IEEE Int. Conference Neural Networks & Signal Processing
Zhenjiang, China, June 8~10, 2008

MODELING MALWARE PROPAGATION IN WIRELESS SENSOR NETWORKS USING CELLULAR AUTOMATA

Yurong Song, Guo-Ping Jiang
Center for Control and Intelligence Technology
Nanjing University of Posts and Telecommunications
Nanjing, Jiangsu,210003,China
songyr@njupt.edu.cn , jianggp@njupt.edu.cn

ABSTRACT

A model based on cellular automata is proposed to investigate and analyze the process of malware propagation over wireless sensor networks by way of multi-hop broadcast protocols. The model captures the inherent characteristics of wireless sensor networks, i.e. limited resource, channel contention (MAC mechanism), and also reflects the self-organization, spatio-temporal correlation of process of malware propagation. The simulation results show that MAC mechanism and density of sensor node greatly influence the malware propagation over wireless sensor network, and the malware propagation diffuses continuously from infected nodes toward outside which is spatially bounded.

Key Words —— wireless sensor networks (WSN); malware propagation; cellular automata; multi-hop broadcast; energy consumption

1. INTRODUCTION

Wireless sensor networks(WSNs) have been widely used for many interesting and new applications such as environmental monitoring, patient health care monitoring, detection of chemical or biological threats, and military surveillance, tracking and targeting[1]. One key issue in wireless sensor networks which are highly distributed and resource constrained environments is security treat[2, 3], where a sensor node can be attacked by malwares such as virus, worms and Trojan. An appropriate model is necessary to characterize an evaluate propagation of malwares over wireless sensor networks.

Worm and virus attacks on the Internet have been widely studied[4-8]. Although many models of malware propagation have been proposed for Internet, they are not well suited to the unique features and application requirements of sensor networks. Wireless sensor networks differ from traditional computer networks in various aspects: First, WSNs are highly distributed system and consist of a great number of distributed nodes (sensor nodes) with the ability to monitor its surroundings. Second, sensor nodes are limited in power, computational capacities, and memory[1]. Finally, minimal (or no)human interaction for the sensors and self-organization is a fundamental feature of wireless sensor networks[9]. Generally, the investigation of malware spreading over WSNs is in its initial stages. Nekovee et al. [10] develop a new model for epidemic spreading of these worms and investigate their spreading in WiFi-based wireless ad hoc networks via extensive Monte Carlo simulations. They incorporate the spatial topology of these networks via a RGG model and also consider the impact of the medium access control (MAC).De et al. [11]consider the distance and pairwise key restricted communication pattern in wireless sensor networks and propose an epidemiological model to investigate the probability of a breakout (compromise of the entire network). Khayam and Radha [12] apply signal processing techniques to model space-time propagation dynamics of topologically-aware worms in a WSN with uniformly distributed nodes. They integrate physical, data link, network and transport protocol characteristics into the proposed model of worm propagation. They focus on the propagation dynamics of unknown worms. Namely, recovery mechanism is paid no attention in the model.

Cellular automata (CA) is a mathematical model for complex natural systems[13-16], containing large numbers of simple identical components with local interactions. CA is a discrete dynamical system with simple construction but complex self-organizing behavior and has been an efficient method to investigate the evolving rules of self-organizing system. We believe that CA can be appropriate model to simulate malware propagation in WSNs.

In this paper, based on epidemic theory, the process of malware propagation in wireless sensor networks is modeled and analyzed using cellular automata. The model focuses on capturing the inherent characteristics of wireless sensor networks to investigate malware propagation. The model allows us to conveniently fit parameters of different scenarios of network. The simulation results show malware propagation

* This work was supported by the Program for New Century Excellent Talents in University of China, under the Contract NCET-06-0510

978-1-4244-2311-8/08/$25.00 ©2008 IEEE.

over wireless sensor network are greatly different from that over Internet.

The paper is organized as follows. Next Section, the Cellular Automata is introduced briefly. In Section 3, we propose our methodology and model. A few of key parameters and assumption are considered in Section 4. In section 5, the simulation results and the related discussions are presented. Finally, we give our conclusions and future work.

2. DEFINING THE CELLULAR AUTOMATA

A 2-D cellular automata is a discrete dynamical system formed by a finite number of $l \times r$ identical objects called cells which are arranged uniformly in a two-dimensional cellular space. Each cell is endowed with a state (from a finite state set Q), that changes at every step of time accordingly to a local transition rule.

In this sense, the state of a particular cell at time t depends on the states of a set of cells, called its neighborhood, at the previous time step t-1. More precisely, a CA is defined by the 4-uplet (C,Q,V, f), where:

$$C = \{(i, j), 1 \le i \le l, 1 \le j \le r\}; \tag{1}$$

Q is the finite state set whose elements are the all possible states of the cells; $V = \{(x_k, y_k), 1 \le k \le n\} \subset Z \times Z$, is the finite set of indices defining the neighborhood of each cell, such that the neighborhood of cell(i,j) is

$$V_{ij} = \{(i + x_1, j + y_1), ..., (i + x_n, j + y_n)\}; \tag{2}$$

Finally, the function f is the local transition function:

$$s_{ij}^t = f(s_{i+x_1, j+y_1}^{t-1}, ..., s_{i+x_n, j+y_n}^{t-1}) \subset Q, \tag{3}$$

where s_{ij}^t denotes the state of cell(i,j) at time t.

3. MODELING THE PROBLEM

The proposed models focus on the stochastic properties of malware propagation and the intrinsic characteristic of wireless sensor networks. We utilize cellular automata to describe the proposed model.

We consider a flat WSN composed of the maximum N stationary and identical sensors which are randomly placed on rectangular 2-D grid composed of L*L cells. We assume that each cell is occupied by at most one sensor node. Thus, we can make simulations with fewer nodes than the maximum number of cells. Let ρ denote the density of sensor nodes, $\rho = N / L^2$. So, the infrastructure of the flat WSN constructs the cellular space and a sensor node denotes a cell in the space. Each sensor node can establish wireless links with only those nodes within a circle of radius r due to the limited power of sensor nodes. To simplify analysis, we assume that all sensor nodes are equipped with omnidirectional antennas that have a maximum transmission range *r*. The horizontal and vertical coordinates of a sensor node are represented by i and j in the 2-D grid(cellular space) respectively. So, node(i,j) is denoted a node located in (i,j)coordinates

According to the corresponding transmission range r, we define the neighborhood of each sensor as shown in Figure 1. Without lack of generality, let the length of a cell of grid be 1, if r=1, each node/cell can have no more than 4 nodes as neighbors, namely the neighborhood of Von Neumann, and if r=1.5, each node/cell can have no more than 8 nodes as neighbors, namely the Moore neighborhood. It is obvious that a node should have more neighbors with the value of communication range r increasing.

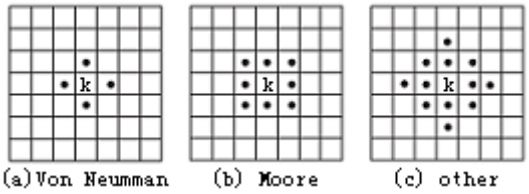

Figure 1. Cell *k* and its possible neighborhoods in a 2-D cellular automata.

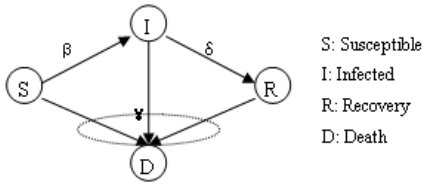

Figure 2. Process of states transforming of sensor nodes

Borrowing the concept of epidemiology, the state of a sensor node can be one of following these states: susceptible, infected, recovery or death.

Let, and denote $\mathbf{s}_{i,j}(t) \in Q$ the state variable of cell/node with coordinates i and j at time t, we define

$$s_{i,j}(t) = \begin{cases} 0, & cell(i,j) \ \ is \ \text{susceptible} \ at \ time \ t \\ 1, & cell(i,j) \ \ is \ \ \text{infected} \ at \ time \ t \\ 2, & cell(i,j) \ \ is \ \ \text{recovery} \ at \ time \ t \\ -1, & cell(i,j) \ \ is \ \ \text{death} \ at \ time \ t \end{cases}. \tag{4}$$

Infected nodes try to spread the malware to their neighbors at each time step. Susceptible sensor nodes become infected with a probability β when they received a packet containing a copy of the malware from an infected neighbor. In addition, infected sensors get patch and recover from infected state with the probability δ . Considering restrained power of sensors and the consumptions during communications among sensor nodes, some sensor nodes can

be dead nodes (nodes with no residual energy on their batteries) at a rate γ .

Furthermore, a cell of the grid that does not contain any node is equivalent to a cell where there is a dead node. The transforming process of states can be shown in Figure 2. S(t), I(t), R(t) and D(t) are denoted the population of susceptible, infected, recovery and death nodes, respectively.

$$\begin{aligned} S(t) &= \frac{1}{N}\sum_{i,j}(s_{ij}(t)=0), \\ I(t) &= \frac{1}{N}\sum_{i,j}(s_{ij}(t)=1), \\ R(t) &= \frac{1}{N}\sum_{i,j}(s_{ij}(t)=2), \\ D(t) &= \frac{1}{N}\sum_{i,j}(s_{ij}(t)=-1), \\ S(t)+I(t)+R(t)+D(t) &= 1 \end{aligned} \tag{5}$$

4. RELATED PARAMETERS AND ASSUMPTION

In the following, propagation parameters are deliberated according to the inherent characteristics of WSN and the spatial-temporal correlation of malware propagation over WSN.

4.1 Infected rate

The infected rate should relate to many factors such as: which type of broadcast protocol being used, MAC mechanism, authentication mechanism for securing data exchanging, attack characteristic of malware and communication pattern. For simplicity, these factors are integrated into a parameter namely, the infected rate β whose value ranges from 0 to 1.

4.2 Death rate

It is well accepted that sensor nodes are severely restricted in terms of computation power and communication capability, especially energy. Malware propagation between nodes results in nodes consuming continuously energy and tending toward death. So, the death rate of nodes is defined by:

$$\gamma_{ij}(t) = c\varepsilon_{ij}(t) \tag{5}$$

where c is a const, $\varepsilon_{ij}(t)$ denotes the cumulative consumption of energy of cell(i,j) until t.

4.3 Routing mechanism

In general, a more robust mechanism for packets routing in wireless sensor networks is by multi-hop broadcasts. Since the transmission power of a wireless radio is attenuated in a squared or even higher order with the distance, multi-hop routing will consume less energy than direct communication. The attackers take advantage of the broadcast mechanism to propagate malicious codes such that malware spreads quickly to the entire network.[17]. We assume that infected nodes adopt broadcasting strategy to spread malware to their neighbors.

4.4 Media access control (MAC)

Malwares over wireless sensor networks will face channel collision, which should in theory reduce the spreading rate of malwares. The MAC protocol specifies a set of rules that enable nearby sensor nodes coordinate their transmissions in a distributed manner[1]. In our model, a MAC table is designed to solve the problem of channel collision. If a sensor node is transmitting a packet, the states of its neighbors should be set block (denoted by '1') in MAC table, which means neighbors can not transmit packets at the same time. Each sensor node checks its state in the MAC table before starting a data transmission. The sensor nodes transmit packets when the channels are idle (denoted by '0' in MAC table). In contrast, the transmission is restrained if the channels are busy.

5. SIMULATION STUDIES

We simulate the malware propagation in a wireless sensor network consisting of N sensors distributing uniformly and randomly in an automata space with $L^2 = 100 \times 100\ unit^2$ cells. $\rho = N / L^2$ denotes the sensor density. The communication range of each sensor is defined r units. Just as presented in proposed model, r=1 denotes each sensor has the Von Neumann neighborhood with 4 neighbors, and r=1.5 denotes each sensor has the Moore neighborhood with 8 neighbors. The spreading process was simulated on above network. Each simulation starts by infecting a few of chosen sensors and evolves following the rules described in section 4.

First we consider the simplest scenario in which the state of a sensor node can be one of both states: susceptible or infected, namely $\delta = 0$ and $\gamma(t) = 0$ in simulations. The scenario aims at discussing the impact of sensor node density on the speed and prevalence of malware propagation over wireless sensor networks. At the same time, we also compare the evolving trend both in the absence and presence of MAC. Each simulation starts from 3 different initial infected nodes with parameters values $\beta = 0.5$ and Moore neighborhood.

Just as shown in Figure.3 (a) and (b), the density of sensor nodes has a great impact on the speed of malware propagation. It can be seen that the infected speed and prevalence show a quick increase with increasing node density. Furthermore, the MAC mechanism is effective in slowing down the malware propagation. Especially, the density is lower the MAC mechanism is more effective.

Next considering immunity and death of nodes, we discuss the propagation dynamics of malware propagation based on the proposed model.

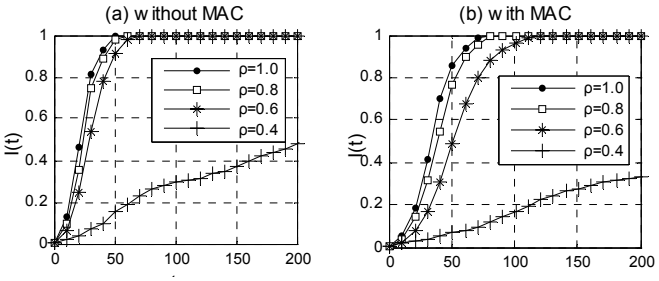

Figure 3. Infected scale under different density

As an example, Figure 4(a)-(d) shows time evolutions of the fraction of nodes with different states under the different scenarios. We set δ =0.001, c=5$\times 10^{-6}$. From Figure 4(a)-(d), we observe that the fraction of nodes that get maximally infected is greatly lowered with a simultaneous recovery procedure and death procedure caused by continuous energy consumption. Comparing (a) and (b) in Figure 4, the spreading speed under β =0.8 is faster than that under β =0.3, and the peak of the curve of I(t) under β =0.8 is higher than that under β =0.3. Simultaneously, the fraction of recovery changes with the fraction of infected. However, there is little change in the fraction of death from β =0.3 to β =0.8. When changing the communication range r, i.e. comparing Moore neighborhood and Nov Neumman neighborhood form Figure 4(a) and (c), the speed propagation is faster in the former than that in the latter. Under same parameters, lower node density slows down the speed as shown in Figure 4 (a) and (d).

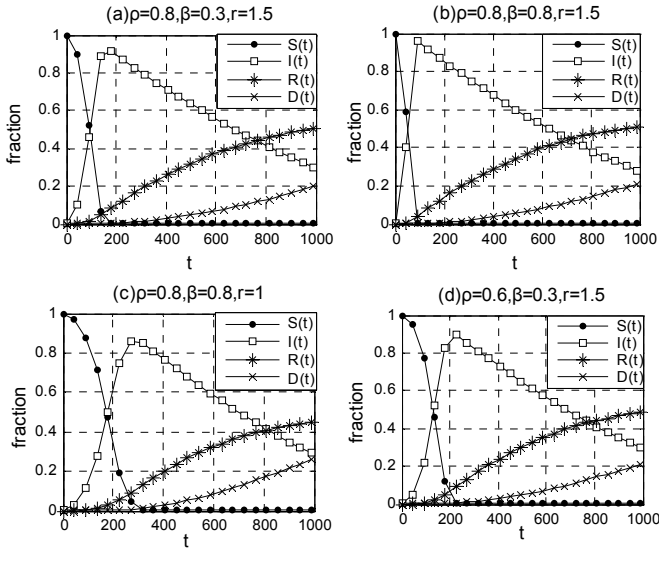

Figure 4. Evolution of propagation

Figure 5 shows the wireless sensor network evolving pattern in the 2-D cellular space with ρ =0.8, β =0.3, δ =0.001, c=5 $\times 10^{-6}$, r=1 under different time t. From the evolution snapshots (Figure 5 (a)-(d)), the epidemic diffuses continuously from infected source toward outside and the propagation happens along a circular front which is spatially bounded. It is means that an infected node in inside of the circular has no chance to infect other susceptible nodes that lie outside of the circular. The characteristic of propagation results in slowing the speed of propagation, which is greatly different to the propagation over Internet.

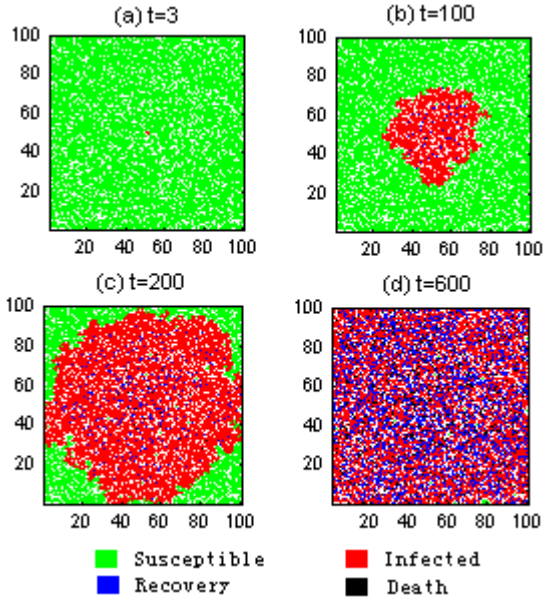

Figure 5. Evolution pattern in the 2-D space

6. CONCLUSIONS AND FUTURE WORK

A model based on cellular automata has been proposed to investigate and analyze the process of malware propagation over wireless sensor networks by way of multi-hop broadcast protocols. The model focuses on capturing the inherent characteristics of WSNs and reflects the self-organization, spatio-temporal correlation of process of malware propagation. The model allows us to conveniently fit parameters of different scenarios of network. The simulation results show malware propagation over wireless sensor network are greatly different from that over Internet. The infected speed and prevalence show a quick increase with increasing node density, infected rate and communication range. Furthermore, the MAC mechanism is effective in slowing down the malware propagation. Especially, the density is lower the MAC mechanism is more effective. The malware propagation diffuses continuously from infected nodes toward outside and happens along a circular front which is spatially bounded, which slows the spreading speed of malware over the wireless sensor network.

As future work, we are planning to evaluate the infected rate affected by the various factors (e.g. security mechanism, attacking characteristic of malware and routing protocol).

7. REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," Computer Networks, vol. 38, pp. 393-422, 2002.

[2] A. S. K. Pathan, H. W. Lee, and C. S. Hong, "Security in Wireless Sensor Networks: Issues and Challenges," Proceedings of the 8th IEEE ICACT, vol. 2, pp. 1043–1048, 2006.

[3] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," Communications of the ACM, vol. 47, pp. 53-57, 2004.

[4] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," Usenix Security, 2002.

[5] C. C. Zou, D. Towsley, and W. B. Gong, "Modeling and simulation study of the propagation and defense of internet e-mail worms," IEEE Transactions on Dependable and Secure Computing, vol. 4, pp. 105-118, 2007.

[6] C. C. Zou, W. Gong, and D. Towsley, "Code red worm propagation modeling and analysis," presented at Proceedings of the 9th ACM Conference on Computer and Communications Security, 2002.

[7] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," Physical Review Letters, vol. 86, pp. 3200-3203, 2001.

[8] M. E. J. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," Physical Review E, vol. 66, pp. 35101, 2002.

[9] K. L. Mills, "A Brief Survey of Self-Organization in Wireless Sensor Networks," Wireless Communications and Mobile Computing, vol. 7, pp. 823-834, 2007.

[10] M. Nekovee, "Worm epidemics in wireless ad hoc networks," New Journal of Physics, vol. 9, pp. 189, 2007.

[11] P. De, Y. Liu, and S. K. Das, "Modeling Node Compromise Spread in Wireless Sensor Networks Using Epidemic Theory," Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks, pp. 237-243, 2006.

[12] S. A. Khayam and H. Radha, "Using signal processing techniques to model worm propagation over wireless sensor networks," Signal Processing Magazine, IEEE, vol. 23, pp. 164-169, 2006.

[13] S. H. White, A. M. d. Rey, and G. R. Sanchez, "Modeling epidemics using cellular automata," Applied Mathematics and Computation, vol. 186, pp. 193-202, 2007.

[14] I. G. Georgoudas, G. C. Sirakoulis, and I. Andreadis, "Modelling earthquake activity features using cellular automata," Mathematical and Computer Modelling vol. 46, pp. 124-137, 2007.

[15] L. H. Encinas, S. Hoya White, A. M. Del Rey, and G. Rodriguez Sanchez, "Modelling forest fire spread using hexagonal cellular automata," Applied mathematical modelling, vol. 31, pp. 1213-1227, 2007.

[16] E. Ahmed and A. S. Elgazzar, "On some applications of cellular automata," Physica A, vol. 296, pp. 529-538, 2001.

[17] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," Ad Hoc Networks, vol. 3, pp. 325-349, 2005.