Poster: Insight into Virus Spreading in Social Networks via Percolation Theory

A*nahita Davoudi*
Dept. of Electrical Engineering and Computer science
University of Central Florida
Orlando, USA
anahita.davoudi@knights.ucf.edu

Keywords—percolation theory; social networks; virus spreading;

Based on information provided by Microsoft in 2011, one out of every 14 downloads is a piece of malware [1]. Although security of browsers has been improved greatly during the past few years, new forms of virus spreading have been seen in social media, especially in Facebook [2].

Typically, Facebook user gets a message from a friend which asked him/her to click on a video link, or download a picture which they usually use phrases like "It's one of your friends private video or picture" to intrigue him/her. The video requires to install a software which would lead to storing a malware in the user's computer and then the virus starts spreading itself out using user's Facebook account.

To enhance the security and manage the thread of virus infection through the social networks, it is essential to examine the response of networks to virus infection. However, enormous virus attacks in past years provide valuable data to study security of social networks; it always has been a challenging task due to availability of information to researchers and the huge size of real social network.

In this conditions, studying simple complex network (i.e. which can represent the behavior of large social networks) along with taking advantage of Percolation theory, can provide us with valuable insights to security of social network.

Percolation theory [3] & [4] is a probability model which describes the behavior of connected clusters in a random graphs. As an example, suppose there is water on a rock. Percolation theory investigates whether the water would be able to pass hole by hole and make its way through the rock. Percolation theory concept provides probabilistic framework to study the information spreading in complex networks including social networks.

Social networks are a type of scale free Networks [5]. The degree distribution for these networks (i.e. the probability which a node has a certain number of links) follows a power law distribution. This means that the degree distribution of social networks, or scale free networks in general, has no characteristic scale. In the random network, in contrast, each node has a similar number of edges [6].

In a social network graph, each node *i* might get infected by virus through its connection with node *j* (i.e. edge) with probability $P_{ij} \leqslant 1$. Some interesting questions which arise here are, "How the nodes characteristics affects these probabilities?", "How these probabilities along with other characteristics of network can affect the spreading of virus in the network?"

The main goal of this research is to investigate the effect of network characteristics on different aspects of virus spreading. Multiple small complex networks (i.e. graph), with known degree and weight probability distribution, have been studied to find out the rate of virus infection caused by initial infection of a specific node. Some of the major findings of the simulations are as follows;

- The infection rate caused by initial infection of node *i.*
- The portion of infected nodes caused by initial infection of node *i.*
- The pprobability that the virus of node *i* reaches to all of its neighbors (i.e. all friends of the infected Facebook account get infected).
- The pprobability that the virus of node *i* reaches to each fraction of its neighbors.
- The probability that all the neighbors within two hops with respect to the source get infected.

In each simulation, node *i* is infected initially and starts spreading massages containing the virus to all neighboring nodes. Each node *j* that has received the massage open the massage (with a pre-assigned probability) and gets infected. The infected nodes start spreading out the virus from the beginning of the next time step.

Studying the result of the simulations can help us to identify the vulnerable nodes based on their characteristics (i.e. number of edges and so on). This finding can be applied to real social networks to simply find vulnerable parts of the network to viruses and then try to modify network characteristics in order to enhance the security in that part.

References

[1] Ben Rooney, “Malware Is Posing Increasing Danger”, The wall street Journal, May 23rd, 2011.

[2] Kate Knibbs, “A new Facebook virus has already infected 800,000 users”, Digital Trends, 28 August, 2013.

[3] Daniel Genin, “Percolation: Theory and Applications”, NIST, 2007.

[4] Duncan S. Callaway, M. E. J. Newman, Steven H. Strogatz, and Duncan J. Watts, “Network Robustness and Fragility: Percolation on Random Graphs”, Physical Review Letters. 85, 5468, 2000.

[5] Barbara R. Jasny, Laura M. Zahn, and Eliot Marshall, “Scale-Free Networks: A Decade and Beyond”, Science Magazin, 24 July 2009, Vol. 325 no. 5939

[6] Lun Li, David Alderson, John C. Doyle, and Walter Willinger, “Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications”, Internet Mathematics, 16 March 2006, Vol. 2, No. 4: 431-523