Chapter 8
Virus Propagation Modeling in Facebook

Wei Fan and Kai-Hau Yeung

Abstract In recent years, online social network services have become part of people's life. One of the consequences these services bring about is the security problem. In this paper, we propose a virus model based on the application platform of Facebook. We also model virus propagation through emails and compare the behaviors of virus spreading in Facebook and email network. It is found that while Facebook provides a platform for application developers, it also provides the same chance for virus spreading. Virus will spread faster in Facebook network if users of Facebook spend more time on it. Moreover, users' network generated with BA scale-free model is compared with some sampled networks of Facebook in this paper. The results show that applying BA model in simulations will overestimate the number of infected users a little, but still reflect the trend of virus spreading.

8.1 Introduction

Computer viruses can spread through the Internet in many ways, such as email, instant messengers (IM), and P2P file sharing. Currently, online social network service (SNS) becomes popular. This kind of service provides a platform for people to have fun. People can share interesting things in their life with their friends on these websites, and they can also take part in some activities or join groups online. But these characteristics give hackers opportunity to attack these users. The virus spreading in SNS network is similar to that in an email network or an instant messenger network. All of them can spread virus by sending or sharing files which contain malicious codes. If a user of these networks gets infected, the infected

W. Fan (✉) · K.-H. Yeung
Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China
e-mail: fanwei.fw@gmail.com; eeayeung@cityu.edu.hk

T. Özyer et al. (eds.), *The Influence of Technology on Social Network Analysis and Mining*, Lecture Notes in Social Networks 6, DOI 10.1007/978-3-7091-1346-2_8,

account will automatically send the same email or file to the users in his/her contact list, which helps virus spread quickly.

There have been some models to simulate the virus propagation in those networks. C. Zou et al. describe their email virus propagation model in [19–21]. They assume that email virus spreads though network by containing virus in email attachments. This model accounts for users' email checking time intervals and the probability of users to open these attachments. And they found that as users' email checking time becomes more variable, the virus spreads faster. In this model, the probability that a certain user gets infected is a constant. In [9,10] T. Komninos et al. propose a worm propagation model for email, IM and P2P networks. They consider that as time grows, users' behaviors should be different. So the probability that a user opens a malicious attachment is not a fixed value, but will decrease as time goes, which is different from Zou's model. And the model of virus propagation in P2P network in [17] assumes that the probability of a user to download an infected file is a function of the ratio of infected files to total files. And this ratio can be affected by users' downloading and executing behaviors.

Although virus propagation models have been proposed for email, IM and P2P networks, these models are not suitable for SNS networks. Besides sending messages while using email services, users of SNS networks can upload files to the Internet. To ensure that users can share their life in time, the activities that a user takes will appear in his/her friends' news feed, so all the friends can read the news when they are online. Moreover, different from email or IM networks, some people use SNS for entertainment, and spend many hours on that every day. These features are helpful for virus propagation. As the behavior of SNS users can be more complex than that in other networks, it is necessary to construct new models for virus propagation in SNS networks. Recently, there are reports that some SNS websites, such as Facebook and MySpace, were attacked by hackers [8,18]. Among these SNS providers, Facebook attracts the largest population. Also, there is an application platform in Facebook. Everyone can build their applications with this platform, which can be easily utilized by hackers. So in this paper Facebook is the focused network in our study.

In this paper, two models of virus propagation in Facebook are proposed. The first is based on the Facebook application platform. In this model, it is assumed that hackers may utilize this platform to post applications along with viruses. Users may install these malicious applications in turn. After being infected, these users will send invitations to their friends. As will be reported later, installations of malicious application spread through network faster than other normal applications with the same initial conditions. In the second model, virus spreads through sending messages to friends. It is a traditional way, just like sending email with malicious attachments. More generally, hackers can post pictures or links that contain Trojans or worms. For simplicity, we assume that these methods are similar and we will describe them as sending messages. As will be reported later from our simulation results, it is found that if some people take SNS networks as entertainment tools and spend more time on them, virus will spread faster.

8.2 Facebook User Network Topology

We present the users of Facebook as a network with N_{user} nodes in this paper. Users are nodes in the network and each node is assigned a number $i, i = 1, 2, \ldots N_{user}$. An edge between two nodes i and j means these two users are friends. In Facebook, user i becomes a friend of user j with their names in each other's friends list, so this network is undirected. We also define that the node degree is the number of friends a user has. Some results have shown that email networks have a scale-free topology [5, 14]. And recently, researchers have studied the structure of some online social networks topology, such as MySpace, orkut, cyworld, and so on [2, 11, 13]. They found that these networks all have power-law degree distributions. Their degree distribution can be described as $P(k) \sim k^{-\gamma}$, here the probability that a node connects to k other nodes is $P(k)$, and the power-law exponent $\gamma > 0$. Most nodes of these networks have small degrees but a few nodes have many connections. This structure has been demonstrated to be vulnerable and the well connected nodes are crucial in epidemic spreading [4, 12, 15, 16].

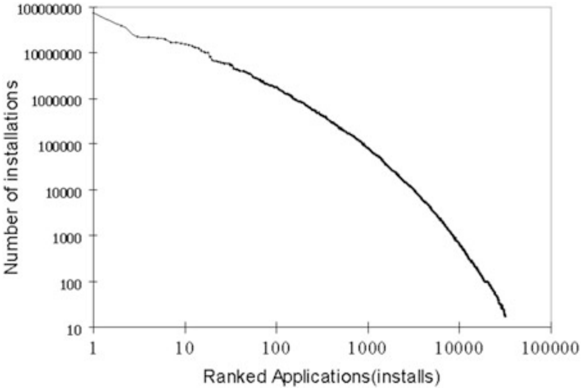
In our simulations, we assume that the nodes' degrees of Facebook users' network exhibit the power-law distribution, as it is also one of the online social networks. Firstly, we will construct the users' network with Barábasi-Albert (BA) scale-free network model [3] to study both virus propagation models. The algorithm of BA model is as follows: (1) Starts with m_0 connected nodes, and $m_0 > 1$. (2) At each time step, add a new node to this network, and this node is connected to m existing nodes ($1 \leq m \leq m_0$). The probability that a new node connects to node i is $\prod(k_i) = k_i/(\sum_j k_j)$. In BA model, $\gamma = 3$. Edges are added with a preferential attachment, so the nodes with greater degrees will get more connections.

In this paper, the actual Facebook network will be sampled to verify that whether BA model can fit Facebook well. Firstly we take a user from Facebook randomly. This user is the source node and then snowball method is used to gain users. After enough nodes are obtained, connections among these nodes are kept, but the edges pointing to the nodes outside the sampled network will be removed. This method is repeated and several sampled networks are obtained. We will apply two of these sampled networks in simulations. The sampled networks and BA networks will be compared in this paper.

8.3 A Model Based on Facebook Application Platform

One of the Facebook's successes is its application platform. By using this platform, companies and individuals can develop their applications. Users of Facebook can add these applications to their accounts. And if they like the applications which they installed, they can send invitations to their friends and invite them to join. Because there will be more fun if more friends are using the same applications. Most applications are also designed to remind users to invite their friends. More

Fig. 8.1 The distribution of number of installations of each application



than 95 % of users have used at least one application, and on average, every day there are 140 new applications added [6]. However, it has been reported that some new applications become available along with virus [8]. If a user installs this kind of application, his/her account is infected and fake messages are forwarded to all his/her friends automatically to persuade them to install the same application. This will increase the probability that a user installs it. Considering the great number of daily installations, damage of this kind of applications may be severe. So it is necessary to construct a virus propagation model based on these applications. It has been shown that the installation of applications has a preferential characteristic in [7]. That means an application with greater number of installations can attract more new users. This is because this application has more chances to be seen by people, and people are more willing to install such a popular application. Moreover, a user who has installed more applications has a higher probability to install new applications. The distribution of number of installations per application is shown in Fig. 8.1. The data is obtained from Adonomics [1].

Gjoka developed a model to simulate the users' number of installations of Facebook applications [7]. With the input of the list of applications, number of installations per application and number of users, this model can generate a graph which shows the power-law distribution of number of installations per user. This model is helpful for us to model the existing number of installations for each user, but it cannot reflect the behavior when a user encounters an application invitation. In this paper we proposed a Facebook virus propagation model based on the application platform. This model not only follows the spreading law of normal applications, but also contains the characteristics of virus spreading.

Our Facebook user network has N_{user} nodes, and each node is assigned a number $i, i = 1, 2, \ldots N_{user}$. We assume that the total number of available applications is N_{app}, and each one is denoted by $k, k = 1, 2, \ldots N_{app}$. Each application has a number of installations to show how many users have installed it. Since there are new installations every day, the number of installations per application is not a fixed value. So the number of installations of application k at time step t is $Install_k(t)$,

$k = 1, 2, \ldots N_{app}$. From the data of Adonomics we can get the knowledge of initial number of installations per application before the virus begins attacking. We assume that the virus starts spreading at $t = t_0$. We will study our model in a network which is much smaller than the real Facebook network, so we will not use the original data of installations. We scale down the size of the network, total number of applications, and the number of installations per user/application. We can create a list of applications and assign a $Install_k(t_0)$ for each application k. The distribution of $Install_k(t_0)$ is similar to the curve in Fig. 8.1. Next we will model the behaviors of applications as described below.

1. Construct the initial number of installations for users

   With the list of existing applications and $Install_k(t_0)$, $k, k = 1, 2, \ldots N_{app}$, we can construct the initial number of installations per user using the model defined in [7]. In the beginning, $Install_k(t_0)$ is ready for each application, but all the users in our network have not installed any application. We need to distribute each installation to users. At each step, each installation of $Install_k(t_0)$ over all the N_{app} applications is assigned to one of the users with a probability. The probability of one installation to be installed by user i is

$$P_{user}(i, t) = \frac{Apps_i(t)^{\rho} + init_{user}}{\sum\limits_{j=1}^{N_{user}} (Apps_j(t)^{\rho} + init_{user})}. \tag{8.1}$$

   In [7] it is found that the number of installations per user has power-law distribution. That means a user who installs more applications has more chances to be selected to install other applications in our simulations. So we use preferential selection in this equation. Here $Apps_i(t)$ is the number of applications that user i has installed at time step t. The parameter ρ reflects the effect of preferential installation. $init_{user}$ is used to show the initial probability $P_{user}(i, t)$ for a user i who does not install any application at that time step. That is, if $Apps_i(t) = 0$, the initial probability for user i is $init_{user}/(\sum_{j=1}^{N_{user}}(Apps_j(t)^{\rho} + init_{user}))$. This probability would make sure that, even a user has no installation can also have the opportunity to be chosen in our simulations. This step is repeated until all $Install_k(t_0)$ installations of all N_{app} applications are exhausted. When the initialization is completed, the simulation steps described in Part 2 will be run.

2. Virus propagation

   The propagation of virus follows the steps described below:

   (a) Select the users who are infected in the beginning:

      The virus spreading starts from I_0 infected users at time step t_0. We randomly select the I_0 initial infected users from the network. Then the total number of applications N_{app} is added by 1. We label the malicious application by M. And we have $I(t_0) = I_0$, here $I(t)$ is the number of infected users in the network at t-th time step. The initial infected user(s) will send invitation messages to their friends.

(b) Maintain the distribution of installations:

The statistics of Facebook shows that there are many new installations every day, but the curve showed in Fig. 8.1 does not change significantly. The distributions of installation of applications are similar from time to time. So we should maintain the preferential characteristic of installations of applications. As we mentioned before, from Adonomics we can know how many installations are taken every day. We assume that in our simulations the number of new installations per day is m. The value of m is also scaled down due to the small size of our network. In this step, we select one application from the application list, the probability of application k being selected is

$$P_{app}(k,t) = \frac{Install_k(t) + init_{app}}{\sum\limits_{j=1}^{N_{app}} (Install_j(t) + init_{app})}. \tag{8.2}$$

Here $init_{app}$ defines the initial probability $P_{app}(k,t)$ for an application without any installation, which has the same function as $init_{user}$. Figure 8.1 in this chapter shows that the number of installations per application also has power-law distribution. Therefore, this equation in our model uses preferential selection, as well. It means that an application with more installations has higher probability to be selected to be installed by users. Then a user i is selected from the network with the probability $P_{user}(i,t)$. If this selected user i has installed this application k before, we will pick another user. We select application and assign it to a user for m times in this step. So we have m new installations at this step b. And if the malicious application is selected in this step, we will change the value of $I(t)$, and the infected users would send invitations to his/her friends.

(c) Dealing with the invitations:

Every user who has received c invitation(s) at this time step will install the malicious application with the probability

$$P_{virus} = \frac{\sigma}{(1 - \frac{Install_M(t)}{N_{user}} \times \frac{Apps_i(t)}{N_{app}})^c}. \tag{8.3}$$

In this equation, the numerator $\sigma = 0.05$, which is the real data we obtain from an application. And we use the denominator to reflect the number of installations' influence on the probability of invitations acceptance. The reason is similar to that of equations (1) and (2): the increments of installations of application/user can raise the possibility that a user accepts invitations. If the user i accepts the invitations, we will change the value of $I(t)$. In our model, a user receives more invitation messages means that he/she has higher risk to be infected.

(d) Time step t is added by 1 and we repeat step b and c for the next time step.

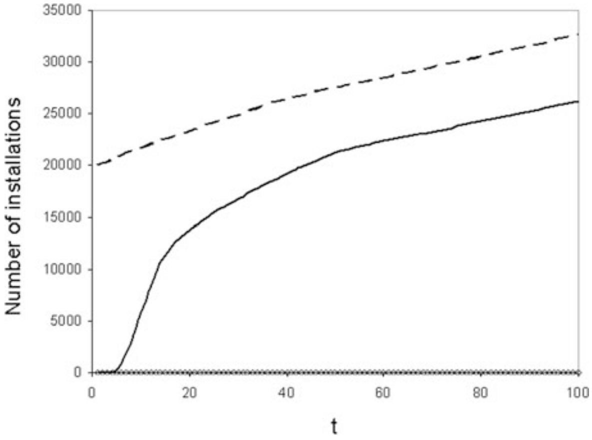


Fig. 8.2 The behavior of three applications over time. In this simulation, $N_{user} = 50,000$, $N_{app} = 100$ before the virus spreads, and $I_0 = 10$

8.3.1 Comparison of Malicious Application and the Top One

In this simulation, we record $I(t)$ at every time step t to show the virus spreading process in the network and the size of infected users in the end. In Fig. 8.2 we plot the behavior of malicious application, as the solid line shows. The dash line shows the behavior of the application which attracts the most users in the beginning. In this figure, the number of installations of the malicious application increases rapidly, and comes close to that of the top application. That is because the step c helps the virus to spread. But when the number of infected users reaches a certain value, its growth mainly comes from step b. So we can see from the figure that the growth rate is as the same as that of the top application after $t = 50$. And we also record the behavior of the application which has the same number of installations at $t = t_0 = 1$, as the diamond spots show. The line composed by these spots is nearly parallel and close to the x-axis. That is because its installations do not increase much. This implies that the number of installations of the application which has the same condition in the beginning does not change significantly.

8.3.2 Comparison of BA Networks with Different User's Average Degree

Users' decisions can be affected by their friends' behaviors. A user with higher degree may be infected easier, since he/she would receive more invitations. In BA scale-free network we can change the users' average degree, which is the average number of friends. Figure 8.3 plots the installations of malicious application in two different networks with average degree $< k >= 6$ and $< k >= 14$. This figure shows that a network with greater $< k >$ has more infected users, and the virus can spread faster in it.

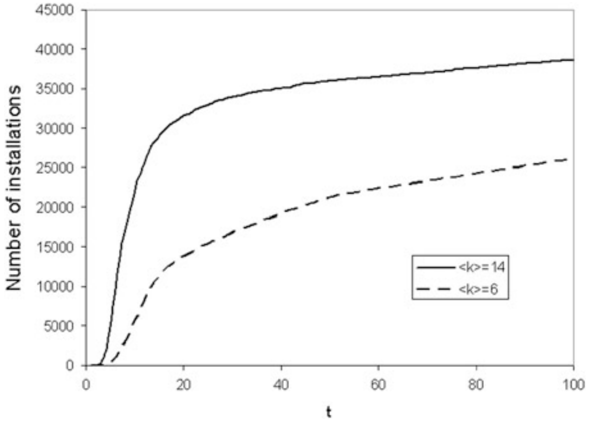


Fig. 8.3 The behavior of malicious application in networks with different average degrees

8.3.3 Comparison of Sampled Facebook Networks and BA Networks

In order to know whether BA model fits the Facebook network well, we take part of the users of Facebook and simulate our model in these sampled networks. Figure 8.4 shows the virus spreading in two sampled networks. These two networks have small power-law exponents that $\gamma \approx 1.5$, which is much smaller than BA network ($\gamma = 3$). Their average degrees are almost the same ($< k >= 34.2$ and $< k >= 35$), and the numbers of users of these two networks are 18,802 and 18,249, respectively. Their curves in this figure are close. We also draw the virus spreading processes in corresponding BA networks, which own the same average degrees and numbers of users. Although the clustering coefficients of the sampled networks are much larger ($c \approx 0.42$ while this value is about 0.009 in BA networks), the virus spread faster in BA networks. The reason is their small power-law exponents. The sampled networks are more heterogeneous than the BA networks, so more users have low degrees, which would slow down the speed of virus spreading.

In our model, infected users send invitations to their friends. This affects the behavior of users significantly, as we see from the results. Users who are friends in Facebook may be also friends in real life, so they will easily trust the invitations they received. So the virus can spread rapidly even there are only a few users installing it in the beginning. By comparing virus propagation in BA networks and that in sampled networks, it is found that the spreading in BA networks shows the general trend of the real situation. However, it is necessary to find a better network model if more accurate simulation results are required. This is because the number of infected users is overestimated on condition that BA model is used.

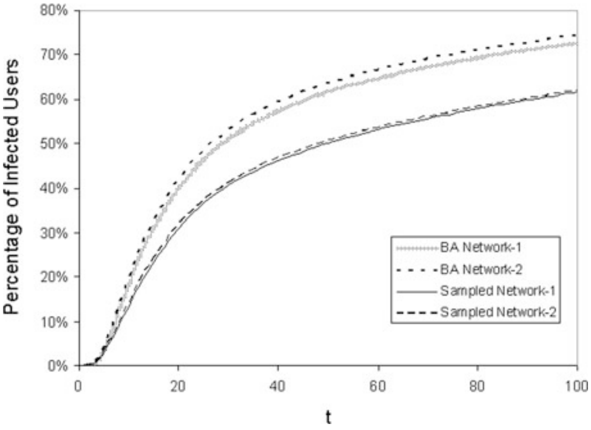


Fig. 8.4 The behavior of malicious application in BA network and sampled networks

8.4 A Model Based on Sending Messages

Our proposed second model is similar to an email virus propagation model. Email virus spreads by sending mails which contain attachments to users. If users open these email attachments, they will become infected. However, users can not add attachments to messages on Facebook, so hackers try to lead users to some third-party websites. These websites will inform users that their flash players are out of date, so users may download the new version of flash player, which is actually the virus, and install it [8]. The virus spreading process can be described like that: users log in Facebook from time to time. When a user is online and receives a message with a link to malicious website, he/she may delete this message without clicking this link, or click it and then become infected. If a user receives this message when he/she is offline, this user will process this message next time he/she logs in. And if a user is infected, the virus will send the same messages to all the friends of this user.

We can see that this spreading process seems to be the same as the email virus described in [19]. Both of them depend on users' interactions. And users check their accounts with dynamic time intervals. However, they have differences. In normal cases, while using email application,[1] people only check that if there is new mail and then log out. They may check email many times in a day, but will not stay on the web page and "play" their mailbox. On the contrary, people spend more time on Facebook and play online games. Every day more than 23 billion minutes are spent on Facebook which has 500 million active users [6]. That means on average each user spends more than 40 min on Facebook per day. If users get new messages when they are online, they can check mailbox immediately. For the ones who are online for hours every day, virus can spread faster. As a result, besides the time between two

[1] In this paper we only consider the case of Web mails.

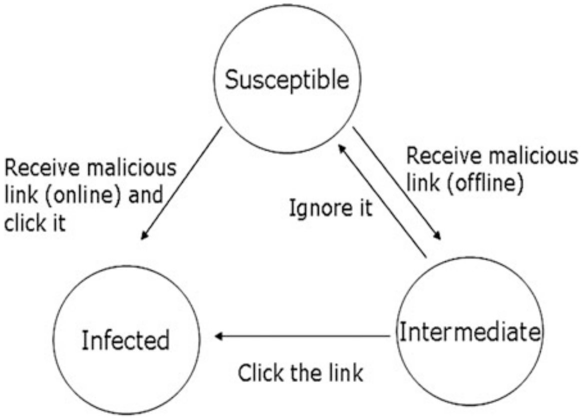

Susceptible
Receive malicious link (online) and click it
Receive malicious link (offline)
Ignore it
Infected
Intermediate
Click the link

Fig. 8.5 Three kinds of status of a user: susceptible, intermediate, and infected

log-in attempts of a user and the probability that a user clicks the direct malicious URL, we need the online time as another factor that affects virus propagation. And we assume that the probability that a user click the URL may not be a constant value, which may decreases if some users realize there is virus spreading in the network.

We still present our model in BA scale-free networks and sampled Facebook networks with N_{user} nodes. Users of Facebook are described as $i, i = 1, 2, \ldots N_{user}$. In our model nodes have three kinds of status–susceptible, intermediate, and infected, as Fig. 8.5 shows. In the beginning, all nodes are susceptible. If a node gets a message with malicious link in inbox but it is offline, this node becomes intermediate. An intermediate node can return back to the status of susceptible if it ignores this message. But it will become infected if the user clicks this link with a clicking probability. Infected nodes cannot return other status. The three users' interaction factors are as follows:

- Facebook log-in time $T_{login}(i)$ for node i $(i = 1, 2, \ldots N_{user})$, follows exponential distribution. It is the time intervals between user i's two log-ins. Its mean $E[T_{login}(i)]$ is independent Gaussian random variable that $E[T_{login}] \sim N(\mu_{Tl}, \sigma_{Tl}^2)$. In our simulation $E[T_{login}] \sim N(40, 400)$.
- Facebook online time $T_{online}(i)$ for node i $(i = 1, 2, \ldots N_{user})$, it is independent Gaussian random variable that $T_{online} \sim N(\mu_{To}, \sigma_{To}^2)$. It is reasonable that $T_{online}(i) < T_{login}(i)$. In our simulation we have $T_{online} \sim N(1, 100)$, and we assume that in the original email virus model $T_{online}(i) \equiv 1, i = 1, 2, \ldots N_{user}$.
- The probability that user i clicks a malicious link $P_{click}(i, t)$ is also independent Gaussian random variable. It is assumed that $P_{click}(t) \sim N(\mu_p(t), \sigma_p^2)$, in which $\mu_p(t)$ will decrease as time goes on, because more users would be aware of this scam if the number of infected users increases. We have $\mu_p(t) = \mu_0(1 - N_{infect}(t)/N_{user})$, here μ_0 is a constant and $N_{infect}(t)$ is the number of infected user at time step t. In our simulation $\mu_0 = 0.5, \sigma_p^2 = 0.09$.

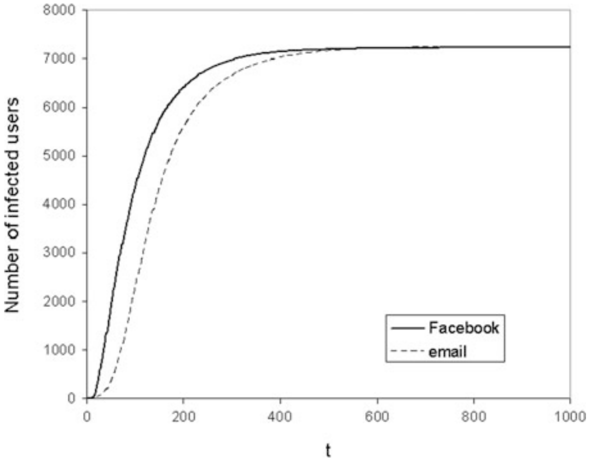


Fig. 8.6 The behavior of number of infected user. In this simulation, $N_{user} = 10,000$, $N_{infect}(t_0) = 10$, and $\mu_p(t) \equiv \mu_0$. The data is averages over 20 simulations

Virus propagation follows the steps below:

1. The propagation begins at $t_0 = 1$. We randomly select $N_{infect}(t_0)$ users who have malicious mails in their mailboxes in the beginning. These mails contain malicious links. If a user clicks the link, he/she will be infected and send the same mails to his/her friends.
2. At every time step, we check all the users who log in Facebook or are online at this step. If a user i has malicious mails in the mailbox, he/she would click the links with $P_{click}(i,t)$, or ignore these mails with probability $1 - P_{click}(i,t)$. After a user i logs out, we will generate new $T_{login}(i)$ and $T_{online}(i)$ for this user's next log-in. All our simulations are under the non-reinfection case, that is, infected users will not send the virus to their friends if they click the malicious link again.
3. Repeat step 2 for the next time step t.

8.4.1 Comparison of Facebook Networks and Email Networks

In this simulation, we present the Facebook network with BA network, and record the number of infected users $N_{infect}(t)$ at each time step. Figure 8.6 plots the $N_{infect}(t)$ in Facebook network with a constant $\mu_p(t) \equiv \mu_0$ and $T_{online}(i) > 1$, as the solid line shows. The dash line is the number of infected users in the original email network, which does not consider the users' online time. By comparing these two lines, we find that the virus will spread faster as some users spend lots of time on Facebook.

But users may be aware of this virus if more and more users are infected. So the probability that they click the links would decrease, and the number of infected users

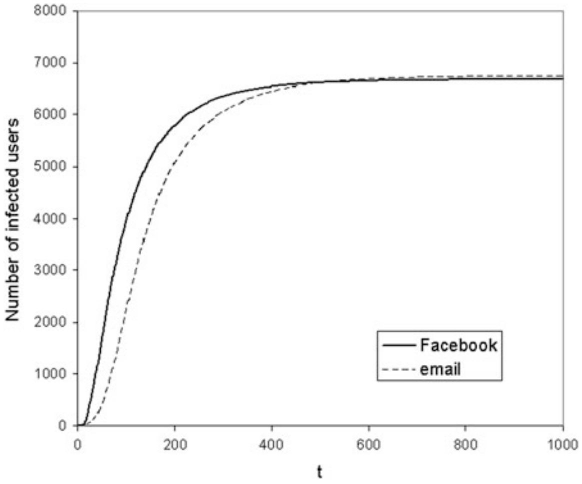


Fig. 8.7 The behavior of number of infected user. In this simulation, $N_{user} = 10,000$, $N_{infect}(t_0) = 10$, and $\mu_p(t) = \mu_0(1 - N_{infect}(t)/N_{user})$. The data is averages over 20 simulations

will be different. We plot it in Fig. 8.7, in which the $\mu_p(t)$ changes over time. We find that the sizes of infected users are smaller in both Facebook and email networks than those in Fig. 8.6. In Fig. 8.7, the solid line still rises faster than the dash line.

Our results indicate that under the same conditions, virus spreads faster in Facebook network than that in the original email network. But if we assume that the probability of a user to click the malicious link is not constant, the size of infected users will be smaller.

8.4.2 *Comparison of BA Networks with Different User Average Degree*

We also compare the virus' behavior in BA networks with different average degrees. Figure 8.8 plots the virus propagation process in two BA networks with $< k >= 6$ and $< k >= 14$. We also get the same conclusion that a network with greater $< k >$ has more infected users.

8.4.3 *Comparison of Sampled Facebook Networks and BA Networks*

The virus spreading in sampled Facebook networks and BA networks are compared, as well. As the characteristics of the two sampled networks that we mentioned in the first model are similar, we only plot the curve of one sampled network in Fig. 8.9. The corresponding BA network owns the same number of nodes and average degree.

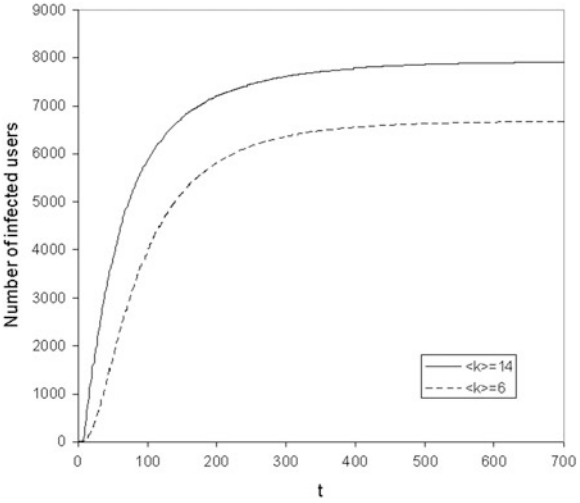


Fig. 8.8 The virus behavior in networks with different average degrees. In this simulation $N_{user} = 10,000$, $N_{infect}(t_0) = 10$, and $\mu_p(t) = \mu_0(1 - N_{infect}(t)/N_{user})$. The data is averages over 20 simulations

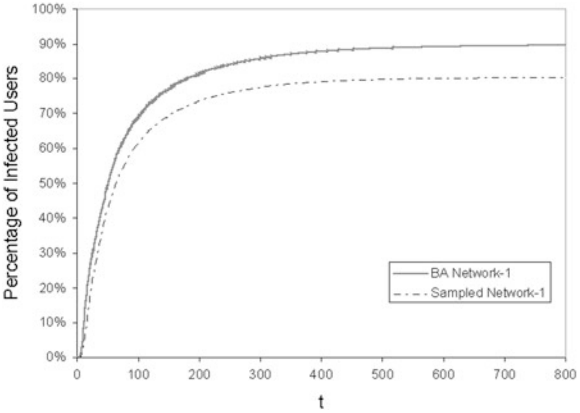


Fig. 8.9 The behavior of virus in BA network and sampled Facebook network

The same conclusion as that of the first model is obtained: virus still makes more users infected in BA network due to the high power-law exponent. But the difference is little.

8.5 Conclusion

In this paper, two models for virus propagation in Facebook network were proposed. In the first model based on Facebook's application platform, if the malicious application is installed by more users, it will become more popular and attract more installations. The results of the simulations showed that, although the malicious application attracts only a few users in the beginning, it can still spread rapidly. That is because users may trust their friends of Facebook and install it. While in the second model, which is similar to the email virus propagation model, the probability that a user is infected becomes smaller as more and more users get infected. We have found that the virus spread faster in Facebook than in the original email network, if we assume that people use Facebook for entertainment and spend more time on it. And it is demonstrated that the virus will spread faster in a network with higher average degree.

In the investigation of both models, we crawled part of the actual Facebook network and simulate our models in the sampled networks. It is found that virus spreads faster and makes more users infected in BA networks, which are more homogeneous than the sampled Facebook networks. This is because users' decisions may be affected by their friends' behaviors. So users with few friends of the sampled networks may slow down the spreading. Our results showed that, from the viewpoint of virus spreading, BA model is a good model and fit the Facebook network approximately. But if we would like to study Facebook from other point of view, we should look for other complex network models which can generate networks of similar power-law exponents and clustering coefficients to those of Facebook.

Acknowledgements This work was supported by Hong Kong Government General Research Fund (grant No. CityU-123608).

References

1. Adonomics: http://www.adonomics.com. Retrieved 10 June 2009
2. Ahn, Y., Han, S., Kwak, H., Moon, S., Jeong, H.: Analysis of topological characteristics of huge online social networking services. In: WWW '07: Proceedings of the 16th International Conference on World Wide Web, Banff (2007)
3. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**, 509–512 (1999)
4. Dezsö, Z., Barabási, A.-L.: Halting viruses in scale-free networks. Phys. Rev. E **65**, 055103(R) (2002)
5. Ebel, H., Mielsch, L.-I., Bornholdt, S.: Scale-free topology of e-mail networks. Phys. Rev. E **66**, 035103 (R) (2002)
6. Facebook: http://www.facebook.com/press/info.php?statistics. Retrieved 20 July 2010
7. Gjoka, M., Sirivianos, M., Markopoulou, A., Yang, X.W.: Poking facebook: characterization of OSN applications. In: ACM SIGCOMM Workshop on Social Networks (WOSN'08), Seattle, Aug 2008
8. Kaspersky Lab: Kaspersky lab detects new worms attacking mySpace and facebook. Message posted to: http://www.kaspersky.com/news?id=207575670 (2008)
9. Komninos, T., Spirakis, P., Stamatiou, Y.C., Vavitsas, G.: A worm propagation model based on scale free network structures and people's email acquaintance profiles. Int. J. Comput. Sci. Netw. Secur. **7**, 2 (2007)
10. Komninos, T., Stamatiou, Y.C., Vavitsas, G.: A worm propagation model based on people's email acquaintance profiles. In: WINE 2006, Patras (2006)

11. Kumar, R., Novak, J., Tomkins, A.: Structure and evolution of online social networks. In: KDD '06: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia (2006)
12. May, R.M., Lloyd, A.L.: Infection dynamics on scale-free networks. Phys. Rev. E **64**, 066112 (2001)
13. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, S.: Measurement and analysis of online social networks. In: IMC '07: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, San Diego (2007)
14. Newman, M.E.J., Forrest, S., Balthrop, J.: Email networks and the spread of computer viruses. Phys. Rev. E **66**, 035101 (2002)
15. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. Phys. Rev. Lett. **86**(14), 3200–3203 (2001)
16. Pastor-Satorras, R., Vespignani, A.: Epidemic dynamics and endemic states in complex networks. Phys. Rev. E **63**, 066117 (2001)
17. Thommes, R.W., Coates, M.J.: Modeling virus propagation in peer to peer networks. In: Information, Communications and Signal Processing, 2005 Fifth International Conference, Bangkok, pp. 981–985 (2005)
18. Ward, M.: Facebook users suffer viral surge. Message posted to http://news.bbc.co.uk/2/hi/technology/7918839.stm (2009)
19. Zou, C.C., Towsley, D., Gong, W.: Email virus propagation modeling and analysis. Technical Report TR-03-CSE-04, Umass ECE Department (2003)
20. Zou, C.C., Towsley, D., Gong, W.: Email worm modeling and defense. In: 13th International Conference on Computer Communications and Networks (ICCCN'04), Chicago, pp. 409–414 (2004)
21. Zou, C.C., Towsley, D., Gong, W.: Modeling and simulation study of the propagation and defense of internet e-mail worms. IEEE Trans. Dependable Secur. Comput. **4**(2), 105–118 (2007)