Mathematical and numerical study of a SIR epidemic model on network with standard incidents

Lukas Podesva[a,*], Milos Koch[a]

[a] *Brno University of Technology, Faculty of Business and Management, Kolejni 2906/4, Brno 61200, Czech Republic*

Abstract

Purpose of the article With the expansion and use of computers and the Internet in everyday life, there has also been an increase in their abuse. The system connection by the Internet increases vulnerability of the connected networks and computers as well as the danger of information abuse or loss, e.g. through computer viruses. As the operating systems are becoming more complicated and complex, it is a bigger challenge for virus authors to attack these systems. In this paper, we formulate a model of computer virus spreading, inspired by the SIR dynamic epidemic model.

Methodology/methods The theory is briefly explained in the opening part and serves as a basis for formulation of the relationships between the quantities investigated in the paper. The results are demonstrated on particular examples and the behaviour of the model is presented using computer simulation. The solution incorporates the theory of mathematical analysis and ordinary differential equations.

Scientific aim The authors' aim is to analyse the computer virus spread model as a system of non-linear differential equations and verify its solvability.

Findings The proposed model is based on a system of non-linear differential equations and allows a qualitative view of its behaviour, simulated by Maple and graphically presented in the application part, for various input values.

Conclusions These days, all users should be familiar with computer security and data protection. The users should know how to protect and secure their data. This requires the knowledge of behaviour of various types of harmful software. Thus the model we have designed may significantly contribute to formulation of a defence strategy against computer virus spreading.

Keywords: computer virus, SIR, epidemic model, network, simulation

JEL Classification: C02, C60, D85

[*] Corresponding author. Tel.:+420 541141 111.
E-mail address: lukas.podesva@vutbr.cz.

Introduction

In today's society, computers have become an integral part of life. As computers and the Internet are used more extensively in everyday life, there has also been an increase in their abues, in particular in the form of a risk of device infection by a harmful code. We use the computer today in nearly all activities – in communication, contract conclusion, etc. Therefore we must be aware of the importance of computer safety and data protection. Users these days are threatened by ever new infiltrations and the attackers' resourcefulness is infinite. For this reason, computer security should not be the sole interest of companies but also of common users who should protect their data susceptible to abuse.

The computer viruses have reached the current form by a long process of evolution. New trends have been added in the course of the development, adjusting to new technologies and types of operating systems. In this paper, we formulate a model of computer virus spreading, inspired by the SIR dynamic epidemic model. The model is described by a system of non-linear differential equations and allows a qualitative view of its behaviour, simulated by Maple and graphically presented in the application part, for various input values. The au-thors' aim is to analyse the computer virus spread model as a system of non-linear differential equations and verify its solvability.

1 Literature Review

Computer viruses are harmful codes or programs that may replicate and spread through cable or wireless connection. With an increasing number of internet applications, computer viruses are a big threat to our work and everyday life. With the introduction of the Internet of Things and 5G, the threat is becoming more and more serious. As a result, it is important to understand how computer viruses spread on the Internet and to propose effective measures to solve this problem. To achieve this goal, it is advisable to get inspiration from biological virus spread in the population, considering their similarity to computer virus spreading (Kephart et al, 1993).

Mahtematical biological epidemic models have been studied for centuries and already in the mid-18th century, David Bernoulli developed one of the first known models inspired by the smallpox virus (Bernoulli, 1760). Infectiousness is one of the main common features shared by both computer and biological viruses (Cohen et al, 1987).

Based on these facts, some traditional epidemic models have been applied, such as SIRS, SEIR, SEIRS, SEIQV a SEIQRS. They display the computer virus spreading as follows: (1) Susceptible persons correspond to non-infected computers, (2) latent patients correspond to infected computers in which all the viruses are latent, and (3) infecting patients correspond to infected computers containing at least one virus. In biological environment, it is well known that an infected latent person cannot infect other individuals. However, in the computer environment, a latent infected computer can infect other computers, e.g. through file copies or downloads. Un-fortunately, all previous computer virus models have not taken this passive infectiousness into account. A reasonable virus computer model should assume that both defective and latent computers are infectious (Yang et al, 2013).

The development and analysis of mathematical models of these systems can allow understanding of virus spreading behaviour, which can bring long-term benefits for the society as a whole. Several traditional epidemic models used for describing the spreading of harmful files – viruses are described below:

- SIR (susceptible-infected-removed) is strictly based on the biological SIR model. This method is restricted to cases when all community members demonstrate the same level of susceptibility to the dis-ease at the beginning (Kermack et al, 1927).
- A fundamental model on which many other models are based is SIS (susceptible-infected-susceptible). In the basic SIS model, in every step, every single node or agent is either infected or susceptible to infection. Any other agent can be infected by the adjacent agents with a given infection rate β, where the network graph structure determines the connection between the individual agents, and therefore plays a direct role in facilitating or preventing infection spread. An infected agent can be cured with a certain rate of treatment δ, by which it returns to the initial susceptibility state. The first SIS model was developed in the mid-20th century by Kermack and McKendrick (Kermack et al, 1932).
- The SEI model (susceptible, exposed, infected) is an adjusted version of the SEI epidemiology model. This model assumes virus spread through P2P (peer to peer) network and classifies each role as one of three categories based on the number of shared infected files (Thommes et al, 2005).

- SEIR (susceptible, exposed, infected, recovered) is a disease spread model with an exponential demographic structure with a natural mortality constant and excessive mortality rate in infectious individuals (Yan et al, 2006).
- SEIRS (susceptible, exposed, infected, recovered, susceptible) is formulated by transmission of harmful objects within the computer network with a mortality rate different from an attack of a harmful object and a mortality rate constant for infectious nodes. Death of a computer network node also means isolation of the node from the computer network, which carries on spreading harmful objects even if an antivirus software is operating constantly (Mishra et al, 2007).
- By extending the SEIRS model and implementing a new quarantine section, Mishra and Jha described the SEIQRS epidemic model. An undesirable file (such as a virus infected file or another harmful object) is moved by the quarantine to a folder that is not easily accessible by ordinary file ad-ministration tools (Mishra et al, 2010).
- The SEIQV (susceptible, exposed, infected, quarantined, vaccinated) epidemic model combines vaccination and dynamic quarantine methods (Wang et al, 2010).
- SLBS (uninfected computers, latent internal computers, breaking-out computers) is a computer model with a gradual treatment speed. It is a dynamic model characterizing computer virus spreading by the Internet, under two assumptions: (1) the computer has infectiousness as soon as it is infected, and (2) latent computers have lower treatment rate than attacked computers (Yang et al, 2013).

The current development in the area focuses on virus spreading in various conditions, when a large number of derived models have been proposed. These include mostly parameter optimization for the individual model situations and their dynamization.

(Chen et al, 2015) for instance, propose a delayed SLBS computer virus model. Using an optimal management strategy, they present an optimal strategy minimizing the total number of computer failures and costs related to toxication or detoxication. (Hong et al, 2015) proposed that node mobility in heterogeneous networks will affect the epidemic process. (Yun-Peng et al, 2017) proposed a dynamic information transfer model based on social impact elements. (Zhang et al, 2018) propose a new virus spreading model based on partial immunization and immune invalidity in a complex network. The new model is based on a combination of the SIS, SIR and SIRS models included in a partial loop of the model, the model is referred to as SISRS. Virus spreading in dynamic graph structure is discussed by (Pare Philip et al, 2018). The extended the studied SIS models to time-variable graphical structures and considered additional stochastic uncertainties. This extension makes the models more realistic an provides advanced knowledge on virus spreading on a larger scale of settings. (Li et al, 2019) examines threshold dynamics and ergodicity of the SIRS stochastic epidemic model with a disease transmission rate controlled by semi-Markov process. The semi-Markov process is used here for the description of randomly changing environment, which is very extensive.

2 Virus Spread Model

The mathematical models of space-time infectious disease spread help to reveal the causes of the disease origin and most often epidemic spread 1. Epidemiology employs models allowing the infection spread to be monitored during interactions among people present in a space or social network. A frequent goal is to establish the critical parameter values at which epidemics start spreading in the population. Apart from the usual use, the epidemiological models can also be used to model fundamentally close processes, such as fire spreading, plant invasion in unoccupied spaces, food chain dynamics, etc.

Let us start with the following virus dynamic model:

$$\frac{dS(t)}{dt} = \alpha - \beta S(t) - \gamma S(t)I(t)$$
$$\frac{dI(t)}{dt} = \gamma S(t)I(t) - \delta I(t) \quad (1)$$
$$\frac{dR(t)}{dt} = \mu R(t) + \rho I(t)$$

where *S(t), I(t)* and *R(t)* mark healthy cells, infected cells and free virus, constant α represents the level of cell regeneration, β, δ, μ mark the death rate of healthy cells, infected cells and free viruses. *Y* is a parameter of the degree of contact between healthy and infected population. $\frac{\rho}{\delta}$is the average number of viruses produced by one infected cell.

Since the behaviour of computer viruses is very similar to the behaviour of biological viruses, the computer virus spread model may be described as follows:

$$\frac{dH(t)}{dt} = \alpha - \beta H(t) - \gamma H(t)I(t)$$
$$\frac{dI(t)}{dt} = \gamma H(t)I(t) - (\vartheta + \beta)I(t) \quad (2)$$
$$\frac{dC(t)}{dt} = -\beta C(t) + \vartheta I(t)$$

where *H(t)*, *I(t)* and *C(t)* are numbers of computers divided into three groups, at time *t*. *H(t)* is the number of healthy computers threatened by infection, *I(t)* is the number of infected computers and *C(t)* are "clean" computers from which viruses have been removed.

Constant α represents the speed of connection of external computers in the network, β is the speed of one computer removal from the network, υ is the speed of virus removal from one computer, γ is the speed of possible infection of a computer connected to the network.

3 Model Analysis

3.1 Positive Solution

For the model to make sense, all classes must be non-negative for each t≥0.

Clause 1

The solution of the (H, I, C) system (4) is positive for ∀t: t>0

Proof

a) Let us now assume that *H(t)* is not always positive, therefore there are t>0 for which H(t)≤0. Since the starting conditions (5) of the system (4) imply that at the start the solution t=0 is positive, there must be the lowest number t_1>0 that H(t_1)=0. If we substitute this value in the equation $\frac{dH(t)}{dt} = \alpha - \beta I(t) + \gamma I(t)H(t)$ the result is $\frac{dH(t_1)}{dt_1}$. With regard to the continuity *H(t)* there is η>0 that for $t \in (t_1, t_1 + \eta)$ it is H(t)<0. However, then for t_1 it is $\frac{dH(t_1)}{dt_1} = \alpha \leq 0$, which is disputable. Therefore for each t>0 there is H(t)>0.

b) For I(t) it may be derived from the equation $\frac{dI(t)}{dt} = \gamma I(t)H(t) - (\vartheta + \beta)I(t)$ that $I(t) = e^{-(\vartheta+\beta)\upsilon} \times \left(\int_0^u \gamma e^{(\vartheta+\beta)} I(v)H(v)dv\right) > 0$

Thus I(t)>0 for each t>0.

Let there now be such t_2 that $C(t_2)=0$. After insertion in $\frac{dC(t)}{dt} = -\beta H(t) + \vartheta I(t)$, we get $\frac{dC(t_2)}{dt} = \vartheta I(t_2) > 0$. Analogically to (a), it may be derived that for each $t>0$ there is $C(t)>0$. Thus for system (2) it applies that its solution the (H, I, C) system is positive for ∀t: t>0

3.2 Equilibrium

Equilibrium refers to a constant differential equation solution. For our model this means that transfers might occur between the respective computer groups; however, their sizes remain constant over time. In general, there are two model equilibrium types: non-indicative, when the disease/infected computers disappear completely and the number of infected individuals is zero, and endemic, which is a state when the disease remains in the population but does not spread into an epidemic.

Equilibriums for our model can be calculated as:

$$E_0 = \left(\frac{\alpha}{\beta}, 0, 0\right)$$
$$E_1 = \left(\tilde{H} = \frac{\vartheta + \beta}{\vartheta}, \tilde{I} = \frac{\beta}{\vartheta} \times \tilde{C}, \tilde{C} = -\frac{\alpha}{\vartheta \tilde{C}} + \frac{\beta}{\gamma}\right) \quad (3)$$

3.3 Numerical Experiment

Mathematic modelling is currently an integral part of various fields of natural, technical, economic and social sciences and an important instrument for modelling and simulations of systems, analyses and forecasting of their processes, phenomena, behaviour of species and states of societies. The use incorporates present information and communication technologies (ICT), in particular application mathematical software (e.g. Maple, Matlab). In this chapter, we will simulate various situations that can occur in computer networks and using suitable software, we will monitor the behaviour of the system in time.

3.3.1 Software Used

For calculation realization, we have selected Maple, which is a computer environment developed at the University of Waterloo in Canada, suitable for solving complex mathematical problems. It belongs to a group of computer algebra systems, allows solving of problems from various fields of mathematics from basics of linear algebra and mathematical analysis, differential and integral count to differential equations, geometry to logic. The system is aimed primarily for symbolic operations in mathematics, numerical calculations and graph displays. For the solution of ordinary differential equations, we will use the dsolve command and several related commands.

3.2.2 Example 1

A situation when a large portion of computers in the computer network have viruses. However, the network parameters reflect a situation when the network should gradually stabilize at E_0. As we can see in Fig. 1, the system positively converges to equilibrium.

Parameters: α=20; β=4; γ=1,2; d=2

Initial function values: H=20; I=10; C=1

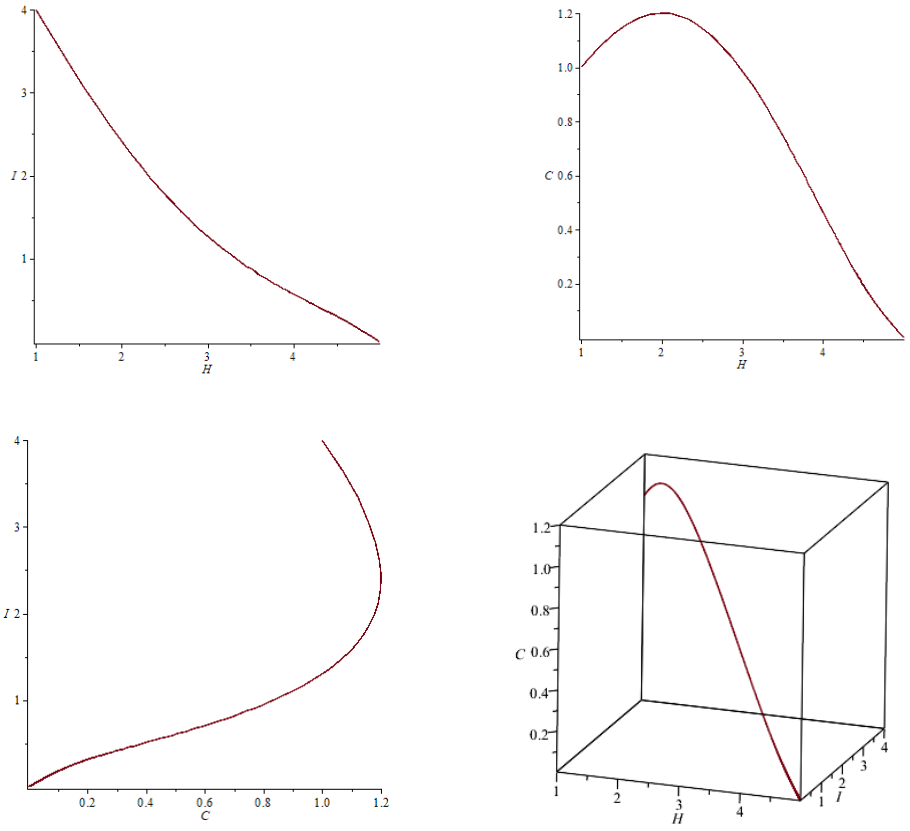

Figure 1 Example 1

3.3.3 Example 2

A situation when the majority of computers in the network are "healthy", only one computer is infected. The network parameters reflect a situation when the network should gradually stabilize at E_1. As we can see in Fig. 2, the system positively converges to equilibrium, as expected.

Parameters: α=4; β=0,3; γ=1,2; d=12

Initial function values: H=12; I=1; C=0

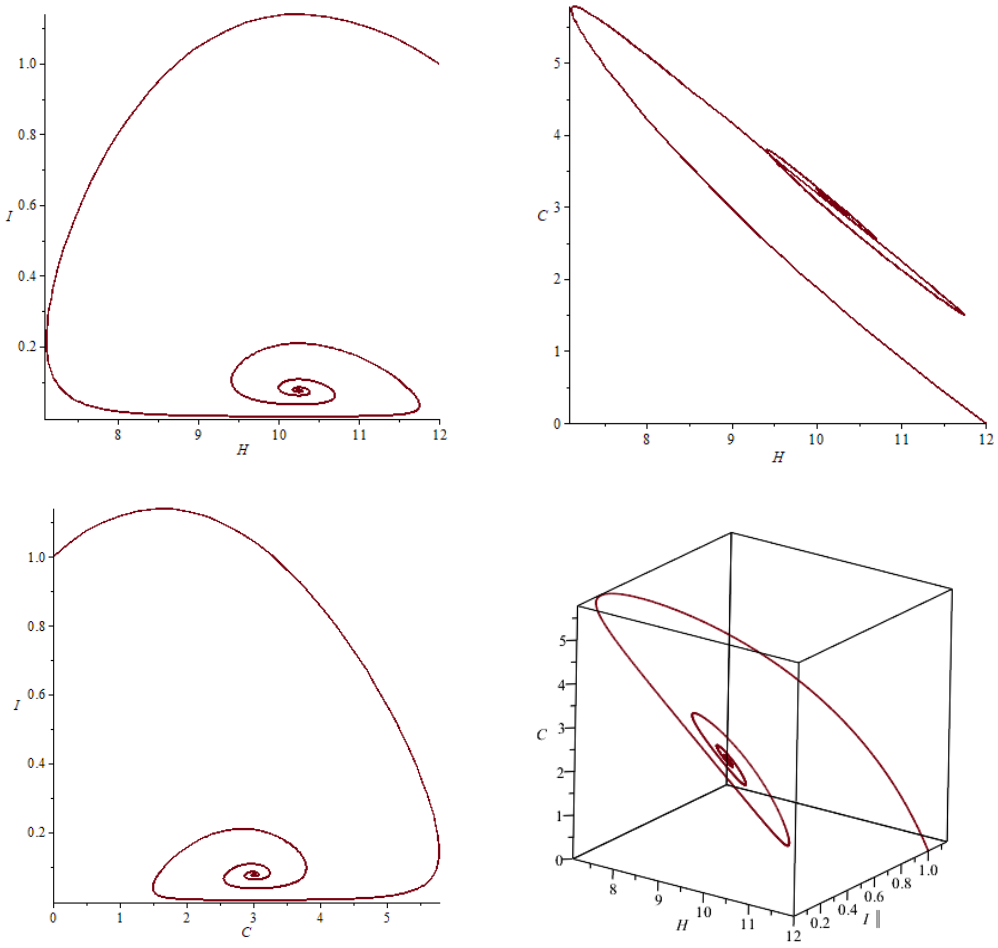

Figure 2 Example 2

3.3.4 Example 3

A situation when the majority of computers in the network are "healthy", only one computer is infected. The network parameters reflect a situation when the network should gradually stabilize at E_1. As we can see in Fig. 3, the system positively converges to equilibrium, as expected.

Parameters: α=4; β=0,3; γ=1,2; d=12

Initial function values: H=2; I=1; C=10

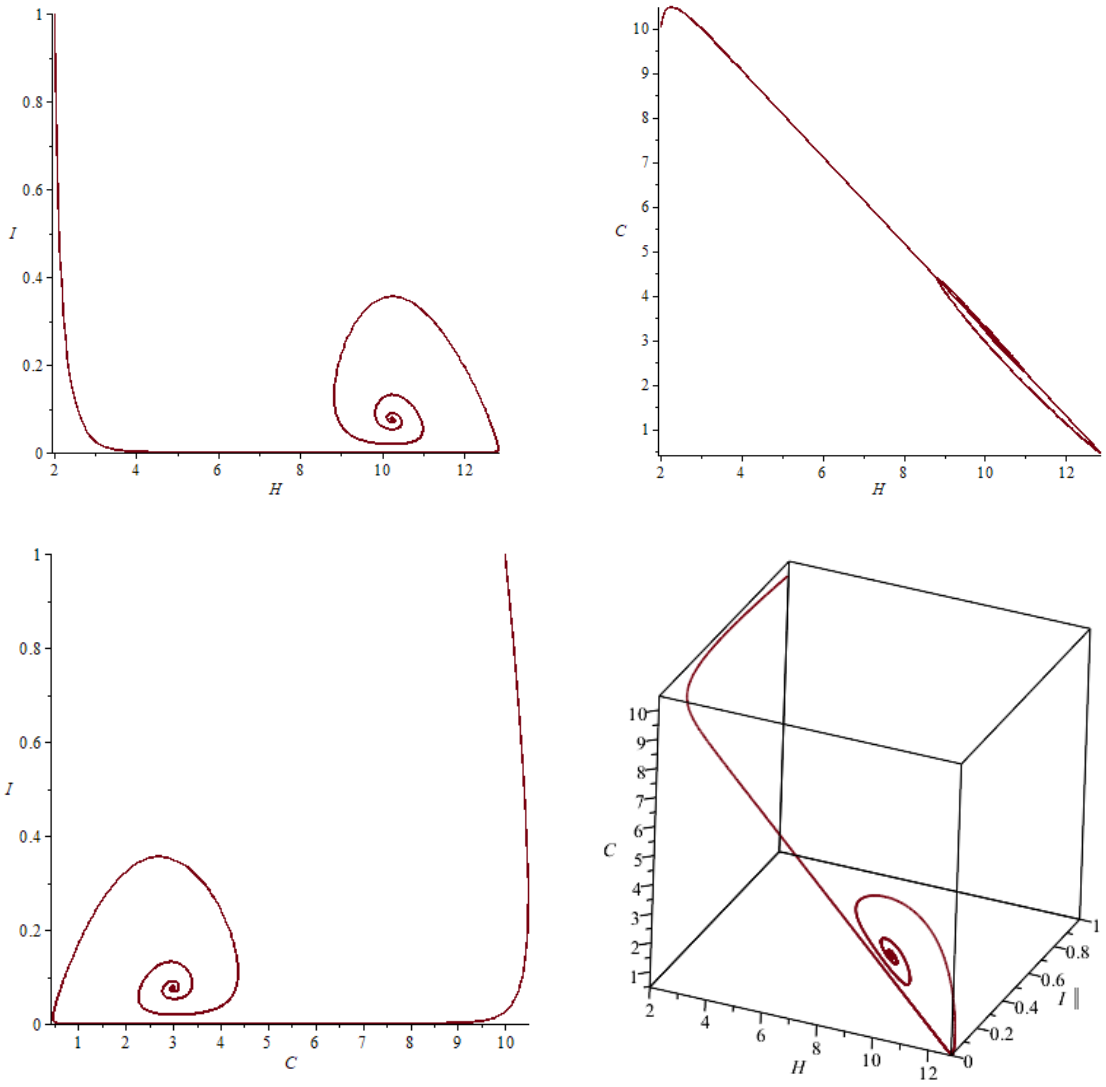

Figure 3 Example 3

Conclusion

Together with the increasing need to use computer technology and the increasing computer network integration, there is an increased risk of infiltration with one of the ever growing range of emerging viruses. Thus this danger should not be underestimated and computer data protection should be of the highest possible quality. If we understand how computer viruses can spread in computer network, it will be easier for us to decide what type of protection to choose.

References

BERNOULLI, D. (1760). *Essai d'une nouvelle analyse de la mortalité causé par la petite vérole et des avantages l'inoculation pour la prevenir*. Paris: Histoire de l'académie royale des sciences. Mémoires de mathématiques et de physique. Histoire de l'académie royale des sciences. p.1-45.

COHEN, F. (1987). Computer viruses: Theory and experiments. *Computers & Security*, 6(1), 22-35. Doi 10.1016/0167-4048(87)90122-2.

HONG, S., YANG, H., ZHAO, T., MA, X. (2015). Epidemic spreading model of complex dynamical network with the heterogeneity of nodes. *International Journal of Systems Science*, 47(11), 2745-2752. Doi 10.1080/00207721.2015.1022890.

CHEN, L., HATTAF, K., SUN, J. (2015). Optimal control of a delayed SLBS computer virus model. *Physica A: Statistical Mechanics and its Applications*, 427, 244-250. Doi 10.1016/j.physa.2015.02.048.

KEPHART, J. O., WHITE, S. R., CHESS, D. (1993). Computers and epidemiology. *IEEE Spectrum*, 30(5), 20-26. Doi 10.1109/6.275061.

KERMACK, W. O., McKENDRICK, A. G: (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 115(772), 700-721. Doi 10.1098/rspa.1927.0118.

KERMACK, W. O., McKENDRICK, A. G. (1932). Contributions to the Mathematical Theory of Epidemics. II. The Problem of Endemicity. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 138(834), 55-83. Doi 10.1098/rspa.1932.0171.

LI, D., LIU, S., Jing'an CUI, J. (2019). Threshold dynamics and ergodicity of an SIRS epidemic model with semi-Markov switching. *Journal of Differential Equations*, 266(7), 3973-4017. Doi 10.1016/j.jde.2018.09.026.

MISHRA, B. K., JHA, N. (2010). SEIQRS model for the transmission of malicious objects in computer network. *Applied Mathematical Modelling*, 34(3), 710-715. Doi 10.1016/j.apm.2009.06.011.

MISHRA, B. K., SAINI, D. K. (2007). SEIRS epidemic model with delay for transmission of malicious objects in computer network. *Applied Mathematics and Computation*, 188(2), 1476-1482. Doi 10.1016/j.amc.2006.11.012.

PARE, P. E., BECK, C. L., NEDIC, A. (2018). Epidemic Processes Over Time-Varying Networks. *IEEE Transactions on Control of Network Systems*, 5(3), 1322-1334. Doi 10.1109/TCNS.2017.2706138.

THOMMES, R.W., COATES, M. J. (2005). Modeling Virus Propagation in Peer-to-Peer Networks. *IEEE*, 981-985. Doi 10.1109/ICICS.2005.1689197.

WANG, F., ZHANG, Y., WANG, CH., MA, J., MOON, S. J. (2010). Stability analysis of a SEIQV epidemic model for rapid spreading worms. *Computers & Security*, 29(4), 410-418. Doi 10.1016/j.cose.2009.10.002.

YANG, L.-X., YANG, X., ZHU, Q., WEN, L. (2013). A computer virus model with graded cure rates. *Nonlinear Analysis: Real World Applications*, 14(1), 414-422. Doi 10.1016/j.nonrwa.2012.07.005.

YAN, P., LIU, S. (2006). SEIR epidemic model with delay. *The ANZIAM Journal*, 48(01). Doi 10.1017/S144618110000345X.

YUN-PENG, X., SONG-YANG, L., YAN-BING, L. (2017). An information diffusion dynamic model based on social influence and mean-field theory. *Acta Physica Sinica*, 66(3), 030501/1-030501/13. Doi 10.7498/aps.66.030501.

ZHANG, X., WU, J., ZHAO, P., SU, X., CHOI, D. (2018). Epidemic spreading on a complex network with partial immunization. *Soft Computing*, 22(14), 4525-4533. Doi 10.1007/s00500-017-2903-1.