism in the state of malware propagation, which can efficiently evaluate system availability and reliability. Shen et al. (2014) proposed a differential game model for malware propagation in WSNs. In the process of the game between the system and the malware, the defense strategy can be changed dynamically, so that the total cost can be minimized. In addition, the author also considered the node sleep state in the process of propagation.

The current WSN malware propagation model is based on a flat network structure; in which all nodes in the network are equal and have completely consistent functional characteristics. However, the actual application scenarios of WSNs typically adopt a hierarchical network structure, and nodes are deployed in clusters. There are two types of nodes in WSNs, cluster head nodes and common nodes. This cluster-based network topology has many advantages over flat topology, such as easy expansion, convenient centralized management, low system construction cost, high network coverage, and reliability (Huang, Jie & Guizani, 2014). Because different network topologies adopt different data-transmission rules, the propagation mechanisms of malicious network software can vary substantially. Thus, previous studies cannot be applied to cluster-based hierarchical networks. In addition, existing epidemiological studies do not consider the strategies of infection rate and recovery rate, but only use a fixed parameter to express it, thus ignoring the impact of the attack and defense game process on the dynamics of malware propagation. In the actual propagation of WSN malware, both the cost and benefits of malware infection and system defense will be considered. For malware, if the expected benefit of launching an attack is greater than the cost, it will launch an infection attack.
Otherwise, its malicious intentions will be hidden. For system defense, the data is received by the node detected only when the system benefit of detection and repair is greater than the cost of detection. In response to these problems, in this study the propagation processes of malware in WSNs are analyzed, and a more effective formal model established to accurately determine the propagation dynamics of malware with cluster-based hierarchical network structure.

PROPOSED MODEL

The proposed identification method is presented in this section, but first the notations and their definitions that will be used in this paper are listed in Table 1. A cluster-based hierarchical WSN can be divided into cluster head nodes and common nodes according to the function of sensor nodes as shown in Fig. 1. Each cluster head node contains the same functional protocol, such as MAC address, routing, nodes management, or security protocol, while common sensor nodes usually do not have functions such as routing, management, and aggregation processing. The common nodes send collected data to the cluster head node. After data-fusion processing, the cluster head node transfers the data through a multi-hop routing and forwarding mechanism, and finally uploads it to the network base station. Therefore, common nodes can only communicate with the cluster head node of the cluster, but cluster head nodes can communicate with one another. It is assumed that sensor nodes are evenly distributed and deployed in a certain area, the communication radius of each sensor node is $r$, and the deployment density of the cluster head nodes is $\sigma$.

According to the classic SIR epidemiological model, the network nodes can have the following three states.

- **Susceptible ($S$):** The susceptible nodes are in a healthy state but can be easily infected by malware. Before the system is attacked by malware, the susceptible state is the initial state of all sensor nodes. In the proposed model, nodes in the susceptible state are divided into susceptible cluster head nodes and susceptible common nodes.

- **Infected ($I$):** The infected nodes are in a state of being contaminated by malware and can send multiple copies of the malware to nearby nodes by data interaction, thereby infecting the nodes in state $S$. Similar to the susceptible nodes in the proposed model, the nodes in state $I$ are divided into infected cluster head nodes and infected common nodes.

- **Recovered ($R$):** The recovered nodes are infected nodes that have been restored to a healthy state through security measures such as virus detection and vulnerability patching and will acquire immunity.

Game between malware and defend system for WSNs

Before studying the propagation mechanism of WSNs malware, one must first consider the attack and defense strategy between malware and a defend system. For sensor nodes infected by malware, it can propagate to more surrounding nodes, which can make the network invalid in a large area. However, if the malware propagates too frequently, it is
Table 1 Table of notations and definitions.

| Notation | Definition |
|----------|------------|
| $\mathcal{M}/D$ | Malware/defend system |
| $C_M/C_D$ | Malware/defend system’s strategy space |
| $U_M/U_D$ | Malware/defend system’s utility |
| $\tilde{S}(t)/S(t)$ | Number of susceptible cluster head nodes/common nodes at time $t$ |
| $\tilde{I}(t)/I(t)$ | Number of infected cluster head nodes/common nodes at time $t$ |
| $R(t)$ | Number of recovered nodes at time $t$ |
| $\tilde{S}_c/S_c$ | Number of effective contacts of cluster head nodes/common nodes |
| $N_1/N_2$ | Number of cluster head nodes/common nodes in cluster-based WSNs |
| $\beta$ | Infection rate of malware |
| $\gamma$ | Recovery rate of infected nodes |
| $r$ | Communication radius of sensor nodes |
| $b$ | Birth rate of sensor nodes |
| $\sigma$ | Density of cluster head nodes |
| $\varepsilon$ | Cost of attack by infected node |
| $\tau$ | Cost of system repairing infected node |
| $v$ | Cost of security detection by susceptible nodes on received data packets |
| $e$ | After infected node is restored, cost to malware or gain to system |
| $R_0$ | The basic reproductive number of malware propagation |
| $E_0$ | The malware-free equilibrium malware propagation |
| $E_1$ | The endemic equilibrium of malware propagation |

easy to be detected by susceptible nodes and repaired. At the same time, the energy of the infected node will be consumed every time the malware is propagated. Therefore, the problem for infected nodes is maximizing the propagation of malware while minimizing their own energy consumption. As for the defend system, the data received by the node can be detected for security. If abnormal malware packets are found, the data will be discarded, and the last hop node will be installed with patches and restored. However, these methods will inevitably bring interference to the normal operation of WSNs. For example, detecting data and installing security patches will consume energy and occupy bandwidth. Therefore, the defend system will also detect the received data with a certain probability, so as to maximize its own benefits. This shows that the attack and defense process between malware and defend systems is actually a game. Therefore, in this paper game theory is used to analyze the attack and defense strategies between them.

**Definition:** The strategic game between malware and a defend system in WSNs is composed of a ternary $G = (J, \{C_j\}, \{U_i\})$, where

- $J = \{\mathcal{M}, D\}$ is the set of players in the game, where $\mathcal{M}$ and $D$ represent the malware and the defend system, respectively.
• $C_i$ is the strategy set of player $i$ when $i$ is the defend system; $C_D = \{\text{detecting, no detecting}\}$, when $i$ is malware; $C_M = \{\text{propagation, no propagation}\}$.

• $U_i$ is the utility obtained by the player $i$ during the game, and its value is determined by the strategies adopted by both sides of the game.

According to the attack and defense game in Definition 1, the payoff matrix of the game can be obtained, as shown in Table 2. If the susceptible node performs security detection on the received data packet, it must consume energy and may block the channel, and the resulting system cost is recorded as $\nu$. The cost of malware caused by the energy consumption of the infected node sending the data packet containing the malicious software to the next hop node is recorded as $\epsilon$. After the susceptible node is detected, if malware is found, the system will repair the last hop node, which will consume energy and bandwidth, and set the total cost as $\tau$. After the repair is successful, the utility to the system and the cost to the malware are both $e$. When the malware initiates an attack and the system initiates detection, the overall utility of the system is $-\nu - \tau + e$, and the utility of the infected node is $-e - \epsilon$. When the system is not detected, the utility of the system and that of the malware are $-e$ and $e - \epsilon$, respectively. In addition, when the system initiates detection but no malware is found, the system’s utility will be $-\nu$. In other cases, the utility of both parties is zero.

The notion of mixed strategy Nash equilibrium (MNE) captures a steady state of a strategic game in which each player holds the correct expectation about other players’ behavior and acts rationally. The main feature of MNE is that the opponent adopts any pure strategy, and the player’s expected utility is the same. According to game theory, there
must be MNE in the game of limited players, so one can solve the MNE of the game \( \mathcal{G} \). Assume that malware attacks and propagates with a probability of \( p \), and does not attack with a probability of \( 1 - p \). The system detects the received data with probability \( q \), and does not detect with probability \( 1 - q \). The utility maximization method to can be used to solve the MNE. According to the utility matrix, the expected utility of the system \( E(U_D) \) is

\[
E(U_D) = pq(-v - \tau + e) + q(1 - p)(-v) + p(1 - q)(-e)
\]

Letting \( \frac{\partial E(U_D)}{\partial q} = 0 \), one has

\[
p = \frac{v}{2e - \tau}
\]  

Similarly, the expected utility of malware \( E(U_M) \) is

\[
E(U_M) = pq(-e - \varepsilon) + p(1 - q)(e - \varepsilon)
\]

Letting \( \frac{\partial E(U_M)}{\partial p} = 0 \), one has

\[
q = \frac{e - \varepsilon}{2e}
\]

When malware initiates an attack with probability \( \frac{v}{2e - \tau} \), the expected utility of the malware is the same regardless of whether the system is detected or not. Similarly, when the probability of system detection is \( \frac{e - \varepsilon}{2e} \), the expected utility is the same regardless of whether the malware initiates an attack. Therefore, \( (p^* = \frac{v}{2e - \tau}, q^* = \frac{e - \varepsilon}{2e}) \) is the MNE of the attack and defense game between malware and system. When both parties are rational, they will use this strategy to attack and defend. When the malware initiates an attack and the system does not detect it, the malware will spread successfully; when the malware initiates an attack and the system detects it, the system will find the malware and repair the data source node. Therefore, the malware infection rate and system recovery rate under MNE conditions can be obtained:

\[
\beta = p^*(1 - q^*) = \frac{v(e + \varepsilon)}{2e(2e - \tau)}
\]

\[
\gamma = p^* q^* = \frac{v(e - \varepsilon)}{2e(2e - \tau)}
\]
Malware propagation model

In this subsection, a malware propagation model is established under the condition of MNE. Let \( \tilde{S}(t), \tilde{I}(t), S(t), I(t), \) and \( R(t) \) denote the susceptible cluster head nodes, infected cluster head nodes, susceptible common nodes, infected common nodes, and number of recovered nodes at time \( t \), respectively. According to the malware attack strategy under MNE, the infected nodes infect surrounding susceptible nodes with infection rate \( \beta \) that can communicate with each other. Moreover, according to the system defense strategy, the infected nodes recover at a rate of \( \gamma \).

In a cluster-based hierarchical network, it is assumed that the number of cluster head nodes and common nodes are \( N_1 \) and \( N_2 \), respectively; that is, there are \( N_1 \) clusters in the network, and each cluster contains \( N_2/N_1 \) common nodes with the node status conversions. Letting \( \tilde{S}_r(t) \) and \( S_r(t) \) represent the effective contact number of the infected cluster head node against the susceptible cluster head node and the susceptible common node, respectively. The formula is expressed as follows:

\[
\tilde{S}_r(t) = \frac{\sigma \pi r^2}{N_1} \tilde{S}(t) \tag{7}
\]

\[
S_r(t) = \frac{S(t)}{N_1} \tag{8}
\]

Each infected node can infect \( \beta \tilde{S}_r \) susceptible nodes per unit time. Therefore, the conversion rate of susceptible cluster head nodes to infected cluster head nodes is \( \beta \tilde{S}_r(t) \tilde{I}(t) \) at time \( t \). Since common nodes cannot communicate with each other but can only be infected by cluster head nodes, the infection rate of common nodes is also determined by \( \tilde{I}(t) \). The conversion rate from susceptible common node to infected common node is \( \beta S_r(t) \tilde{I}(t) \) at time \( t \). At the same time, due to the existence of the defense system, the conversion rate of infected nodes \( I \) and \( \tilde{I} \) to immune group \( R \) is \( \gamma I(t) \) and \( \gamma \tilde{I}(t) \), respectively. Considering that the node cannot continue to work due to physical device damage or exhaustion of battery power, the death and births rate of the node are both set to \( b \), which ensures that the number of nodes in the network remains constant. In this way, we can obtain the state transition relationship between groups in the cluster network shown in Fig. 2.

In accordance with the rates of change between different states shown in Fig. 2, one can establish a mathematical model of a system of differential equations based on cluster-based WSNs for malware propagation:

\[
\begin{align*}
\frac{d\tilde{S}(t)}{dt} &= bN_1 - \beta \frac{\sigma \pi r^2}{N_1} \tilde{S}(t) \tilde{I}(t) - b\tilde{S}(t) \\
\frac{d\tilde{I}(t)}{dt} &= \beta \frac{\sigma \pi r^2}{N_1} \tilde{S}(t) \tilde{I}(t) - \gamma \tilde{I}(t) - b\tilde{I}(t) \\
\frac{dS(t)}{dt} &= bN_2 - \beta \frac{S(t)}{N_1} I(t) - bS(t) \\
\frac{dI(t)}{dt} &= \beta \frac{S(t)}{N_1} I(t) - \gamma I(t) - bI(t) \\
\frac{dR(t)}{dt} &= \gamma \tilde{I}(t) + \gamma I(t) - bR(t)
\end{align*}
\tag{9}
\]
Since the total number of sensor nodes in the system is a fixed value \((N_1 + N_2)\), one can obtain
\[
R(t) = N_1 + N_2 - \tilde{S}(t) - S(t) - \tilde{I}(t) - I(t)
\]  
(10)

Therefore, primarily the first four equations in system (9) are considered. Letting
\[
\alpha_1 = \beta \frac{\sigma \pi r^2}{N_1}
\]  
(11)
\[
\alpha_2 = \frac{\beta}{N_1}
\]  
(12)

The system model Eq. (9) can finally be simplified as
\[
\begin{align*}
\frac{d\tilde{S}}{dt} &= bN_1 - \alpha_1 \tilde{S} \tilde{I} - b\tilde{S} \\
\frac{d\tilde{I}}{dt} &= \alpha_1 \tilde{S} \tilde{I} - \gamma \tilde{I} - b\tilde{I} \\
\frac{dS}{dt} &= bN_2 - \alpha_2 S \tilde{I} - bS \\
\frac{dI}{dt} &= \beta \alpha_2 \tilde{S} \tilde{I} - \gamma I - bI
\end{align*}
\]  
(13)

**EXISTENCE AND STABILITY OF EQUILIBRIUM**

In this section equilibrium points of system Eq. (13) are derived and their stability proved. If an equilibrium point is globally stable, then the final state of the system is certain under any initial conditions. For equilibrium points of system Eq. (13), one has
\[
\frac{d\tilde{S}}{dt} = 0, \frac{d\tilde{I}}{dt} = 0, \frac{dS}{dt} = 0, \frac{dI}{dt} = 0
\]  
(14)
Letting
\[ \alpha_1 \tilde{S} - \gamma \tilde{I} - b \tilde{I} = 0 \] (15)
one has (i) \( \tilde{I} = 0 \) or (ii) \( \tilde{S} = (\gamma + b) / \alpha_1 \) and \( \tilde{I} > 0 \). For the case of \( \tilde{I} = 0 \), one has malware-free equilibrium
\[ E_0 = (\tilde{S}_0, \tilde{I}_0, S_0, I_0) = (N_1, 0, N_2, 0) \] (16)
For the case of \( \tilde{S} = (\gamma + b) / \alpha_1 \) and \( \tilde{I} > 0 \), one has endemic equilibrium:
\[ E_1 = (\tilde{S}_1, \tilde{I}_1, S_1, I_1) \] (17)
where
\[ \tilde{S}_1 = \frac{\gamma + b}{\alpha_1} \] (18)
\[ \tilde{I}_1 = \frac{bN_1}{\gamma + b} - \frac{b}{\alpha_1} \] (19)
\[ S_1 = \frac{bN_2}{\alpha_2 \tilde{I}_1 + b} \] (20)
\[ I_1 = \frac{bN_2}{(\gamma + b)(1 + \frac{b}{\alpha_2 \tilde{I}_1})} \] (21)
Letting
\[ R_0 = \frac{N_1 \alpha_1}{\gamma + b} = \frac{\beta \sigma r^2}{N_1 \alpha_1} \] (22)
\( R_0 \) is the basic reproductive number of system Eq. (13), and only if \( R_0 > 1 \); thus, \( \tilde{I}_1 > 0 \), and the endemic equilibrium \( E_1 \) is meaningful.

**Malware-free equilibrium and its stability**

To investigate the local stability of the equilibrium points of the system Eq. (13), one must calculate the corresponding Jacobian matrix as
\[ J = \begin{bmatrix}
-\alpha_1 \tilde{I} - b & -\alpha_1 \tilde{S} & 0 & 0 \\
\alpha_1 \tilde{I} & \alpha_1 \tilde{S} - \gamma - b & 0 & 0 \\
0 & -\alpha_2 S & -\alpha_2 \tilde{I} - b & 0 \\
0 & \alpha_2 S & \alpha_2 \tilde{I} & -\gamma - b
\end{bmatrix} \] (23)

**Lemma 1.** The malware-free equilibrium \( E_0 \) of system Eq. (13) is locally asymptotically stable if \( R_0 < 1 \) and unstable if \( R_0 > 1 \).

**Proof.** Calculating the Jacobian matrix Eq. (25) at malware-free equilibrium, one obtains
\[ J(E_0) = \begin{bmatrix}
-b & -\alpha_1 N_1 & 0 & 0 \\
\alpha_1 \tilde{I} & \alpha_1 N_1 - \gamma - b & 0 & 0 \\
0 & -\alpha_2 N_2 & -b & 0 \\
0 & \alpha_2 N_2 & 0 & -\gamma - b
\end{bmatrix} \] (24)
To prove the stability at the point $E_0$, one will find all the eigenvalues ($\lambda$) of the matrix Eq. (26). The eigenvalues are given as
\[ \lambda = -b, \quad -\gamma - b, \quad \alpha_1 N_1 - \gamma - b \] (25)
where $-b$ is a double eigenvalue.

Hence, when $R_0 < 1$, all eigenvalues of the matrix (26) have no positive real part, and the malware-free equilibrium $E_0$ is locally asymptotically stable; when $R_0 > 1$, Eq. (26) has a positive eigenvalue $\alpha_1 N_1 - \gamma - b$. Thus, malware-free equilibrium $E_0$ is unstable. **Theorem 2.** The malware-free equilibrium $E_0$ is globally asymptotically stable if $R_0 < 1$.

**Proof.** Let
\[ D = \{ (\tilde{S}, \tilde{T}, S, I) \in \mathbb{R}^4 \mid 0 \leq \tilde{S} + \tilde{T} \leq N_1, \quad 0 \leq S + I \leq N_2, \quad \tilde{S}, \tilde{T}, S, I \geq 0 \} \] (26)
Obviously, $D$ is the positive invariant set of system Eq. (13). When $R_0 < 1$, construct the Liapunov function
\[ V(t) = \tilde{T}(t) \] (27)
The derivative of $V(t)$ along the trajectory of system Eq. (13) is\[ \frac{dV(t)}{dt} = \frac{d\tilde{T}(t)}{dt} = \left[ \alpha_1 \tilde{S} - (b + \gamma) \right] \tilde{T} \leq \left[ \alpha_1 N_1 - (b + \gamma) \right] \tilde{T} \] (28)
Since $R_0 < 1$, then $\alpha_1 N_1 - (b + \gamma) < 0$. Thus,
\[ Q = \left\{ (\tilde{S}, \tilde{T}, S, I) \in D \mid \frac{dV(t)}{dt} = 0 \right\} = \{ \tilde{T} = 0 \} \] (29)
Therefore, the maximum invariant set of the system in $Q$ is $\{ \tilde{T} = 0 \}$. According to the principle of Lasalle invariance,
\[ \lim_{t \to \infty} \tilde{T}(t) = 0 \] (30)
Substituting the preceding equation into system Eq. (13).
\[ \lim_{t \to \infty} \tilde{S}(t) = N_1, \quad \lim_{t \to \infty} S(t) = N_2, \quad \lim_{t \to \infty} I(t) = 0 \] (31)
Therefore, $E_0$ is globally attractive. Combined with the local stability of $E_0$, it can be seen that $E_0$ is globally asymptotically stable.

**Endemic equilibrium and its stability**

**Lemma 3.** The endemic equilibrium $E_1$ of system Eq. (13) is locally asymptotically stable if $R_0 > 1$.

**Proof.** At the endemic equilibrium point, the Jacobian matrix is
\[ J(E_1) = \begin{bmatrix} -\alpha_1 \tilde{L}_1 - b & -\gamma - b & 0 & 0 \\ \alpha_1 \tilde{L}_1 & 0 & 0 & 0 \\ 0 & -\frac{b N_1}{\tilde{L}_1 + \frac{b}{\alpha_2}} & -\alpha_2 \tilde{L}_1 - b & 0 \\ 0 & \frac{b N_1}{\tilde{L}_1 + \frac{b}{\alpha_2}} & \alpha_2 \tilde{L}_1 & -\gamma - b \end{bmatrix} \] (32)
Two eigenvalues are given as
\[ \lambda = -\alpha_2 \tilde{I}_1 - b, \quad -\gamma - b \] (33)
and are negative. The remaining two eigenvalues are given by
\[ \lambda^2 + (\alpha_1 \tilde{I}_1 + b) \lambda + (\gamma + b) \alpha_1 \tilde{I}_1 = 0 \] (34)
According to the Routh–Hurwitz criteria, since all the coefficients of Eq. (36) are positive, there is no positive real part eigenvalue. When \( R_0 > 1 \), all eigenvalues of the matrix Eq. (34) have no positive real part, and endemic equilibrium \( E_1 \) is locally asymptotically stable.

**Theorem 4.** The endemic equilibrium \( E_1 \) is globally asymptotically stable if \( R_0 > 1 \).

**Proof.** When \( R_0 > 1 \), construct the Liapunov function
\[ V(t) = \frac{1}{2} \omega_1 \left( \tilde{S} - \tilde{S}_1 \right)^2 + \omega_2 \left( \tilde{I} - \tilde{I}_1 - \tilde{I}_1 \ln \frac{\tilde{I}}{\tilde{I}_1} \right) \] (35)
where \( \omega_i > 0, i = 1, 2 \). The derivative of \( V(t) \) along the trajectory is
\[
\frac{dV(t)}{dt} = \omega_1 \left( \tilde{S} - \tilde{S}_1 \right) \frac{d\tilde{S}(t)}{dt} + \omega_2 \left( 1 - \frac{\tilde{I}_1}{\tilde{I}} \right) \frac{d\tilde{I}(t)}{dt}
\]
\[
= \omega_1 \left( \tilde{S} - \tilde{S}_1 \right) \left( bN_1 - \alpha_1 \tilde{I} - b\tilde{S} \right) + \omega_2 \left( \tilde{I} - \tilde{I}_1 \right) \left( \alpha_1 \tilde{S} - \alpha_1 \tilde{S}_1 \right)
\]
\[
= -\omega_1 \left( \tilde{S} - \tilde{S}_1 \right)^2 \left( \alpha_1 \tilde{I} + b \right) + \omega_1 \left( \omega_2 - \omega_1 \tilde{S}_1 \right) \left( \tilde{I} - \tilde{I}_1 \right) \left( \tilde{S} - \tilde{S}_1 \right)
\] (36)
Letting \( \omega_2 = \omega_1 \tilde{S}_1 \), and any value \( \omega_1 > 0 \), one thus obtains
\[
\frac{dV(t)}{dt} = -\omega_1 \left( \tilde{S} - \tilde{S}_1 \right)^2 \left( \alpha_1 \tilde{I} + b \right) \leq 0
\] (37)
and, obviously,
\[ Q = \left\{ (\tilde{S}, \tilde{I}, S, I) \in D \left| \frac{dV(t)}{dt} = 0 \right. \right\} = \{ \tilde{S} = \tilde{S}_1 \} \] (38)
where \( D \) is determined by Eq. (28). Therefore, the maximum invariant set of the system in \( Q \) is \{ \( \tilde{S} = \tilde{S}_1 \) \}. According to the principle of Lasalle invariance,
\[ \lim_{t \to \infty} \tilde{S}(t) = \tilde{S}_1 \] (39)
Substituting this into system Eq. (13), one has
\[
\lim_{t \to \infty} \tilde{I}(t) = \tilde{I}_1, \quad \lim_{t \to \infty} S(t) = S_1, \quad \lim_{t \to \infty} I(t) = I_1
\] (40)
Therefore, \( E_1 \) is globally attractive. Combined with the local stability of \( E_1 \), it can be seen that \( E_1 \) is globally asymptotically stable.

**SIMULATION AND NUMERICAL ANALYSIS**

According to Theorem 2 and Theorem 4, the size of basic reproductive number \( R_0 \) is of great significance in determining whether WSN malware will continue to propagate. When \( R_0 < 1 \), the system reaches global stability at the malware-free equilibrium point regardless of the initial state of each group in the network, and the malware will eventually disappear.
Figure 3  Dynamics showing that number of nodes in different states changes with time. (A) $R_0 < 1$ ($\beta = 0.0061$, $\gamma = 0.0044$), the number of infected nodes is eventually reaches zero. (B) $R_0 > 1$ ($\beta = 0.3$, $\gamma = 0.0333$), the number of infected nodes is eventually reaching a constant value at the disease equilibrium point $E_1$.

However, when $R_0 > 1$, the system reaches global stability at the epidemic equilibrium point, and the number of infected nodes in the final system will be maintained at a normal state. Therefore, one can manipulate the size of the basic reproductive number $R_0$ by changing network parameters, for example node density, communication radius, and death rate, to determine the effects of these parameters on malware propagation. To verify the results, a numerical simulation experiment was conducted using the system dynamics modeling software Vensim.

It is assumed that the number of cluster head nodes $N_1 = 100$, and the number of common nodes $N_2 = 1000$ in a WSN, and a set of simulation parameters was chosen as follows: $b = 0.01$, $\gamma = 0.01$, $\sigma = 0.5$, $r = 1$, $\epsilon = 8$, $\nu = 1$, $\tau = 5$, $e = 50$, and $\beta = 0.01$. Initial values of susceptible, infected, and recovered nodes in WSNs are $\tilde{S}(0) = 75$, $\tilde{I}(0) = 25$, $S(0) = 1000$, $I(0) = 0$, and $R(0) = 0$. According to Eqs. (5) and (6), one can obtain the infection and recovery rates as 0.0061 and 0.0044 under the game of the malware and the defend system, respectively, and further calculate that $R_0 < 1$.

The dynamics that the number of nodes in different states changes with time is shown in Fig. 3. Based on Fig. 3A, throughout the propagation of WSN malware, the number of infected cluster head nodes is monotonously decreasing, and eventually reaches zero; the number of infected common nodes increases rapidly in the initial stage, but as the number of cluster head infected nodes decreases, the number of common nodes infected gradually decreases, and eventually reaches zero. The other three state groups eventually reach a stable level. Therefore, the system state eventually reaches the malware-free equilibrium point $E_0$, which is consistent with Theorem 2.

To further verify the propagation dynamics when $R_0 > 1$, the loss detected by the system is reduced to $\nu = 5$ and the revenue of node recovery increased to $e = 10$. At this time, the infection and recovery rates will increase to 0.3 and 0.0333, respectively. In this situation, one can obtain $R_0 > 1$. The numerical simulation results are shown in Fig. 3B. It is found that the number of infected nodes increases rapidly in the initial stage, and then begins...
Figure 4  Dynamics of infected nodes for different communication radii $r$. (A) Infected cluster head nodes $I(t)$, (B) Infected common nodes $I(t)$. When $r < r_{th}$, the malware in the system will eventually disappear; when $r > r_{th}$, the number of infected nodes eventually tends to a constant value.

Full-size DOI: 10.7717/peerjcs.728/fig-4

to decline before eventually reaching a constant value. Malware continues to propagate among network nodes, which shows the influence of the infection and recovery rates on malware propagation in WSNs. This is also in line with the conclusion obtained in Theorem 4, namely that the system is ultimately at the disease equilibrium point $E_1$. Next, the effects of sensor-node density, node communication radius, and node death rate on WSN malware infection are studied.

**Node communication radius $r$**

Letting $R_0 = 1$ for Eq. (24), one can obtain the node communication radius threshold of the malware propagation in a cluster-based WSN:

$$
r_{th} = \sqrt{\frac{\gamma + b}{\beta \pi}}
$$

That is to say, when $r < r_{th}$, $R_0 < 1$, according to Theorem 2, system Eq. (13) will stabilize at the malware-free equilibrium $E_0$, and the malware in the system will eventually disappear. When $r > r_{th}$, $R_0 > 1$, and, according to Theorem 4, system Eq. (13) will stabilize at the endemic equilibrium $E_1$, and malware in WSNs will exist consistently. According to the aforementioned WSN parameters, one can calculate $r_{th} = 1.2259$. As shown in Fig. 4, when $r = 0.5 < r_{th}$ and $r = 1 < r_{th}$, system Eq. (13) stabilizes at malware-free equilibrium. The number of infected nodes eventually reaches zero, and convergence speed increases as $r$ decreases. When $r = 2 > r_{th}$ and $r = 3 > r_{th}$, system Eq. (13) stabilizes at the endemic equilibrium. The number of infected nodes eventually tends to a constant value, and the constant value increases with $r$.

**Node distributed density $\sigma$**

One can also obtain the node distributed density threshold of malware propagation in a cluster-based WSN according to Eq. (24):

$$
\sigma_{th} = \frac{\gamma + b}{\beta \pi r^2}
$$
That is to say, when $\sigma < \sigma_{th}$, $R_0 < 1$, and the malware in the system will eventually disappear. When $\sigma > \sigma_{th}$, $R_0 > 1$, and system Eq. (13) will stabilize at the endemic equilibrium $E_1$, and malware will exist consistently. By calculation, $\sigma = 0.7514$. As shown in Fig. 5, when $\sigma = 0.1 < \sigma_{th}$ and $\sigma = 0.5 < \sigma_{th}$, the malware will eventually disappear, and the convergence speed increases as $\sigma$ decreases. When $\sigma = 1 > \sigma_{th}$ and $\sigma = 5 > \sigma_{th}$, the number of infected nodes eventually tends to a constant value, and the constant value increases with $\sigma$.

**Node death rate $b$**

The threshold of malware propagation about the death rate of nodes is

$$b_{th} = \beta \sigma \pi r^2 - \gamma$$

(43)

That is to say, when $b > b_{th}$, $R_0 < 1$, the malware in the system will eventually disappear. When $b < b_{th}$, $R_0 > 1$, system Eq. (13) will stabilize at the endemic equilibrium $E_1$, and malware will exist consistently. By calculation, $b_{th} = 0.0052$. As shown in Fig. 6, when $b = 0.001 < b_{th}$ and $b = 0.003 < b_{th}$, system Eq. (13) stabilizes at endemic equilibrium. When $b = 0.01 > b_{th}$ and $b = 0.03 > b_{th}$, system Eq. (13) stabilizes at malware-free equilibrium. In contrast with communication radius $r$ and node density $\sigma$, even when $R_0$ is greater than 1, the number of infected cluster head nodes decreases monotonically and eventually tends to a constant value. In Figs. 4A and 5A, it can be seen that when $R_0 > 1$ the number of infected cluster head nodes increases rapidly in the initial stage and then decreases to a constant level.

**CONCLUSION**

In this paper, a differential equation model is proposed to analyze the propagation dynamic of malware based on epidemiology and game theory for cluster-based WSNs. The game between malware and the WSN defend system is established, and the model’s mixed strategy Nash equilibrium obtained. Different from the malware infection rate and recovery rate assumed in other existing propagation models, one can calculate the specific
parameter expression of the infection and recovery rates according to the Nash equilibrium. When infectious disease theory is used to build a dynamic model of transmission, the communication methods of cluster head nodes and common nodes in a cluster-based network is considered and the malware propagation characteristics determined. The equilibrium point of the model is derived and the stability of the equilibrium point analyzed to determine conditions for avoiding the continuous propagation of malware in WSNs. The theoretical analysis and numerical simulations performed in this paper show that the propagation dynamics of malware in WSNs is closely related to node communication radius, node density, and node death rate. To effectively prevent and control the propagation of malicious software, cluster-based WSNs should set reasonable network parameters so that the basic reproductive number of malicious software propagation is less than 1, thereby improving WSN defense capabilities. In future work, we will further study the malware detection and defense issues in the WSN system, so that the system can detect malware in time and prevent damage.

**ADDITIONAL INFORMATION AND DECLARATIONS**

**Funding**
This work was supported by the National Key Research and Development Program of China (Grant No. 2018YFB2100403). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Grant Disclosures**
The following grant information was disclosed by the authors: National Key Research and Development Program of China: 2018YFB2100403.

**Competing Interests**
The authors declare there are no competing interests.
Author Contributions

- Xuejin Zhu conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Jie Huang conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The Vensim numerical simulation file of the malware propagation model is available in the Supplementary File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.728#supplemental-information.

REFERENCES

Abdalzaher MS, Seddik K, Muta O, Abdelrahman A. 2016. Using Stackelberg game to enhance node protection in WSNs. In: IEEE consumer communications and networking conference. Piscataway: IEEE.

Ahmad A, Yu M, Sagnik C, Saima I. 2017. A comprehensive survey on real-time applications of WSN. Future Internet 9(4):77 DOI 10.3390/fi9040077.

Cai M, Mirrabbasi S. 2021. A self-sustained smart monitoring platform for capacitive de-ionization cell in wireless sensor network. IEEE Transactions on Industrial Electronics 68(5):4164–4172 DOI 10.1109/TIE.2020.2982104.

Chen Z, Ji C. 2005. Spatial-temporal modeling of malware propagation in networks. IEEE Transactions on Neural Networks 16(5):1291 DOI 10.1109/TNN.2005.853425.

Farjamnia G, Gasimov Y, Kazimov C. 2019. Review of the techniques against the wormhole attacks on wireless sensor networks. Wireless Personal Communications 105(4):1561–1584 DOI 10.1007/s11277-019-06160-0.

García L, Parra L, Jimenez JM, Lloret J, Lorenz P. 2020. IoT-based smart irrigation systems: an overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture. Sensors 20(4).

Huang L, Jie L, Guizani M. 2014. Secure and efficient data transmission for cluster-based wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems 25(3):750–761 DOI 10.1109/TPDS.2013.43.

Khalifeh A, Darabkh KA, Khasawneh AM, Alqaisieh I, Rajendiran K. 2021. Wireless sensor networks for smart cities: network design, implementation and performance evaluation. Electronics 10(2):218 DOI 10.3390/electronics10020218.

Khayam SA, Radha H. 2005. A topologically-aware worm propagation model for wireless sensor networks. In: IEEE international conference on distributed computing systems workshops. Piscataway: IEEE.
Lazarescu MT, Poolad P. 2021. Asynchronous resilient wireless sensor network for train integrity monitoring. *IEEE Internet of Things Journal* **8**(5):3939–3954 DOI 10.1109/JIOT.2020.3026243.

Liping F, Lipeng S, Qingshan Z, Hongbin W. 2015. Modeling and stability analysis of worm propagation in wireless sensor network. *Mathematical Problems in Engineering* **2015**:1–8.

Liu X, Yu J, Li F, Lv W, Wang Y, Cheng X. 2020. Data aggregation in wireless sensor networks: from the perspective of security. *IEEE Internet of Things Journal* **7**(7):6495–6513 DOI 10.1109/JIOT.2019.2957396.

Mishra BK, Keshri N. 2013. Mathematical model on the transmission of worms in wireless sensor network. *Applied Mathematical Modelling* **37**(6):4103–4111 DOI 10.1016/j.apm.2012.09.025.

Nwokoye CN, Umeh I. 2018. Analytic-agent cyber dynamical systems analysis and design method for modeling spatio-temporal factors of malware propagation in wireless sensor networks. *Methods* **5**:1373–1398 DOI 10.1016/j.mex.2018.10.005.

Ojha RP, Srivastava PK, Sanyal G, Gupta N. 2021. Improved model for the stability analysis of wireless sensor network against malware attacks. *Wireless Personal Communications* **116**:2525–2548 DOI 10.1007/s11277-020-07809-x.

Qiao LI, Zhang B, Cui L, Fan Z, Vasilakos AV. 2014. Epidemics on small worlds of tree-based wireless sensor networks. *Journal of Systems Science and Complexity* **27**(6):1095–1120 DOI 10.1007/s11424-014-1178-1.

Queiruga-Dios A, Encinas AH, Martín-Vaquero J, Encinas LH. 2017. Malware propagation models in wireless sensor networks: a review. In: *International conference on european transnational education international workshop on soft computing models in industrial and environmental applications computational intelligence in security for information systems conference*.

Shen S, Huang I, Liu J, Adam C., Shui Y, Cao Q. 2016. Reliability evaluation for clustered WSNs under malware propagation. *Sensors* **16**(6):855.

Shen S, Li H, Han R, Vasilakos AV, Wang Y, Cao Q. 2014. Differential game-based strategies for preventing malware propagation in wireless sensor networks. *Information Forensics and Security IEEE Transactions on* **9**(11):1962–1973 DOI 10.1109/TIFS.2014.2359333.

Shen S, Ma H, Fan E, Hu K, Yu S, Liu J, Cao Q. 2017. A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion. *Journal of Network and Computer Applications* **91**(Aug.):26–35 DOI 10.1016/j.jnca.2017.05.003.

Souissi I, Ben Azzouna N, Ben Said L. 2019. A multi-level study of information trust models in WSN-assisted IoT. *Computer Networks* **151**(MAR. 14):12–30 DOI 10.1016/j.comnet.2019.01.010.

Tang S. 2011. A modified SI epidemic model for combating virus spread in wireless sensor networks. *International Journal of Wireless Information Networks* **18**(4):319–326 DOI 10.1007/s10776-011-0147-z.
Tang S, Mark BL. 2009. Analysis of virus spread in wireless sensor networks: an epidemic model. In: 2009 7th international workshop on design of reliable communication networks. Piscataway: IEEE, 86–91 DOI 10.1109/DRCN.2009.5340022.

Wang X, Li Q, Li Y. 2010. EiSIRS: a formal model to analyze the dynamics of worm propagation in wireless sensor networks. Journal of Combinatorial Optimization 20(1):47–62 DOI 10.1007/s10878-008-9190-9.

Wang XM, Li YS. 2009. An improved SIR model for analyzing the dynamics of worm propagation in wireless sensor networks. Chinese Journal of Electronics 18(1):8–12.

Wang Y, Li D, Dong N. 2018. Cellular automata malware propagation model for WSN based on multi-player evolutionary game. Iet Networks 7(3):129–135 DOI 10.1049/iet-net.2017.0070.

Wang Y, Wen S, Xiang Y, Zhou W. 2014. Modeling the propagation of worms in networks: a survey. IEEE Communications Surveys and Tutorials 16(2):942–960 DOI 10.1109/SURV.2013.100913.00195.

Zhou H, Shen S, Liu J. 2020. Malware propagation model in wireless sensor networks under attack-defense confrontation. Computer Communications 162:51–58 DOI 10.1016/j.comcom.2020.08.009.

Zhu L, Zhao H, Wang X. 2015. Bifurcation analysis of a delay reaction-diffusion malware propagation model with feedback control. Communications in Nonlinear Science and Numerical Simulation 22(1-3):747–768 DOI 10.1016/j.cnsns.2014.08.027.