Neural, Parallel, and Scientific Computations, **27**, No. 1 (2019), 23-34 ISSN: 1061-5369

SOME NEW APPROACHES FOR MODELLING LARGE–SCALE WORM SPREADING ON THE INTERNET. II

ANTON ILIEV[1], NIKOLAY KYURKCHIEV[2],
ASEN RAHNEV[3], AND TODORKA TERZIEVA[4]

[1,2,3,4]Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In this paper we receive models that in some situations can be applied to the theory of computer viruses propagation. As the authors in [3] mention: "It is very hard to use some real-world worm traffic traces or realistic parameters for research. Even traffic traces used in research papers (e.g. Slammer [4] and Code-red [5]) are not public. From the published papers [4], [5] we are not able to find parameters that can be used in our model.". Many researchers make a great efforts to describe adequately situation connected to worm propagation [14]–[58].

AMS Subject Classification: 97N50
Key Words: Verhulst logistic model, power law logistic model, Hausdorff distance, cooperative distribution of traffic filtering policies, IDS rules

Received: September 29, 2018 ; **Accepted:** February 16, 2019 ;
Published: February 22, 2019. **doi:** 10.12732/npsc.v27i1.3
Dynamic Publishers, Inc., Acad. Publishers, Ltd. https://acadsol.eu/npsc

1. COOPERATIVE DISTRIBUTION OF TRAFFIC FILTERING POLICIES

The epidemic infection models that can be used to study the propagations of the worms as it infects Internet hosts are divided into two directions: discrete stochastic (time-stepped) and deterministic (using differential equations).

For large enough populations it is common to approximate the stochastic model by the better continuous state continuous time deterministic model.

Very often deterministic infection Kermack–Mckendrick SIR model for worm propagation dynamics is used.

This model assumes that individuals in a populations of N hosts fall into three classes, susceptible individuals $S(t)$, infected individuals $I(t)$, and removed individuals $R(t)$ who were infected but have recovered (the recovery, or removal rate being γ).

Infections occur randomly and homogeneously transmission coefficient β, that is, the pairwise rate of infection.

The epidemic process is described by the following differential equations:

$$\begin{aligned}\frac{\partial S}{\partial t} &= -\beta IS,\\ \frac{\partial I}{\partial t} &= \beta IS - \gamma R,\\ \frac{\partial R}{\partial t} &= \gamma I.\end{aligned} \tag{1}$$

We will continue the work in [1] we will study the filter propagation via the Border Gateway Protocol (BGP).

If the population is split into K separate strata, each corresponding to a single autonomous system (AS), the same state variables as above can be defined for each stratum, with a subscript notation, i.e.,

$$N = \sum_{k=1}^{K} N_k = \sum_{k=1}^{K} (S_k + I_k + R_k).$$

In the stratified epidemic model, the equation for the susceptibles in stratum k is

$$\frac{\partial S_k}{\partial t} = -S_k \sum_{j=1}^{K} \beta_{kj} I_j \tag{2}$$

and the other equations are derived similarly. Now, from the point of view of filtering, we need to consider two things.

First, interactions within a stratum (AS) are unconstrained and only depend on the worm scanning capabilities, whereas interactions across strata also depend on filtering rules in place at the gateways.

We can assume that the typical filtering rules are precise about the characteristics of packets. Using approach given by [1] we set $\beta_{kj} \equiv \beta\mu_{kj}$, with $\mu_{kk} = 1$ for $k = 1, 2, ..., K$ and denote $\mu_k \equiv \mu_{kj}$, $k \neq j$ the permeability of the gateways in the kth AS with respect to scans. The equations thus become

$$\begin{aligned}\frac{\partial S_k}{\partial t} &= -\beta S_k(I_k + \mu_k(I - I_k)),\\ \frac{\partial I_k}{\partial t} &= \beta S_k(I_k + \mu_k(I - I_k)) - \gamma I_k,\\ \frac{\partial R_k}{\partial t} &= \gamma I_k.\end{aligned} \tag{3}$$

We note that the separation of infection rate and permeability enables us to provide a rough approximate model of the local preference scan pattern, such as those observed in Code Red II, Nimda and Blaster worms, by appropriately choosing initial values for the permeability.

Dynamically filtering is as desirable as dynamic routing. The authors in [1] conceived the notion of a Dynamically Distributed Traffic Filter (DTF). A DTF contains indication of a network activity that should be blocked.

The so-called "stratified epidemic model" gives good results, but numerical analysis is very difficult.

The conducted serious research, connected to the data analysis which are object to the explorations in this paper and the possibility of their good approximation with over fifty "sigmoidal functions" show that there are models which are preferable in comparison to seemingly much more sophisticated models, as an example stratified epidemic model and its modifications. We will give a short look at such one model.

The following modified form of the Verhulst logistic model is called *a power law logistic model*, see Banks [6]:

$$\frac{dM}{dt} = kM\left(1 - \left(\frac{M}{m}\right)^{\theta}\right). \tag{4}$$

Integrating (2.4) with initial condition $M(0) = m_0$ we have

$$M(t) = m\left(\frac{1}{1 + \left(\left(\frac{m}{m_0}\right)^{\theta} - 1\right)e^{-k\theta t}}\right)^{\frac{1}{\theta}}. \tag{5}$$

The logistic function (5) finds applications in many scientific fields, including population dynamics, bacterial growth, population ecology, plant biology, chemistry and statistics.

Let

$$h_{t_0}(t) = \begin{cases} 0, & \text{if} \quad t < t_0, \\ [0,1], & \text{if} \quad t = t_0, \\ 1, & \text{if} \quad t > t_0, \end{cases} \tag{6}$$

is the *Heaviside function* for

$$t_0 = \frac{1}{k\theta}\ln\left(\frac{1}{\theta}\left(\left(\frac{m}{m_0}\right)^{\theta} - 1\right)\right).$$

Let

$$\begin{aligned} p &= -1 + \frac{m}{(1+\theta)^{\frac{1}{\theta}}}, \\ q &= 1 + \frac{km\theta}{(1+\theta)^{\frac{1+\theta}{\theta}}}, \\ r &= qm^{-1}(1+\theta)^{\frac{1}{\theta}}. \end{aligned}$$

For some conditions for q, m, for the one–sided Hausdorff distance d [8] between $h_{t_0}(t)$ and the sigmoid (5) the following inequalities hold [7]:

$$d_l = \frac{1}{r} < d < \frac{\ln r}{r} = d_r. \quad (7)$$

The estimates for the value of the Hausdorff approximation is reliable when assessing the important characteristic - "saturation".

We use the following model:

$$M^*(t) = \omega \left(\frac{1}{1 + \left(\left(\frac{1}{x_0} \right)^{\theta} - 1 \right) e^{-k\theta t}} \right)^{\frac{1}{\theta}}.$$

For contemporary applied research on sigmoids and some of their applications see the monographs [9]–[13].

So, we will study how we can effectively approximate propagation of SQL slammer the worm infection dynamics with and without DTF form [1], see Fig. 1 where it can be seen that the exponential growth in the early propagation stages will be evidently smoothed after the DTF application (after 10 and 15 sec., see Fig. 2 and Fig. 3 respectively).

2. PERCENTAGE OF TRAFFIC EXPLAINED BY AUTOMATICALLY GENERATED IDS RULES IN EACH ITERATION

In the paper [2] the authors proposed a novel framework for automatically discovering and analyzing of traffic generated by computer worms and other anomalous behaviors that interact with a non-solicited traffic monitoring system.

Network packets are analyzed by an Intrusion Detection System (IDS), and new signatures are generated clustering those which remain unknown for the IDS.

Furthermore, the framework provides a mechanism to cluster the alarms produced by the IDS producing a correlated vision of the traffic observed.

Both the automatic signature generation and the alarm clusters are accomplished using data mining techniques.

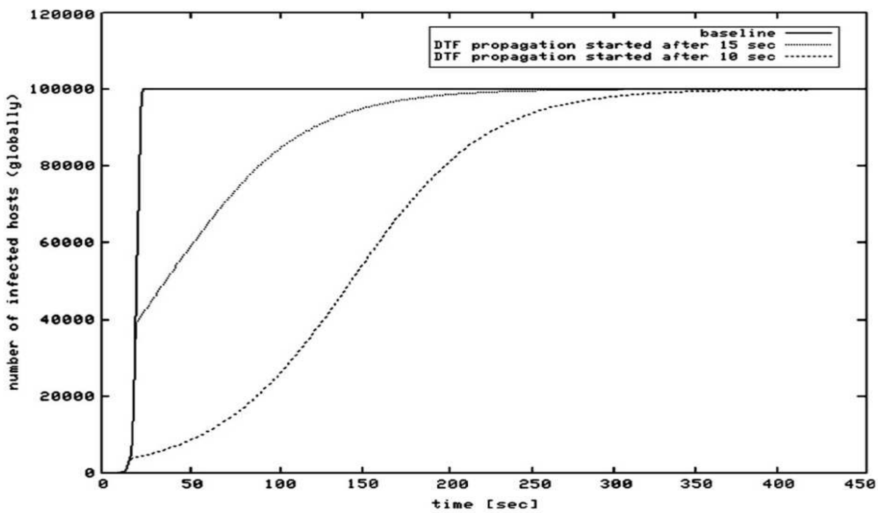

Figure 1: SQL slammer – the worm infection dynamics with and without DTF [1].

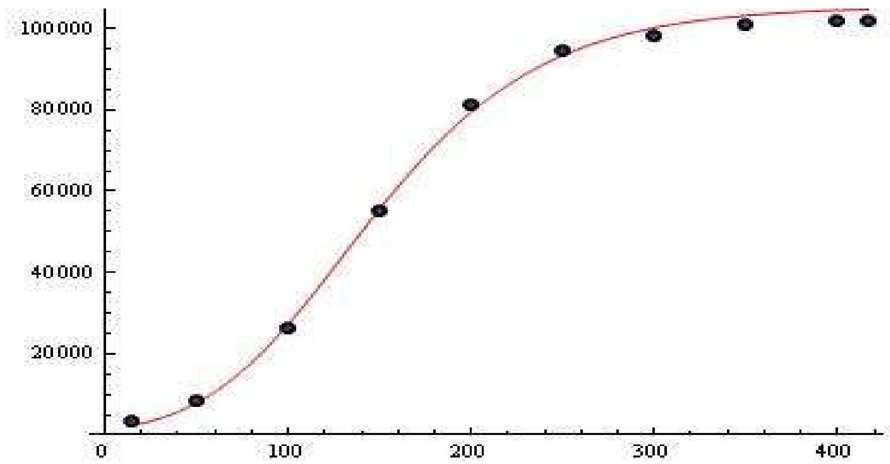

Figure 2: The model $M^*(t)$ for $\omega = 105450$; $x_0 = 0.0125657$; $k = 0.0529382$; $\theta = 0.330662$.

The framework [2] relies on four components (see Fig. 4):

1) a Worm Detection System (WDS) responsible for interacting with worm infected machines;

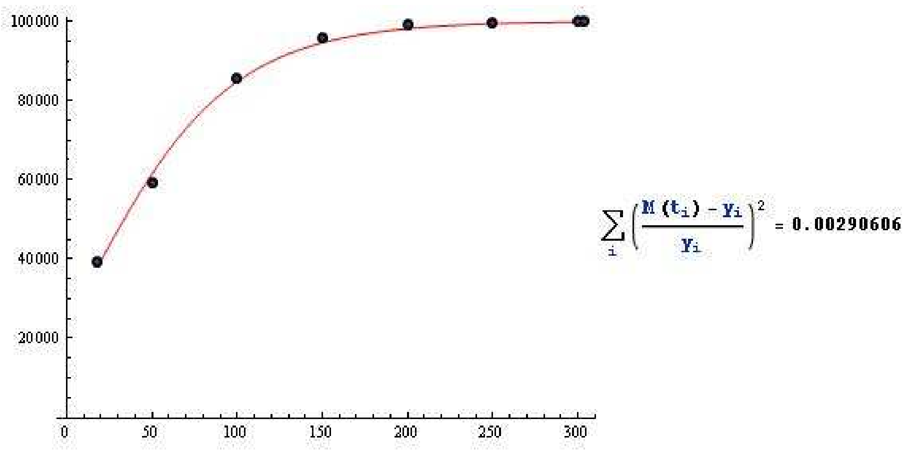

Figure 3: The model $M^*(t)$ for $\omega = 100000$; $x_0 = 0.240324$; $k = 4.79847$; $\theta = 0.004498$.

2) a knowledge-based Intrusion Detection System (IDS) which discerns the data between known and unknown traffic patterns;

3) a data mining tool;

4) an automatic signature generation system.

Results of the knowledge discovery of the unknown traffic dataset [2] show that in only three iterations more than 95% of the data captured by the system can be explained, using 69 new IDS rules.

On the 5th iteration, 99% of the data is explained with 86 signatures.

Figure 5 shows which percentage of the traffic is explained by the automatically-generated IDS rules over the experiment.

The process of IDS rules for recognizing known and unknown traffic patterns iteration is a random value.

It is turns out that this process is well modelled with model $M^*(t)$, see Fig. 6.

The received result and explicit type of this approximate model can be used to control and adequate intervention in pointed out mechanism described in [2].

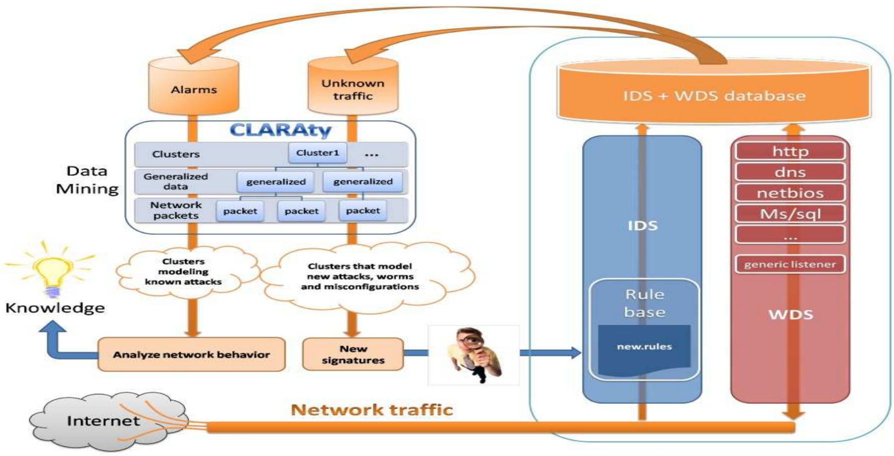

Figure 4: Framework for the analysis of worm activity [2].

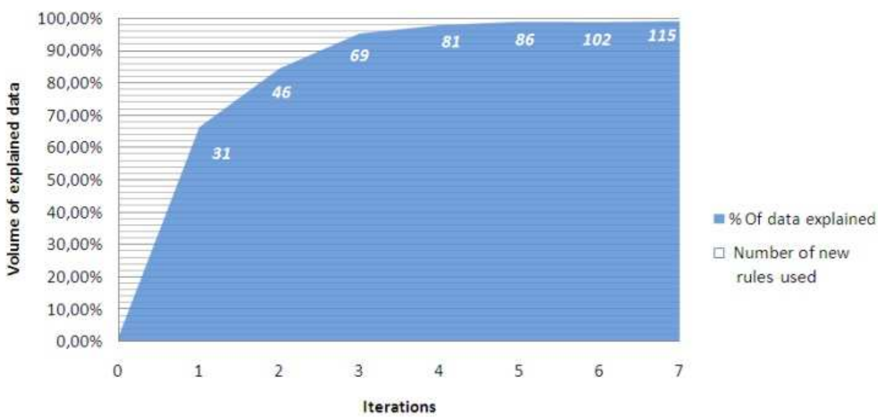

Figure 5: Percentage of traffic explained by automatically generated IDS rules in each iteration [2].

3. ACKNOWLEDGMENTS

This paper is supported by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)", financed by the Ministry of Education and Science.

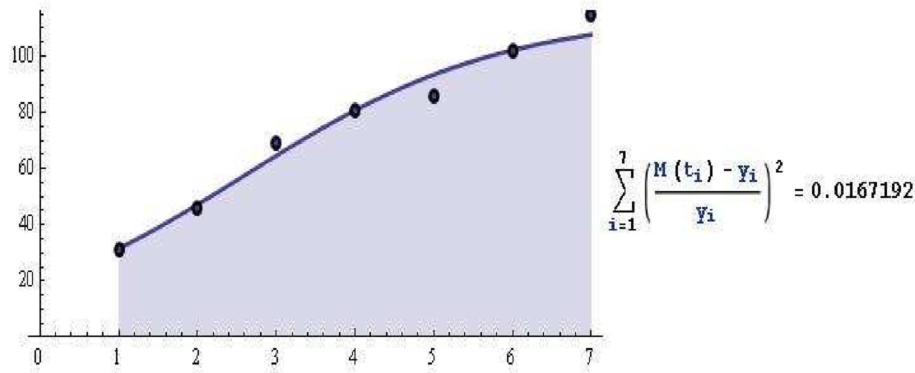

Figure 6: The model $M^*(t)$ for $\omega = 115$; $x_0 = 0.170588$; $k = 0.610897$; $\theta = 0.997403$.

REFERENCES

[1] F. Palmieri, U. Fiore, Containing large-scale worm spreading in the Internet by cooperative distribution of traffic filtering policies, *Computers & Security*, **27** (2008), 48–62.

[2] U. Zurutuza, D. Zamboni, A Data Mining Approach for Analysis of Worm Activity Through Automatic Signature Generation, *AISec'08 Proceedings of the 1st ACM workshop on Workshop on AISec*, (2008), 61–70.

[3] O. A. Toutonji, S.-M. Yoo, M. Park, Stability analysis of VEISV propagation modeling for network worm attack, *Applied Mathematical Modelling*, **36** (2012), 2751–2761.

[4] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, N. Weaver, Inside the slammer worm, *IEEE Magaz. Secur. Privacy*, **1** No. 4 (2003), 33–39.

[5] C. Shannon, D. Moore, The Spread of the Code-Red Worm, < http : //www.caida.org/analysis/security/code−red/coderedv2_analysis.xml >.

[6] R. Banks, Growth and Diffusion Phenomena: Mathematical Frameworks and Applications, Springer Verlag, Berlin (1991).

[7] N. Kyurkchiev, A. Iliev, A note on the power law logistic model, Proc. of the NTADES Series of AIP, (2019). (to appear)

[8] F. Hausdorff, *Set theory (2 ed.)*, Chelsea Publ., New York (1962) [1957], ISBN 978-0821838358, Republished by AMS-Chelsea (2005).

[9] N. Kyurkchiev, S. Markov, *Sigmoid functions: Some Approximation and Modelling Aspects*, LAP LAMBERT Academic Publishing, Saarbrucken (2015), ISBN

978-3-659-76045-7.

[10] N. Kyurkchiev, A. Iliev, S. Markov, *Some Techniques for Recurrence Generating of Activation Functions: Some Modeling and Approximation Aspects*, LAP LAMBERT Academic Publishing (2017), ISBN: 978-3-330-33143-3.

[11] N. Kyurkchiev, A. Iliev, *Extension of Gompertz-type Equation in Modern Science: 240 Anniversary of the birth of B. Gompertz*, LAP LAMBERT Academic Publishing (2018), ISBN: 978-613-9-90569-0.

[12] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, *Some software reliability models: Approximation and modeling aspects*, LAP LAMBERT Academic Publishing (2018), ISBN: 978-613-9-82805-0.

[13] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, *Nontrivial Models in Debugging Theory: Part 2*, LAP LAMBERT Academic Publishing (2018), ISBN: 978-613-9-87794-2.

[14] D. Moore, C. Shannon, J. Brown, Code-Red: a case study on the spread and victims of an Internet worm, *Internet Measurement Workshop* (IMW), (2002), 273–284.

[15] C. Zou, W. Gong, D. Towsley, Code red worm propagation modeling and analysis, *CCS '02 Proceedings of the 9th ACM conference on Computer and communications security*, (2002), 138–147.

[16] C. Zou, W. Gong, D. Towsley, Worm propagation modeling and analysis under dynamic quarantine defense, *Proceedings of the 2003 ACM workshop on Rapid malcode*, October 27–27, (2003), Washington, DC, USA.

[17] C. Zou, D. Towsley, W. Gong, On the performance of internet worm scanning strategies, *Performance Evaluation*, **63**, No. 7 (2006), 700–723.

[18] C. Zou, W. Gong, D. Towsley, L. Gao, The monitoring and early detection of internet worms, *IEEE/ACM Transactions on Networking (TON)*, **13**, No. 5 (2005), 961–974.

[19] P. Wang, L. Wu, R. Cunningham, C. Zou, Honeypot detection in advanced botnet attacks, *International Journal of Information and Computer Security*, **4**, No. 1 (2010), 30–51.

[20] A. Visheratin, M. Melnik, D. Nasonov, N. Butakov, A. Boukhanovsky, Hybrid scheduling algorithm in early warning systems, *Future Generation Computer Systems*, **79**, No. P2 (2018), 630–642.

[21] J. Jerkins, J. Stupiansky, Mitigating IoT insecurity with inoculation epidemics, *Proceedings of the ACMSE 2018 Conference*, March 29–31, (2018), 1–6, Richmond, Kentucky.

[22] Q. Xiao, S. Chen, M. Chen, Y. Ling, Hyper-Compact Virtual Estimators for Big Network Data Based on Register Sharing, *ACM SIGMETRICS Performance Evaluation Review*, **43**, No. 1 (2015), 417–428.

[23] H. Asghari, M. Ciere, M. Van Eeten, Post-mortem of a zombie: conficker cleanup after six years, *Proceedings of the 24th USENIX Conference on Security Symposium*, August 12–14, (2015), 1–16, Washington, D.C.

[24] A. Dainotti, A. King, K. Claffy, F. Papale, A. Pescape, Analysis of a "/0" stealth scan from a botnet, *IEEE/ACM Transactions on Networking (TON)*, **23**, No. 2 (2015), 341–354.

[25] D. Lee, J. Kim, K. Kim, A study on abnormal event correlation analysis for convergence security monitor, *Cluster Computing*, **16**, No. 2 (2013), 219–227.

[26] E. Magkos, M. Avlonitis, P. Kotzanikolaou, M. Stefanidakis, Toward early warning against Internet worms based on critical-sized networks, *Security and Communication Networks*, **6**, No. 1 (2013), 78–88.

[27] S. Xu, W. Lu, L. Xu, Push- and pull-based epidemic spreading in networks: Thresholds and deeper insights, *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, **7**, No. 3 (2012), 1–26.

[28] C. Shannon, D. Moore, The Spread of the Witty Worm, *IEEE Security & Privacy*, **July/August**, (2004), 46–50.

[29] A. Mohammed, S. Nor, M. Marsono, Analysis of Internet Malware Propagation Models and Mitigation Strategies, *IRACST International Journal of Computer Networks and Wireless Communications (IJCNWC)*, **2**, No. 1 (2012), 16–20.

[30] S. Staniford, V. Paxsony, N. Weaver, How to own the Internet in Your Spare Time, *Proceedings of the 11th USENIX Security Symposium, San Francisco*, California, USA, August 5-9, (2002).

[31] S. Fei, L. Zhaowen, M. Yan, A survey of Internet Worm Propagation Models, *Proceedings of IC-BNMT2009*, (2009), 453–457.

[32] S. Fei, L. Zhaowen, M. Yan, Worm Propagation based on Two-Factor Model, *Proceedings of 2009 5th International Conference on Wireless Communications, Networking and Mobile Computing*, (2009), 4 pp.

[33] D. Smith, L. Moore, The SIR model for the Spread of Diseases, *JOMA*, (2004).

[34] J. Kim, S. Radhakrishnan, S. Dhall, Measurement and Analysis of worm propagation on Internet network topology, *Proceedings of 13th International Conference on Computer Communications and Networks (IEEE Cat. No.04EX969)*, 495–500.

[35] T. Li, Z. Guan, Y. Wang, The Stability of a Worm Propagation Model with Time Delay on Homogeneous Networks, *Proceedings of International Conference on Intelligent Control and Information Processing*, August 13-15, (2010) - Dalian, China, 753–755.

[36] T. Li, Z.-H. Guan, Y. Wang, Y. Li, Impulsive Control of the Spread of worm with Nonlinear Incidence Rates, *Proceedings of 2010 Chinese Control and Decision Conference*, (2010), 966–969.

[37] Y. Wang, Z.-H. Guan, T. Li, S. Zhang, Modeling and Analyzing the Spread of Worm with Impulsive Effect on Homogeneous Network, *Proceedings of 2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, (2010), V7-501–V7-504.

[38] C. Junhua,W. Shengjun, Modeling and Analyzing the Spread of worms with Bilinear Incidence Rate, *Proceedings of 2009 Fifth International Conference on Information Assurance and Security*, (2009), 167–170.

[39] W. Shaojie, L. Qiming, D. Bo, M. Weining, Analysis of a Mathematical Model for Worm Virus Propagation with time delay, *Proceedings of 2009 Second International Conference on Environmental and Computer Science*, (2009), 375–379.

[40] D. Zhang, Y. Wang, SIRS: Internet Worm Propagation and Application, *Proceedings of 2010 International Conference on Electrical and Control Engineering*, (2010), 3029–3032.

[41] Q. Liu, R. Xu, S. Wang, Modeling and Analysis of an SIRS Model for worm Propagation, *Proceedings of 2009 International Conference on Computational Intelligence and Security*, (2009), 361–365.

[42] S. Fei, L. Zhao-wen, M. Yan, Modeling and Analysis of Internet worm propagation, *The Journal of China Universities of Posts and Telecommunications*, **17**, No. 4 (2010), 63–68.

[43] J. Wang, C. Xia, Q. Liu, A novel Model for the Internet Worm Propagation, *Proceedings of 2010 Sixth International Conference on Natural Computation (ICNC 2010)*, (2010), 2885–2888.

[44] F. Wang, J. Song, Y. Dong, J. Gu, Epidemic models applied to worms on internet, *Proceedings of 2009 Second International Conference on Intelligent Networks and Intelligent Systems*, (2009), 160–163.

[45] Z. Wei, Q. Facheng, C. Shiqi, W. Ruchuan, The Study of Network Worm Propagation Simulation, *Proceedings of 2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, (2010), V9-295–V9-299.

[46] M. Liuqi, The research and development of worm defense strategies, *Proceedings of 2010 3rd International Conference on Computer Science and Information*

Technology, (2010), 168–171.

[47] F.Wang, Y. Zhang, C. Wang, J. Ma, S. Moon, Stability analysis of a SEIQV epidemic for rapid spreading worms, *Computer & Security*, **29** (2010), 410–418.

[48] Y. Yao, H. Guo, F. Gao, G. Yu, The Worm Propagation Model with pulse Quarantine Strategy, *Proceedings of 2010 International Conference on Multimedia Information Networking and Security*, (2010), 269–273.

[49] H. Zhang, W. Su, W. Quan, *Smart Collaborative Identifier Network: A Promising Design for Future Internet*, Springer-Verlag, Berlin (2016).

[50] X. Wang, J. Zhu, H. Lin, X. Su, Y. Jiang, Modeling Propagation of Active P2P Worm in Chord Network, In: *Advances in Intelligent and Soft Computing*, J. Kacprzyk eds., **133** (2012), S. Sambath & E. Zhu (Eds.), *Frontiers in Computer Education*, 383–390.

[51] Y. Xiao, F. Li, H. Chen, eds., *Handbook of Security and Networks*, World Scientific, Singapore (2011).

[52] S. Sellke, N. Shroff, S. Bagchi, Modeling and Automated Containment of Worms, *Proceedings of the 2005 International Conference on Dependable Systems and Networks (DSN05)*, (2005), 10 pp.

[53] W. Yu, C. Boyer, S. Chellappan, D. Xuan, Peer-to-peer system-based active worm attacks: modeling and analysis, *IEEE International Conference on Communications, 2005*, (2005), 295–300.

[54] S. Zhang,Z. Jin, J. Zhang, The Dynamical Modeling Analysis of the Spreading of Passive Worms in P2P Networks, *Discrete Dynamics in Nature and Society*, **2018**, Article ID 1656907, (2018), 13 pp.

[55] G. Yan, S. Eidenbenz, Modeling Propagation Dynamics of Bluetooth Worms (Extended Version), *IEEE Transactions on Mobile Computing*, **8**, No. 3 (2009), 353–367.

[56] S. Sellke, N. Shroff, S. Bagchi, Modeling and Automated Containment of Worms, *IEEE Transactions on Dependable and Secure Computing*, **5**, No. 2 (2008), 71–86.

[57] S. Peng, M. Wua, G. Wang, S. Yu, Propagation Model of Smartphone Worms Based on Semi-Markov Process and Social Relationship Graph, *Computers & Security*, **44** (2014), 92–103.

[58] N. Kyurkchiev, A. Iliev, A. Rahnev, T. Terzieva, A new analysis of Code Red and Witty worms behavior, Communications in Applied Analysis, 23 (2), 2019, 267–285.