Study of Virus Propagation Model in Cloud Environment

Tongrang Fan, Yanjing Li and Feng Gao

1School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China
Fantr2009@126.com, 532465444@qq.com, f.gao@live.com

Abstract

The influence of the changes about the number of the nodes with different states by the external factors is not considered in the traditional model. So, traditional models are not suitable for the cloud network which is dynamic. A virus propagation model of computer, HSIR propagation model, which is suitable for the cloud environment is proposed in this paper. The model considers the dynamic changes of the cloud environment, the implementation of the immunization strategy and the time difference between the virus spread and the implementation of the immunization strategy. Additionally, the balance point of the HSIR model is found and the stability of the balance point is proved by the mathematical theory. The HSIR virus propagation model can describe the process of the virus propagation in the cloud environment more realistic.

Keywords: cloud network, HSIR, balance point, stability

1. Introduction

Computer viruses have many similarities with the biological viruses, so the propagation model of the computer is established by using the propagation model of the biology [1-4]. The SIS model which describes the virus propagation of the computer was proposed by the Kephart and White for the first time. The nodes in the SIS model have two states, the susceptible nodes and infected nodes. The infected nodes can change into the healthy nodes by the antivirus software in the computer network. So, the SIR model was proposed by J Kim [5]. The nodes in the SIR model have three states, susceptible nodes, infected nodes and recovered nodes.

The SEIRS model is proposed by BK Mishra and DK Saini [6]. The model describes the propagation of the worm, studies its disease-free equilibrium and explains the stability of the simulation results based on the threshold parameters. BK Mishra and Navnit Jha [7] proposes the dynamic model of the virus propagation, and gives the estimate of the time evolution of the infected nodes. Then they discuss the behaviors of the nodes with different states by using the SEIQR model. At the same time, the threshold, equilibrium and the stability of the virus propagation in the computer network are studied by them. BK Mishra and SK Pandey [8, 9] analysis the fuzzy transmission of the worms in the computer network and propose the SIRS model of fuzzy transmission. The impact of the virus for the immune response of the computer network is studied by paper [10] by using the nonlinear mathematical model. They find that the immune system response changes with the concentration of the virus. The characteristics of the transmission dynamic of computer virus are proposed by Z Dezso [11], Newman and Forrest [12]. The behaviors of the transmission dynamic of network are analyzed by using the SIR and SIS model by Grassberger [13]. J. Kleinberg [14] point out that the propagation of the network virus need to consider the features of data communication.
of nodes. The impact of the connection mode in the computer network for the propagation of the computer virus is studied in paper [15-19].

The impacts of the changes about the number of the nodes with different states by the external factors are not considered in the above models. So, they are all passive propagation model. Cloud environment which is a typical dynamic environment is accessed and quit by lots of nodes at any time. The number of nodes in the cloud environment is changing all the time. Therefore the above models are all not suitable for the cloud network. A virus propagation model of computer which applies to the cloud environment is proposed in this paper. The dynamic changes of the cloud environment, the implementation of the immunization strategy and the time difference between the virus spread and the implementation of the immunization strategy are all take into account. So, it can describe the process of the virus propagation in the cloud environment more realistic.

2. HSIR Model

The paper only analyzes the propagation process of one virus in order to facilitate the research. If only one virus is considered, then the nodes in the network can be divided into four states: H represents the healthy nodes which are not immune, this kind of nodes are healthy nodes and their neighbors are not infected, if their neighbors become infected nodes then the nodes become susceptible nodes; S represents the susceptible nodes, this kind of nodes are healthy nodes, but there are infected nodes in their neighbors; I represents the infected nodes; R represents the immune nodes, this kind of node are immune, never to be infected.

The relevant definitions and assumptions of the model are as follows:

**Definition 1:** propagation rate

Virus will be propagated from the infected nodes to its neighbors in the unit time, the neighbor nodes which are infected take the proportion of the total neighbors is the propagation rate, represented by $\theta$.

**Definition 2:** immunization rate

The nodes which are not infected (include the healthy nodes and the susceptible nodes) and the infected nodes in the network will change into immune nodes by a certain probability respectively after injected the immunization strategy into the network at a unit time. The total probability of these kinds of nodes is called immunization rate and represented by $\rho$, in which $\rho = \gamma + \omega + \delta$.

**Definition 3:** immune delay

There is a delay between the virus outbreak and the immune. This delay is called immune delay, represented by $\mu$.

**Hypothesis 1:** there is only virus in the network at the initial time; all the nodes can be infected at that time; the healthy nodes $H_0 = N$, the susceptible nodes $S_0 = 0$, the infected nodes $I_0 = 0$ and the immune nodes $R_0 = 0$.

**Hypothesis 2:** the node is infected instantaneously. The process of the virus propagation in the node is not considered. The time which virus propagate from one node to the other one is fixed, called a unit time.

**Hypothesis 3:** the cloud network studied in this paper is the network of logic-level.
**Hypothesis 4:** the virus is always in active state. The infected node infects its neighbors by the probability $\theta$. Different virus has different infection rate. The infection rate is fixed in the propagation process for one virus.

The state transition of this model is shown in Figure 1:

![Figure 1. The State Transition of HSIR Virus Propagation Model](image)

The virus can spread freedom when the immunization strategy does not exist in the network that is the time $t \leq \mu$. The immunization strategy is injected into the network when $t > \mu$, then the propagation of the virus and the immunization strategy both exist in the network. $N$ represents the total number of the nodes in the cloud at the initial time. $A$ represents the nodes added into the cloud at a unit time. $\beta C(N)$ represents the contact rate, in which $C(N)$ is the contact number that the infected nodes contact to the susceptible nodes; $\beta$ represents the infection rate of the infected node at a unit time. $\beta C(N) SI < k > / N$ represents the probability that the healthy node change into the susceptible node at a unit time. $< k >$ represents the average degree of the nodes in the network. $\alpha$ represents the probability that the susceptible node change into the infected node. $\gamma$ represents the probability that the infected nodes are immune; $\omega$ represents the probability that the susceptible nodes are immune; $\delta$ represents the probability that the healthy nodes are immune; $\sigma$ represents the rate of the nodes exited from the cloud. $d$ represents the rate of the dead nodes.

The dynamic differential equation of the USIR propagation after $t > \mu$ is as follows:

\[
\begin{align*}
\frac{dH(t)}{dt} &= A - \beta C(N) \frac{< k >}{N} H(t) S(t) I(t) - \delta H(t) - \sigma H(t) \\
\frac{dS(t)}{dt} &= \beta C(N) \frac{< k >}{N} H(t) S(t) I(t) - \omega S(t) - \alpha S(t) - \sigma S(t) \\
\frac{dI(t)}{dt} &= \alpha S(t) - \gamma I(t) - dI(t) - \sigma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t) + \omega S(t) + \delta H(t) - \sigma R(t)
\end{align*}
\]

The total number of the nodes in the network is $N$ at the initial time, that is $H(t) + S(t) + I(t) + R(t) = N$. $H(t)$ represents the number of healthy nodes which are not immune at time $t$; $S(t)$ represents the number of susceptible nodes at time $t$; $I(t)$ represents...
the number of infected nodes at time \( t \); \( R(t) \) represents the number of the immune nodes at time \( t \).

Make \( \sigma dt = d\tau \), it has:

\[
\begin{align*}
\frac{dH(t)}{d\tau} &= \frac{A}{\sigma} - \beta_i H(t) S(t) I(t) - \delta_i H(t) - H(t) \\
\frac{dS(t)}{d\tau} &= \beta_i H(t) S(t) I(t) - \omega_i S(t) - \alpha_i S(t) - S(t) \\
\frac{dI(t)}{d\tau} &= \alpha_i S(t) - \gamma_i I(t) - d_i I(t) - I(t) \\
\frac{dR(t)}{d\tau} &= \gamma_i I(t) + \omega_i S(t) + \delta_i H(t) - R(t)
\end{align*}
\]  

(2-2)

This system is equivalent to the system (1), in which, \( \beta_i = \frac{\beta C(N) < k >}{\sigma N} \), \( \delta_i = \frac{\delta}{\sigma} \), \( \omega_i = \frac{\omega}{\sigma} \), \( \alpha_i = \frac{\alpha}{\sigma} \), \( d_i = \frac{d}{\sigma} \), \( \gamma_i = \frac{\gamma}{\sigma} \). The total number of the nodes in the network is represented by the differential equation

\[
\frac{dN(t)}{dt} = \frac{dH(t)}{dt} + \frac{dS(t)}{dt} + \frac{dI(t)}{dt} + \frac{dR(t)}{dt} = \frac{A}{\sigma} - N - d_i I(t).
\]

By using the variable \( N \) instead of the variable \( H \), it has:

\[
\begin{align*}
\frac{dS(t)}{d\tau} &= \beta_i H(t) S(t) I(t) - P_1 S(t) \\
\frac{dI(t)}{d\tau} &= \alpha_i S(t) - P_2 I(t) \\
\frac{dR(t)}{d\tau} &= \gamma_i I(t) + \omega_i S(t) + \delta_i H(t) - R(t) \\
\frac{dN(t)}{d\tau} &= \frac{A}{\sigma} - N(t) - d_i I(t)
\end{align*}
\]  

(2-3)

In which, \( P_1 = \omega_i + \alpha_i + 1 \), \( P_2 = \gamma_i + d_i + 1 \). System (2-3) is equivalent to system (2-1) and system (2-2). The system (2-1) and system (2-2) can be comprehended by studying the system (2-3). From the biology angle, the system (2-3) can be studied in \( M = \{(H, S, I, R) \in R^4_+: 0 \leq S + I + R \leq N \leq \frac{A}{\sigma}\} \). Considering the practical significance of the system (2-3), various initial values are in \( M \), so \( M \) is the maximum positive invariant set of the system (2-3).

270
3. Balance Point and the Proof of its Stability

The balance point of the model is found and its stability is proved in the following content. The related theorem is as follows:

**Theorem 1:** (Lyapunov Stability Theory) in the regional $D = R^n$ of the autonomous system: (1) there is an infinite positive definite (negative definite) function $V(x)$; (2) the total derivative for $t$ of the function $V(x)$ along the solution of the given system is always negative (always positive) in the total phase space; (3) the set of points $F$ which satisfy $\frac{dV}{dt} = 0$ does not contain the other whole trajectory of the system besides the origin. Then the zero solution of the given system is said to be global asymptotic stable.

**Theorem 2:** assume the characteristics root of the coefficient matrix $A$ of the system are $\lambda_1, \lambda_2, \ldots, \lambda_r$, then it has three conclusions as follows:

1. If $\lambda_1, \lambda_2, \ldots, \lambda_r$ all have negative real part, then the zero solution of the system is said to be asymptotically stable.

2. If at least one characteristic root in $\lambda_1, \lambda_2, \ldots, \lambda_r$ has positive real part, then the zero solution of the system is not stable.

3. If there is no root with positive real part in $\lambda_1, \lambda_2, \ldots, \lambda_r$, but there is zero root or the pure imaginary root of zero real part, then the zero solution of the system is stable when the primary factor of zero root or the zero real part are all the first-order. If there is at least one primary factor of zero root or the zero real part greater than 1, then the zero solution of the system is not stable.

**Theorem 3:** (Hurwitz criterion) For an $n$th-degree polynomial with constant coefficient $a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0$, making

$$\begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \\ \vdots & \vdots & \vdots \\ a_{2n-1} & a_{2n-2} & a_{2n-3} & a_{2n-4} & \cdots & a_n \end{vmatrix} = a_n \prod_{i}^{n-1} \lambda_i$$

in which $a_0 > 0$, $a_i = 0$, $i = n + 1, n + 2, \ldots, 2n - 1$ then the sufficient and necessary condition about all the roots of the equation have negative real part is $\prod_i > 0, \prod_2 > 0, \cdots, \prod_n > 0$.

If the right sides of the four differential equations in system (2-3) are equal to 0, then:

$$\frac{A}{\sigma} - \beta_iH(t)S(t)I(t) - \delta_iH(t) - H(t) = 0 \quad (2-4)$$
\[
\beta_i H(t) S(t) I(t) - P_i S(t) = 0 \quad (2-5)
\]
\[
\alpha_i S(t) - P_i I(t) = 0 \quad (2-6)
\]
\[
\gamma_i I(t) + \omega_i S(t) + \delta_i H(t) - R(t) = 0 \quad (2-7)
\]

In addition, there is:
\[
\frac{A}{\sigma} - N(t) - d_i I(t) = 0 \quad (2-8)
\]

By the formula (2-4), (2-6), (2-7), (2-8), one has:
\[
H^* = \frac{A \alpha_i \sigma d_i^2}{\beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_i \sigma d_i^2}, \\
S^* = \frac{P_2 (A - \sigma N)}{\sigma d_i}, \\
I^* = \frac{A - \sigma N}{\sigma d_i}, \\
R^* = \gamma_i I^* + \omega_i S^* + \delta_i H^*
\]

\[P_1 = \omega_1 + \alpha_1 + 1, P_2 = \gamma_1 + d_i + 1.\] Substituting the values of S and I into (2-5) yields:
\[
\frac{P_2}{\alpha_i \sigma d_i} \left( \frac{A \alpha_i \beta_i d_i (A - \sigma N)}{\beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_i \sigma d_i^2} - P_1 \right) (A - \sigma N) = 0 \quad (2-10)
\]

Assume \( F(N) = \frac{P_2}{\alpha_i \sigma d_i} \left( \frac{A \alpha_i \beta_i d_i (A - \sigma N)}{\beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_i \sigma d_i^2} - P_1 \right) \). It is known from (2-10) that the system (2-3) have disease-free equilibrium in the interval \((0, \frac{A}{\sigma})\), that is \( N = \frac{A}{\sigma} \), then the disease-free equilibrium is
\[
P^0 = \left( \frac{A}{\sigma (\delta_i + 1)}, 0, 0, \frac{A \delta_i}{\sigma (\delta_i + 1)} \right)
\]

Also there are:
\[
F\left( \frac{A}{\sigma} \right) = -\frac{P_1 P_2}{\alpha_i \sigma d_i} < 0, \quad F(0) = \frac{P_1 P_2}{\alpha_i \sigma d_i}, \\
\left[ \frac{A \alpha_i \beta_i d_i (A - \sigma N)}{P_1 \beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_i \sigma d_i^2} - 1 \right] = \frac{P_1 P_2}{\alpha_i \sigma d_i} (R_0 - 1).
\]
\[ R_0 = \frac{A\alpha_1\beta_i d_i (A - \sigma N)}{P_1 \left[ \beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_1 \sigma d_i^2 \right]} \] is the threshold of the system (2-3). \( F(N) \) is monotonic decreasing when \( R_0 > 1 \), moreover \( F(0) > 0 \), \( F(\frac{A}{\sigma}) < 0 \), so \( F(N) = 0 \) has a unique positive root in the interval \( \left( 0, \frac{A}{\sigma} \right) \). Thus the system (2-3) has a unique endemic equilibrium \( P^* \left( H^*, S^*, I^*, R^* \right) \). The values of \( H^*, S^*, I^* \) and \( R^* \) are determined by (2-9).

**Theorem 4:** the disease-free equilibrium \( P^0 \) is global asymptotic stable in \( M \) when \( R_0 \leq 1 \). \( P^0 \) is not stable when \( R_0 > 1 \).

**Proof:** take the Lyapunov function when \( R_0 \leq 1 \):
\[ V = \alpha_1 S + P_I \] (2-11)

Then,
\[
\begin{align*}
V'_{\text{system(2-3)}} &= \beta_i \alpha_1 H(i) S(i) I(i) - \alpha_i P_I S(i) + \alpha_i P_I S(i) - P_I P_I I(t) \\
&= \left[ \beta_i \alpha_1 H(i) S(i) I(t) - P_I P_I \right] I(t) \\
&= P_I P_I \left[ \frac{A\alpha_1\beta_i d_i (A - \sigma N)}{P_1 \left( \beta_i P_2 (A - \sigma N)^2 + (\delta_i + 1) \alpha_1 \sigma d_i^2 \right)} - 1 \right] I(t) \\
&= P_I P_I \left( R_0 - 1 \right) I(t)
\end{align*}
\]

So, it has \( V' \leq 0 \) when \( R_0 \leq 1 \). In addition, \( V' = 0 \) only when \( I(t) = 0 \) or \( R_0 = 1 \). Making \( M = \{(H, S, I, R) | V' = 0 \} \), then the maximum positive invariant set of \( M \) is a set of single point \( \{P^0\} \). Therefore it is known from the theorem 1 that the disease-free equilibrium \( P^0 \) is global asymptotic stable. When \( R_0 > 1 \), it has \( V' > 0 \), so \( P^0 \) is not stable.

**Theorem 5:** system (2-3) has a unique endemic equilibrium \( P^* \) and is asymptotically stable when \( R_0 > 1 \).

**Proof:** the Jacobian matrix \( J(P^*) \) of system (2-3) at \( P^* = (S^*, I^*, R^*, N^*) \) is:
\[
\begin{pmatrix}
\beta_i H^* I^* - P_I & \beta_i H^* S^* & 0 & 0 \\
0 & -P_2 & 0 & 0 \\
\alpha_i & \gamma_1 & -1 & 0 \\
0 & -d_i & 0 & -1
\end{pmatrix}
\]

Its characteristic equation is \( \lambda E - J(P^*) = 0 \), in which \( E \) is unit matrix, so the characteristic equation can be reduced to:
\[
(\lambda + 1)^2 \left[ (\lambda + P_1 - \beta_i H^* I^*) (\lambda + P_2) - \alpha_i \beta_i H^* S^* \right] = 0 .
\]
Define $\alpha_1 = P_1 + P_2 - \beta_1 H^* I^*$, $\alpha_2 = P_2 \times \left( P_1 - \beta_1 H^* I^* \right) - \alpha_t \beta_t H^* S^*$, then the above equation can be expressed as:

$$(\lambda + 1)^2 (\lambda^2 + \alpha_1 \lambda + \alpha_2) = 0 \quad (2-12)$$

Because:

$$P_1 - \beta_1 H^* I^* - \alpha_t \beta_t H^* S^*$$

$$= P_1 - \beta_1 H^* \left( \frac{A - \sigma N}{\sigma d_1} - \alpha_t \frac{P_2 (A - \sigma N)}{\sigma \alpha d_1} \right)$$

$$= P_1 - \beta_1 H^* I^* \left( 1 - P_2 \right)$$

$$= P_1 - \beta_1 H^* I^* \left( 1 - \gamma_1 - d_1 - 1 \right)$$

$$= P_1 + \beta_1 H^* I^* \left( \gamma_1 + d_1 \right) > 0$$

Then:

$$\alpha_1 = P_1 + P_2 - \beta_1 H^* I^* > 0;$$

$$\alpha_2 = P_2 \left( P_1 - \beta_1 H^* I^* \right) - \alpha_t \beta_t H^* S^*$$

$$= \left( \gamma_1 + d_1 \right) \left( P_1 - \beta_1 H^* I^* \right) + P_1 - \beta_1 H^* I^* - \alpha_t \beta_t H^* S^* > 0$$

So:

$$\begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} = a_1 a_2 > 0.$$
are considered in the HSIR model. It is accessed and quit by lots of nodes. So the HSIR model can describe the virus propagation in the cloud accurately.

(2) The immune of the health node is considered.

The computer users in the cloud environment often take some anti-virus measures (such as early warning and removal directly after the virus detected, pre-immune measures et al) to prevent the propagation of the virus. The generation of the immune nodes is considered in the HSIR model. The immune nodes not only contain the infected nodes but also contain the healthy nodes. So the environment of the virus propagation is considered more comprehensively in the HSIR model.

(3) The time difference between the virus spread and the implementation of the immunization strategy is considered.

The implementation of the immunization strategy often lags behind the virus propagation in the practical situations. The implementation of the immunization strategy is considered in the SIR and SIRS et al., models, but the time difference between the virus spread and the implementation of the immunization strategy is not considered. The HSIR model considered the above cases, so it can describe the process of the virus propagation in the cloud more accurately.

5. Summary

The impacts of the changes about the number of the nodes with different states by the external factors are not considered in the traditional models. So, they are all passive propagation model. The cloud environment is large, distributed and dynamic. So the traditional models are not suitable for the cloud environment. A virus propagation model of computer which applies to the cloud environment, HSIR propagation model, is proposed in this paper. It takes the dynamic changes of the cloud environment, the implementation of the immunization strategy and the time difference between the virus spread and the implementation of the immunization strategy into account. In addition, the balance point of the HSIR model is found and the stability of the balance point is proved by the mathematical theory. And the threshold of the virus propagation in the HSIR model is found too.

Acknowledgements

This work is supported by Information Fusion and Cloud Services Network Platform project, which is a part of Internet Innovation Union (IIU).(Minister of Education, Science and Technology Development Center[2012]32). The research is also funded by Could Resource Management in Dragon-lab Project (ZH2011102).

References


Authors

Tongrang Fan: Doctor, Professor, Master Tutor, assistant dean of School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, China. Her major field of study are network technology and network security, Information processing and Network and education.

Yanjing Li: Graduate student of School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, China. His major field of study is network security.

Feng Gao: Doctor, Lecturer, teacher of School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, China. His major field of study are Cross-platform mobile development technology, the Internet of things technology.