A Computational Model of Computer Virus Propagation

Li-Chiou Chen
lichiou@andrew.cmu.edu
Department of Engineering and Public Policy
Center for the Computational Model of Social and Organizational Analysis
Carnegie Mellon University

Kathleen M. Carley
KathleenCarley@andrew.cmu.edu
Department of Social and Decision Sciences
Center for the Computational Analysis of Social and Organizational Systems
H.J. Heinz III School of Public Policy and Management
Carnegie Mellon University

Extended Abstract

Computer virus infection is the most common computer security problem. This problem has imposed significant amount of financial losses to organizations (CSI, 2000). Even though most organizations have installed anti-virus software in their computers, majority of them still experienced computer virus infection (ICSA, 2000). Most anti-virus software could not detect a new virus unless it is patched with the new virus definition file. Disseminating the new virus information and patches is hence important to raise user awareness. However, little research has focused on evaluating the effectiveness of disseminating new virus information on reducing virus infection. We hence propose a corporate response model to investigate the effectiveness of warning message propagation. In addition, we use the model to study the influence of social network topology on the virus and warning message propagation.

A computer virus is a segment of program code that will copy its code into one or more larger “host” programs when it is activated. A worm is a program that can run

independently and travel from machine to machine across network connections (Spafford, 1990). We will refer computer viruses to both computers viruses and worms in Spafford's definition since most viruses today can be propagated in both ways.

Epidemic propagation models (Bailey, 1975) have been applied on modeling the propagation of computer viruses (Kephart and White, 1993). Simulation models have been used to discuss the influence of the network topology (Kephart, 1994)(Wang, 2000)(Pastor-Satorras, 2001). However, neither the empirical topology data has been collected nor the characteristics of the topology have been further studied. In addition, the warning message propagation is related to the network topology, which could be a different network from the virus propagation network.

A corporate response model is developed to describe computer viruses propagation and the warning messages propagation. The components of the model include the inter-organizational social network topology, the computer network topology, the virus propagation mechanism, and the node state transition diagram. The four components are described as follows:

1) The inter-organizational social network topology and computer network topology are both represented as a graph G = (V, E, W(i,j)) where V is a set of nodes and E is a set of edges. W(i,j) denotes the link between node i and j where i, j $\in$ V. W(i,j) = 1 if a link exists between node i and node j and W(i,j) = 0 otherwise. We then apply social network analysis (Wasserman, 1994) measures, such as density and centralization, to characterize the network topology in our virtual experiments. A new measure, isolation, is needed to describe the computer virus propagation topology since the isolation nodes are critical to in the propagation process. Isolation of graph G, $I(G) = \frac{|S|}{|V|}$, is defined as the number of isolated nodes divided by the total number of nodes in the graph. Isolated nodes refer to the nodes

that do no have any links with other nodes in the same graph. S , = { i: $\forall$j$\in$ V, $\sum_{j=1}^{N} W(i, j) = 0$ }, denotes a set of isolated nodes.

2) The virus propagation mechanisms are categorized as one-to-one, one-to-many, many-to-one, and many-to-many. The category refers to the number of infection source and the number of infection targets at once.
3) The node state transition diagram is used to describe the dynamics of a node over time, such as if a node is infected by a virus or warned by a warning message. The change of states is determined by the propagation of viruses through the social network and the propagation of warning messages through the computer network. We assume that the nodes will receive an automatic warning message if their computers physically connect to a computer that has already had one.

Virtual experiments are conducted by varying the type of topology, the number of nodes, density and isolation. Experiment results show that random graph topology generated by the same density and isolation as real world data set could be used on modeling the social network of computer virus propagation.

In addition, isolation is not an effective strategy for an organization if warning messages are propagated in the network. Isolating an organization from other nodes in the social network could isolate the node from virus infection but isolate the node from the warning messages as well. This result contradicts with many organizations have assumed.

Bailey, N.J.T. 1975. *The Mathematical Theory of Infectious Diseases and Its Applications*. New York: Oxford University Press.

CSI 2000. 'CSI/FBI Computer Crime and Security Survey' *Computer Security Issues & Trend*.

ICSA 2000. 'ICSA Labs 6th Annual Computer Virus Prevalence Survey 2000': ICSA.net.

Kephart, J.O. and White, S.R. 1993. 'Measuring and Modeling Computer Virus Prevalence' *IEEE Computer Security Symposium on research in Security and Privacy.* Oakland, California.

Kephart, J.O. 1994. 'How Topology Affects Population Dynamics' in Langton, C.G. (ed.) *Artificial Life III*. Reading, MA: Addison-Wesley.

Pastor-Satorras, R. and Vespignani, A. 2001. 'Epidemic Dynamics and Endemic States in Complex Networks' . Barcelona, Spain: Universitat Politecnica de Catalunya.

Spafford, E.H. 1994. 'Computer Viruses as Artificial Life'. *Journal of Artificial Life*.

Wang, C., Knight, J.C. and Elder, M.C. 2000. 'On Computer Viral Infection and the Effect of Immunization' *IEEE 16th Annual Computer Security Applications Conference.*

Wasserman, S. and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.