IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-8, Aug 2018 E-ISSN: 2347-2693

Malware Dissemination and Anticipation Model for Ensuring Privacy in Time-Varying Population Networks

[1*]N. Sindhuja, [2]K. Ravi Kumar

[1,2]Dept. of. Computer Science, Tamil University, Thanjavur-613010, India

[*]*Corresponding Author: sindhu.ndrs@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 12/Aug/2018, Published: 31/Aug/2018

***Abstract*-** In modern days, more and more community joins social networks to contribute to information with others. At the same time, the in sequence sharing/spreading becomes far more frequent and convenient due to the wide usage. The research contented of computer networks comprises arrangement topology, network interchange uniqueness, and the authority of the network behavior on the whole set of connections. The spread and avoidance of network malware knowledge studied in network and have been one of the majority prolific fields in complex network dynamics research. Through our research, we found that some individuality of workstation network virus proliferation is similar to real world outbreak spread. Therefore, any misinformation should be exposed in time when it does not increase to a large group of populace. All preceding works deliberate either how the in succession is extend in the social complex or how to inhibit the further pervasion of an observed misinformation. However, no works considered how to discover the broadcasting of misinformation in time. A possible explanation is to set observers across the network to determine the suspects of misinformation established by the optimization problematic is NP-hard and deliver approximation assurances for an avaricious answer for various meanings of this problem by provides evidence that they are sub modular. In this accomplishment, a novel method to decide on a set of spectator in a social network with the minimum cost, where these observers assurance any misinformation can be discovered with a high likelihood before it reaches a surrounded number of users.

***Keywords*-** Information sharing, Misinformation, Online Social Network, Suspects, Optimization, Privacy.

I. INTRODUCTION

1.1 Information Security

Information security, every now and then shortened to InfoSec, is the practice of put a stop to unauthorized access, use, revelation, disturbance, alteration, examination, footage or obliteration of in order. It is a wide-ranging term that can be used not considering of the form the information may take (e.g., electronic, physical). Information security's primary focus is the evenhanded fortification of the Confidentiality, Integrity and Availability of data (also known as the CIA triad) while preserve a focus on well-organized policy implementation, all without slow down organization efficiency. This is necessarily accomplish from side to side a multi-step risk organization process that identifies possessions, threat foundation, vulnerabilities, potential collision, and potential controls, followed by measurement of the effectiveness of the risk administration plan.

To make conform this obedience, intellectual and professionals work together and seek to set basic leadership, policies, and manufacturing standards on code word, antivirus software, firewall, encryption software, legal liability and user/administrator preparation standards. This standardization may be further driven by a wide variety of laws and system that affect how data is admittance, procedure, store up, and transport. Nevertheless, the accomplishment of any principles and guidance within an entity may have incomplete effect if a culture of continual development isn't adopted.

The social networks are becoming increasingly human centric. In other words, the human social behaviors and activities must be carefully studied in association with such networks. However, it is intractable to conduct rigorous studies of human centric networking and communications over a large-scale virtual social network because of the large scale, complex topology and security problems of network. In addition, it is illegal to carry out special scientific researches and experimental developments on real social networks, such as social-aware routing protocol design, faults and worm propagation, and advertising promotion. As such, the structural modeling and conceptual properties of virtual social networks are well studied as a special form of the networks. A departure from the previous form of social network is the online groups or online communities which

allow users to create, post, comment to and read from their own interest and niche-specific subject.

Their major finding is that people are tempted to interact more with each other when they have similar age, language, and location. Likewise, they discovered that a link is significantly more likely to be friendly when its two endpoints have multiple common neighbors, which means that communities are mostly formed by the principle of ‘‘the friend of my friend is my friend’’. Online Social Network (OSN) websites have turn out to be an attractive objective for these worms (here in after referred to as OSN worms) because of the subsequent properties of the online social set of connections.

First, online social networks are small-world association, which mean they have the possessions of small average undeviating path length and high come together. The small typical unswerving path length property can reduce the proliferation time from one user explanation to another user report. Meanwhile, the high come together property put forward that users are tightly associated together, which make easy the explosion of OSN worms. Second, online social set of connections are also scale-free networks, which are a group of power-law networks where high-degree nodes tend to attach with other high-degree nodes. When an OSN worm infects the account of a well-liked user (i.e., user with a bulky number of friends), this scale-free possessions suggests that the worm can infect another well-liked account shortly. As a consequence, the worm can accomplish exponential enlargement by promulgate to the large acquaintances set of these popular consumer. Moreover, OSN worms also influence social engineering to increase the genuineness of worm communication.

1.2 Problem Specification:

The topologies of social worms consist of a social logical layer and an actual physical layer. The former has following characteristics:

1) they are defined as a” semi-directed network”, in which some edges are directed and others are undirected;

2) The in-degree of nodes tends to match the out-degree, and they both follow the power law distribution;

3) They are assortative, which implies that nodes with a high degree tend to connect with each other;

4) The weight of each edge denotes the propagation probability from user i to user j;

5) Each node in the social logical layer contains a group of nodes corresponding to nodes in the actual physical layer (i.e. hosts of different locations). The latter has some characteristics of power law, disassortative, rich-club and localization. For convenience of description, malicious emails, links or profiles are called” messages”.

II. PROPOSED APPROACH

The proposed system makes available a new analytical model to capture the infrastructure among the infected users by a set of dissimilarity equations, which together describe the overall propagation of the modern malware. Then it commence a new concept of virtual nodes to address the dryness in previous work, which can represent the situation of a user sending out one more round of malware copies each time this user gets contaminated. The perform result of empirical and hypothetical study to investigate why and how the projected SII model is superior to obtainable models.

The basic elements for the proliferation of modern malware are nodes and topology information. A node in the topology corresponds to a user in the network. The Virtual nodes, which can represent the circumstances of a user sending out one more round of malware copies whenever this user gets contaminated. For recent malware, bring to mind that a conciliation user may send out malware copies to neighbors every time the user visits those malware hyperlinks. Malware are also throwing out when certain events are generated. Thus, at an uninformed time t,a user may take delivery of multiple malware copies from an identical bordering user who has been compromised. To represent the repetitious spreading procedure of their disease and the self-start, we commence virtual nodes to present the kth disease caused by contaminated users at whatever time aperture the k th malware copy.

This work adopts theoretical concepts and technique from the field of social network analysis, namely, centrality measures and subgroup analysis, to capture the structural characteristics of both social network and technical network. In exacting, based on the unique properties of the malware dissemination process, we identify random-walk between’s as the appropriate centrality measure to evaluate the structural position of individual nodes. Modularity-maximizing decomposition is then applied to examine the embedded subgroup structure. Based on the derived centrality measures and the discovered subgroup structure, we formulate our structural risk model to examine the collision of individual-, group-, and network-level characteristics on malware dissemination dynamics. In order to calculate approximately and evaluate the projected structural danger representation, it constructs real secretarial networks and simulates the self-replicating malware proliferation development. It considers a real community network structure create from a large social multifaceted site and then maps nodes in the community set of connections to nodes in the technical network within the association. The detectors analyze malware behaviors continuously and try to oppose these techniques and approach hence; we need to allow detection development technique to lead malware

updating from side to side very well analytical process for malware activities and behaviors to fix any possible targeted threats. A new reproduction must be designed to surround real system illustration, to scrutinize the malware behaviors touching these samples after complicated malware keep informed. The objectives of this reproduction are to avoid organization threats before mortal infected by real malware.

III. METHODOLOGY

3.1 Evidence Collecting Module in OSN

The evidence collecting module was made to collecting worm propagation evidence (e.g., worm messages, worm updates).On the other hand, among the enormous amount of information exchanged in an OSN website, the challenge is how to gather only suspicious worm evidence. From the time, when OSN worms go after the social connections in propagation, an ally of an infected user account is more likely to receive worm propagation evidence. To leverage this advantage, we adapt the idea of honey pot here as "decoy friend". A decoy friend is a low-interactive honey pot, and it is created and added into a normal user's friends list by the detection system. When a user account is subjected to infection by an OSN worm, trap friends of that account can be given worm evidence. Similar ideas have been suggested for other types of networks. It distracts only receive malicious messages. However, the same assumption does not hold in our work. In fact, our system treats the collected information from decoys only as suspicious evidence because some normal user activities can also be observed by decoys. Decoys form a disguised surveillance network. We assign each decoy to be friends with several normal users so that a decoy cannot be easily spotted because of its small number of friends. In addition, there are a few practical concerns regarding applying decoy friends in real world OSN websites.

The first potential concern is related to user's information privacy because decoys collect suspicious information in the network. However, since users' data are all stored and kept in the OSN websites, we think our system will not cause new data/information leakage. Nevertheless, to alleviate such possible concern, our system will only keep the suspicious information for a short period of time. The second concern is that users might be reluctant to accept decoy friends. As such, a website will need to consult its users before assigning decoy friends to them. In fact, the OSN websites could provide incentives to encourage users to accept decoy friends. After all, both users and the OSN websites try to avoid worm infections for their own benefits. The third concern is on the number of decoy friends to be deployed in an OSN website. Besides user's reluctance, the population of decoys may negatively impact the popularity of an OSN website, because decoy friends do not contribute to any interactive activities such as discussions or communications.

3.2 Worm Detection Module

This module identifies the infected user accounts based on the suspicious worm propagation evidence. To distinguish actual worm evidence from normal user communications, this module applies correlation test on the suspicious evidence. The correlation test is based on similarities in the content and the structure of worm propagation evidence. One reason behind this similarity is that worm messages or updates composed by the same worm usually serve the same principle (e.g., advertising a malevolent link). Another reason is that the involuntary message production algorithms run by worms tend to use again words and phrases because of the limited size of their contender words set. In this module, we occupy a two-level spatial-correlation scheme, namely local connection and network correlation. To provide necessary in sequence to correlation, our system maintains a data arrangement called suspicious propagation evidence list (SPEL), which is connected with each preferred user.

In SPEL, each piece of substantiation is accumulate as a {decoy friend ID, receiving time, content} tuple. Local Correlation: Local correlation achieve similarity test among suspicious confirmation collected by two decoy friends dispense to the same selected user.

IV. IMPLEMENTATION

4.1 Modules

1. Network Formation

Research on complex networks has demonstrated that the number of hosts of networks follows the power law. People found that the size distribution usually follows the power law, such as population in cities in a country or personal income in a nation.

2. Malware Propagation

a) Premature stage: A premature stage of the running off a malware means only a small percentage of vulnerable hosts have been concession and the propagation follows exponential distributions.

b) Absolute stage: The absolute stage of the dissemination of malware resources that all susceptible hosts of a given arrangement have been compromised.

c) Late stage: A late stage means the time interval between the early stage and the final stage.

3. Filtering Malware Detection Distribution

Malware detection of coexist multiple malware in networks. In reality, multiple malware may coexist at the same networks. Due to the fact that different malware focus on different vulnerabilities, the distributions of

different malware should not be the same. It is demanding and appealing to establish arithmetical models for multiple malware allotment in terms of networks. The two layers in both layers are sufficiently large and meet the conditions for the modeling methods. In order to improve the accuracy of malware propagation, it may extend our work to layers. In another scenario, it may expect to model a malware distribution for middle size networks.

4. Performance Evaluation

The performance of our experiments indicates that this data does not fit the power law. For a given Android malware program, it only focuses on one or a number of specific vulnerabilities. Therefore, all smartphones share these vulnerabilities form a specific network for that Android malware. Advantage: a. our accurate psychotherapy, find that the allocation of a given malware follows an exponential distribution at its early stage, and obeys a power law distribution with a short exponential tail at its late stage, and finally converges to a power law distribution.

V. RESULT ANALYSIS

Here discuss the results presented in the previous sections and their possible extensions. The main characteristic of our setting is that there are infinite fixed points based on the relationship between the various parameters of the problem. This is in contrast to the finite and small number of fixed points in the case of two viruses. One could erroneously think that having two profiles in the network is like having two viruses but the truth is that the introduced heterogeneity of the underlying network adds complexity to the problem of finding the necessary stability conditions for the fixed points.

Compartment	Probability	Min	Max	Mean
Susceptible	$p = 0.25$	8.008	2000	345.9
	$p = 0.5$	5.379	2000	327.1
	$p = 0.75$	8.41	2000	346.4
	$p = 1$	5.797	2000	318.3
Exposed	$p = 0.25$	2.509	1000	296.6
	$p = 0.5$	2.515	1000	322
	$p = 0.75$	2.504	1000	320.2
	$p = 1$	2.506	1000	349.1
Infected	$p = 0.25$	0.5927	1039	144.5
	$p = 0.5$	1.799	1040	163
	$p = 0.75$	1.212	1038	159.4
	$p = 1$	3.047	1039	175
Recovered	$p = 0.25$	8.63	1027	364.7
	$p = 0.5$	3.497	1000	253.9
	$p = 0.75$	9.491	1075	324.4
	$p = 1$	0.154	1018	196.5

Figure 5.1 Statistical Analysis of virus propagation

5.1 Clique

It has the given conditions so that a virus in presence of two different profiles in the network will die out. It also gave conditions so that a particular number of nodes will get infected from each profile, thus connecting the footprint of the virus in the profiles with the parameters of the profiles with respect to the virus. Of course, it provides such results for particular interesting cases since tackling the general case seems much harder. It is a matter of messy computations to do the same for a barbell graph with uniform weights on the edges between its two cliques. However, adding arbitrary weights on the barbell graph or even on the clique requires a more general approach where the characteristics of the adjacency matrix must be taken into account.

5.2 Arbitrary Graph

In this case, it provides a general condition so that the virus will die out or persist in the network. In case the virus persists; prove conditions that should hold for the graph so that the steady-state infection probability or each node is within some pre- specified range. Although this is a useful result, it is not the whole story.

This is because impose that the probabilities of all nodes should be within this range. As a result, fail to catch the case where most of the nodes are within this range but there are some nodes with probabilities that are outside this range. For instance, visualize a faction Kn and a path Pn of n nodes correspondingly so that the path Pn hangs from some node in Kn creating a graph of 2n nodes in total. It is expected that nodes in Pn will have lower probabilities than those in Kn and thus some of them may be out of the pre-specified range. To tackle these cases one needs to fully solve the respective dynamical system.

5.3 Profiling

Take for example an epidemiological scenario where the virus is the flu. The network specifies the contact between people during a day. It is known that there are groups that are more susceptible to the virus than other groups of people (e.g., children and adults). In this case, one would propose to specify profiles based on the age of nodes (as have done in one of our experiments). An interesting approach is presented in where profiling in a social experiment, shed light in potentially influential users. In a social network scenario, one could also specify the affinity towards a particular rumor or idea (e.g., a PS4 game)by looking at relative historical data of each agent and then decide whether each agent is more susceptible or less susceptible to this particular rumor or idea (or even class of rumors and ideas). However, there is still the problem of giving a value that describes the affinity of each agent. This can either be the choice of the researcher or can be accomplished by using a classifier working on relative historical data, if there is such data of course.

VI. CONCLUSION & FUTURE WORK

6.1 Conclusion

In this research, thoroughly explore the problem of malware distribution at large-scale networks. The solution to this problem is desperately desired by cyber defenders as the network security community does not yet have solid answers. Dissimilar from previous model methods, propose a two layer endemic model: the higher layer focuses on network of a large scale set of connections, for example, domains of the Internet; the subordinate layer focuses on the hosts of a specified network. This two layer model get better the accuracy measure up to with the available single layer endemic models in malware model. Moreover, the proposed two layer representation offers us the distribution of malware in terms of the low layer networks.

It carry out a restricted analysis based on the projected model, and obtain three termination: The distribution for a given malware in terms of set of connections follows exponential allocation, power law distribution with a short exponential appendage, and power law giving out, at its early, late, and final stage, respectively. In order to scrutinize our theoretical findings, have demeanor extensive experimentation based on two real-world large-scale malware, and the consequences confirm our hypothetical claim.

6.2 Future Enhancement

Future work involves a much superior scale of experimentation to quantitatively evaluate the efficiency and efficiency of the proposed social media analytics model. Moreover, alternative latent text mining methods such probabilistic latent semantic indexing and sequential pattern mining will be examined and compared with the performance of the LDA-based method. Also, the application of the mined cybercrime related high-level features to cyber-attack prediction and prevention will be studied.

REFERENCES

[1] M. Fossi and J. Blackbird, "Symantec internet security threat report 2013," Symantec Corporation, Tech. Rep., April, 2014.

[2] C. C. Zou, D. Towsley, and W. Gong, "Modeling and simulation study of the propagation and defense of internet e-mail worms," Dependable and Secure Computing, IEEE Transactions on, vol. 4, no. 2, pp. 105–118, 2007.

[3] Z. Chen and C. Ji, "Spatial-temporal modeling of malware propagation in networks," Neural Networks, IEEE Transactions on, vol. 16, no. 5, pp. 1291–1303, 2005.

[4] S. Wen, W. Zhou, J. Zhang, Y. Xiang, W. Zhou, and W. Jia, "Modeling propagation dynamics of social network worms," Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 8, pp. 1633–1643, 2013.

[5] Y. Cao, V. Yegneswaran, P. A. Porras, and Y. Chen, "Pathcutter: Severing the self-propagation path of xss javascript worms in social web networks." in NDSS, 2012.

[6] M. R. Faghani and U. T. Nguyen, "A study of xss worm propagation and detection mechanisms in online social networks," Information Forensics and Security, IEEE Transactions on, vol. 8, no. 11, pp. 1815–1826, 2013.

[7] C. Song, T. Koren, P. Wang, and A.-L. Barab´asi, "Modelling the scaling properties of human mobility," Nature Physics, vol. 6, no. 10, pp. 818–823, 2010.

[8] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. ACM, 2007, pp. 29–42.

[9] M. E. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," Physical Review E, vol. 66, no. 3, p. 035101, 2002.

[10] Y.-Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong, "Analysis of topological characteristics of huge online social networking services," in Proceedings of the 16th international conference on World Wide Web. ACM, 2007, pp. 835–844.

[11] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics in finite size scale-free networks," Physical Review E, vol. 65, no. 3, p. 035108, 2002.

[12] M. Bogun´a, R. Pastor-Satorras, and A. Vespignani, "Epidemic spreading in complex networks with degree correlations," in Proceedings of the XVIII Sitges Conference on Statistical Mechanics, Lecture Notes in Physics, Springer, Berlin, 2003.

[13] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on. IEEE, 1991, pp. 343–359.

[14] D. Chakrabarti, J. Leskovec, C. Faloutsos, S. Madden, C. Guestrin, and M. Faloutsos, "Information survival threshold in sensor and p2p networks," in INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE. IEEE, 2007, pp. 1316–1324.

[15] G. Yan, G. Chen, S. Eidenbenz, and N. Li, "Malware propagation in online social networks: nature, dynamics, and defense implications," in Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security. ACM, 2011, pp. 196–206.

