Chapter

# A Review of Mathematical Model Based in Clustered Computer Network

*Cristiane M. Batistela and José Roberto C. Piqueira*

## Abstract

The threats produced by viruses in computer networks have been frequent and the subject of many studies. Computer viruses share common characteristics with biological viruses, and therefore, one of the ways to study the dynamics of virus propagation has been through biological analogies. Inspired by macroscopic models, the susceptible-infected-removable (SIR) model allowed variations of compartmental models and suggested defenses considering antidotal (SIRA) and quarantined compartments (SIQRA), giving rise to models that evaluate the effectiveness and strategies to control the spread of viruses in networks. Recently, with the rapid popularization and access to networks, new studies have been taken into consideration the clusters of association of networks, indicating new control strategies and particularities of the dynamics. Toward this goal, this chapter presents a review of the mathematical model based in clustered computer network with the brief overview of the mathematical model reviews and providing an integrated framework to clustered model. In this essay, there is a discussion about the several ways of applying compartmental models to study the propagation of computer viruses and malwares through networks, emphasizing the effect of connections between geographically distributed machine clusters.

**Keywords:** bifurcation, cluster, disease-free, equilibrium point, SIR, stability

## 1. Introduction

Computer viruses emerged as programs capable of harming the functioning of a machine. Initially, the damage was minor as well as its proliferation capacity. With the increasing access to communication networks, the great development of hardware and software, and the inclusion of these services as an essential part of daily life, computer viruses have become a threat [1, 2].

Virus program codes are complex and easy to replicate, and detection and removal by antivirus programs are difficult [3]. Some feats of these viral programs are the ability to acquire bank passwords, personal data, and confidential information [4], which can cause immeasurable damage [5].

The increase in the use of mobile devices combined with the increasing access to wireless Internet facilitated the execution of many daily tasks such as accessing e-mail, electronic transactions of various natures and created opportunities for the advent of the Internet of Things (IoT), and connecting sensors and actuators to allow different types of objects to establish connections to the Internet, including home appliances, cars, and even industrial equipment. Therefore, these items are able to collect and transmit data from the cloud, contributing to a digital transformation in the world, and can provide several improvements in human life [6].

Consequently, understanding the spread of viruses in computer networks has become fundamental for the establishment of strategies to control and mitigate the spread of viruses. To improve the security and reliability of networks, a new branch of study, known as cybersecurity, has contributed to guide control strategies in order to minimize losses and one of its approaches is to build mathematical models.

One of the areas of cybersecurity is related to the study of the propagation of viruses in computer networks. Many ways of approaching the problem have been relevant to the understanding of the dissemination of malware, including the mathematical approach.

The mathematical study of computer viruses has an inspiration in biology and can be understood at two levels: microscopic and macroscopic [2, 7]. The tools used to develop antivirus programs, which are programs capable of detecting threats and preventing damage to machines, are concentrated at the microscopic level. In addition, the propagation of viruses in the network can be mitigated by the action of antivirus and quarantines proposed by the software when detecting some unexpected action [8].

The macroscopic level was developed from the classical model of disease propagation whose dynamics indicate the possibility of infection [9] and favor the orientation of strategies to control dissemination. The classic epidemiological model, proposed by Kermack and McKendrick, suggests the division of the population into compartments containing the group of susceptible (S), infected (I), and removed (R), giving rise to the SIR model, whose dynamics and parameters indicate strategies for control [9, 10].

Inspired by the above-mentioned works and based on the compartment level SIR model, this chapter considers the review of relationship between networks and the influence of the biolocical compartmental models for cybersecurity. Different from the conventional compartimental level models, this study shows the issue of how the association of two compartmental models occurs and indicates future prospects.

This work reviews the develop and analyze a model of virus propagation in two independent populations where those who ware infected from one population can come into contact with those who are susceptible from the other, to analyze the effects that one infected population can cause on the other. For this, two clusters were created and each cluster represents a population, both exposed to the same virus and represented by a model with antoidotal compartment. To represent the interaction between the two populations, a new connection between the sets was created and represents the capacity of an infected person come into contact with a susceptible one from the other, and this interaction will be modulated by a parameter.

The remaining chapter is ordered as follows: In the next section, a review of epidemiological models is presented with applications in computers, in section 3, hypothesis and equations are presented for the model with antidotal compartment, and the cluster model is presented, followed by the conclusion.

## 2. Epidemiological models

In 1927 was published for the first time a deterministic model to study the dynamics of virus spread in populations. This was a compartmental model consisting of three compartments (susceptible—infected—removed) [9]. In this work, a theory was developed relating the development of an epidemic to a critical value, later known as the basal reproduction number $R_0$.

The modeling of epidemics is associated with the dynamic behavior of processes where populations are studied according to their epidemiological status, these processes are described by differential equations, and the dynamics between their states is given by different parameters such as birth rate, mortality rate, infection, and recovery rate.

The modeling of the spread of epidemics has been the objective of many works [11–15]. These models allow a better understanding of the mechanisms of disease spread and can lead to more effective control strategies.

The scientific literature in epidemiology is quite diverse. Among the most cited models regarding this topic are the models: susceptible—infectious—susceptible (SIS) models [16–19]; susceptible—infected—recovered (SIR) models [20–25]; susceptible—exposed—infected—recovered (SEIR) models [26–28]; susceptible—exposed—infected—quarantined—recovered (SEIQR) models; susceptible—exposed—infected—quarantined—recovered—susceptible (SEIQRS) [29]; susceptible—exposed—infected—recovered—susceptible—vaccined (SEIRS-V) [30]; susceptible—exposed—infected—susceptible—vaccined (SEIS-V) [31] and others.

As for the way of treating chance, it can be classified into two levels: stocchastic and deterministic. In the first case, the model includes variables, giving a probabilistic distribution to the system, incorporating uncertainty, an intrinsic characteristic of epidemiological systems [32–35]. On the other hand, deterministic models provide the same results every time they are simulated with the same initial conditions [36, 37], being suitable to verify system sensitivity to the variation of the parameters [20, 21, 38].

Adapting the SIR model to computers, a lot of research has contributed to the understanding of virus propagation [26, 28–31, 39, 40] and one of the main goals is to establish effective security strategies [41].

The most explored strategies in cybersecurity are related to the use of antiviral compartments (A) and quarantine (Q). The adaptation of the SIR model gave rise to some robust models, including susceptible—infected—removed—antidotal (SIRA) [21] and susceptible—infected—removed—antidotal (SIQRA) [42].

Following this line, the first analysis of clustered computer networks using an epidemiological model studied the influence between two networks equipped with computers with antivirus and evaluated the dynamics of virus promotion and suggested viral dissemination control strategies.

## 3. Cluster SIRA model: Hypothesis and equations

There is a lot of compartimental models indicated for epidemiology [43] and their origin is Kermack and Mckendrick SIR (susceptible—infected—removed) models [9, 43, 44].

The population is considered constant and is divided into three compartments: Susceptible computers are uninfected and subject to infection (S); infected computers are represented by (I), those removed by infection or not (R), as shown in **Figure 1**.

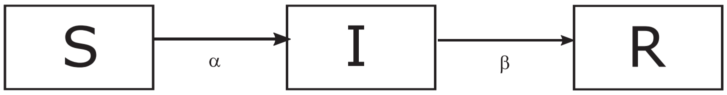

**Figure 1.**
*SIR model.*

As reported by [21] the dynamic equation for the populations $S$, $I$ and $R$ are:

$$\begin{cases} \dot{S} &= -\alpha SI; \\ \dot{I} &= \alpha SI - \beta I; \\ \dot{R} &= \beta I. \end{cases} \tag{1}$$

The susceptible population $S$ is infected with a rate, that is, related to the probability of susceptible individuals to establish effective communications with infected ones. Therefore, this rate is proportional to the product $SI$, with proportion factor represented by $\alpha$ and infected individual can become removed with a rate controlled by $\beta$.

Considering initial conditions $S(0) \geq 0$, $I(0) \geq 0$ and $R(0) \geq 0$, in model such as SIR the interest is to investigate the dynamics of virus propagation indicates whether the virus will remain in the network or if it will naturally be eradicated. One of the ways to evaluate this behavior is to study the basal reproduction rate $(R_0)$. This number indicates whether the virus will continue to be propagated and will be considered a situation analogous to the endemic one, or if it will become extinct in the network.

Based on a model described by (1), a model with a modification, including an antidotal population compartment (A) representing nodes of the network equipped with fully effective antivirus programs, is studied and considering constant population with four compartments: susceptible computers are uninfected and subject to infection (S); infected computers are represented by (I), and those removed by infection or not (R) and (A) are uninfected computers equipped with antivirus, as shown in **Figure 2**.

As reported by [21] the dynamic equation for the populations $S$, $I$, $R$ and $A$ are:

$$\begin{cases} \dot{S} &= N - \sigma_{SA}SA - \beta SI - \mu S + \sigma R; \\ \dot{I} &= \beta SI - \alpha_{IA}AI - \delta I - \mu I; \\ \dot{R} &= \delta I - \sigma R - \mu R; \\ \dot{A} &= \sigma_{SA}SA + \alpha_{IA}AI - \mu A. \end{cases} \tag{2}$$

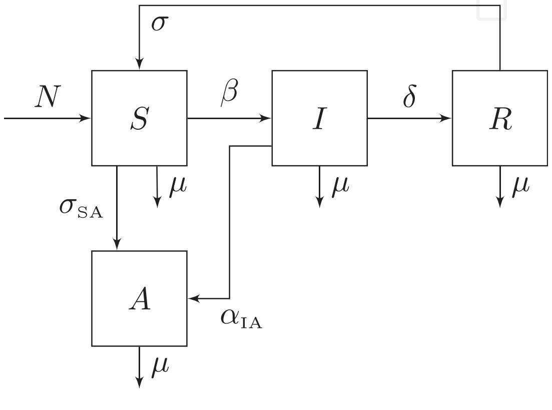

**Figure 2.**
*SIRA model.*

The influx rate is considered to be $N = 0$ because during the propagation of the considered virus, there is no incorporation of new computers in the network. The choice of $\mu = 0$ is justified that the machines obsolescence time is larger than the time of the virus action.

The model represents the spread of a known virus and the conversion of the antidoto to the infected is not considered. In this model, a vaccination strategy can be defined implying a control strategy associated with the economic use of antivirus programs.

The analysis of the SIRA model shows that it is possible to reach an disease-free equilibrium, guaranteeing a good operational performance of the network and that even in a situation of endemic equilibrium, the introduction of at least one machine equipped with antivirus guarantees a good performance of the network, tending to a disease-free equilibrium.

Furthermore, considering a constant total number of machines, the main control parameters are associated with the infection rate and how quickly infected machines are removed for formatting procedure. The other network parameters are associated with the transient response of the network in some small disturbance.

The other variation of the SIRA model is improved by considering that, when the machines pass to the removed condition, a fraction of these machines is recovered and the complement is considered dead. The introduction of the mortality rate results in an increase of the robustness of the disease-free equilibrium point of a computer network [38].

The study of the SIRA model was complemented by considering another control strategy by adding a quarantined compartment. The new compartment can be evaluated for the presence or absence of saturation and both situations indicate robustness in control strategies.

Based on a model described by [20], the virus propagation in a cluster is studied [45]. The model proposed is an association of two networks constituted by the SIRA model that interacts as shown in **Figure 3**.

Considering this hypothesis, adding another compartimental model, and associating a new infection rate, representing the infection capacity to the network, the cluster SIRA model for viruses propagation was proposed has the following dynamical eqs. (3):

$$\begin{cases} \dot{S}_1 &= -\alpha_{SA1} S_1 A_1 - \beta_1 S_1 I_1 - \rho_2 I_2 S_1 + \theta 1 R_1; \\ \dot{I}_1 &= \beta_1 S_1 I_1 - \delta_1 I_1 - \alpha_{IA1} I_1 A_1 + \rho_2 I_2 S_1; \\ \dot{R}_1 &= \delta_1 I_1 - \theta_1 R_1; \\ \dot{A}_1 &= \alpha_{SA1} S_1 A_1 + \alpha_{IA1} I_1 A_1; \\ \dot{S}_2 &= \alpha_{SA2} S_2 A_2 - \beta_2 S_2 I_2 - \rho_1 I_1 S_2 + \theta 2 R_2; \\ \dot{I}_2 &= \beta_2 S_2 I_2 - \delta_2 I_2 - \alpha_{IA2} I_2 A_2 + \rho_1 I_1 S_2; \\ \dot{R}_2 &= \delta_2 I_2 - \theta_2 R_2; \\ \dot{A}_2 &= \alpha_{SA2} S_2 A_2 + \alpha_{IA2} I_2 A_2. \end{cases} \tag{3}$$

For the cluster SIRA model, the susceptible population $S$ is infected with a rate, that is, related to the probability of susceptible elements to establish effective communications with infected ones and this rate is proportional to the product $SI$, with proportion factor represented by $\alpha$ or if infectivity occurs between network, by rate $\rho$ that is related to the probability of infected elements to establish effective communications with susceptible computer in another network.

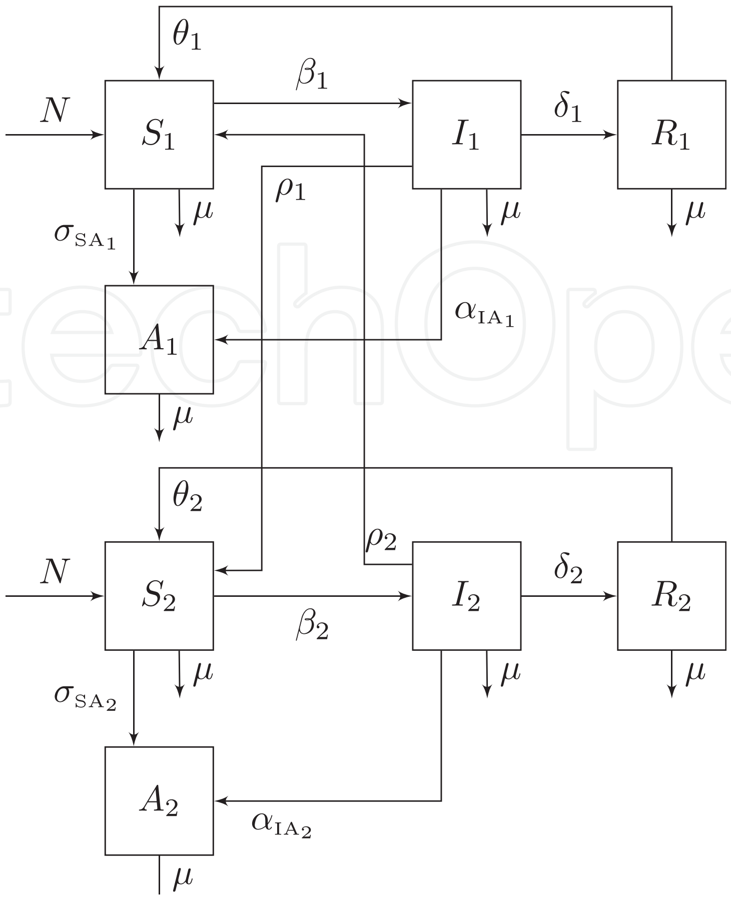

**Figure 3.**
*SIRA cluster model.*

Clustered sets, represented by subscripts 1 and 2, are divided into four groups, as shown in **Figure 3** and populations are considered constant in each cluster.

The SIRA cluster study, despite presenting a simple model composed of two connected grouped networks, points out the main control strategies associated with the control of parameters in order to avoid new forms of attacks.

Among the possibilities for adjustments, we consider infection rates in susceptible populations, due to contact with infected populations from the same cluster ($\beta$); conversion rates are removed by infected ($\delta$) and infection rates in susceptible populations due to contact with the infected population of the other cluster ($\rho$).

The choice of network topology and connection strategies is an effective strategy to reduce the spread of viruses, since infection rates are not known in advance.

Another way to prevent the spread of viruses is to maintain the removal rates of damaged machines, which plays an important role in controlling the spread of networks.

If the virus is known, the best strategy to prevent its spread is to introduce antidote nodes containing programs that can be propagated throughout the network, immunizing the other nodes.

## 4. Conclusions

This chapter reports a detailed survey of compartments model in computer viruses with antidotal machine. The review provided a wide application of epidemiological models in compartmental models applied to computer networks. The focus of the review was to show the infuence of machines equipped with antivirus and their control strategies on the spread of viruses. The study of clustered networks is recent and with the model presented, it is expected that the review provides tools for studies of virus propagation dynamics in more complex networks.

## Acknowledgements

JRCP is supported by the Brazilian Research Council (CNPq), grant number: 302883/2018–5.

## Conflict of interest

The author declares that there is no conflict of interests regarding the publication of this study.

## Author details

Cristiane M. Batistela[1] and José Roberto C. Piqueira[2]*

1 Federal University of ABC – UFABC, São Bernardo do Campo, SP, Brazil

2 Polytechnic School of University of São Paulo – EPUSP, São Paulo, SP, Brazil

*Address all correspondence to: piqueira@lac.usp.br

## References

[1] Denning PJ, editor. Computers under Attack: Intruders, Worms, and Viruses (Vol. 990). New York: ACM Press; 1990

[2] Cohen F. Computer viruses: Theory and experiments. Computers & Security. 1987;**6**(1):22-35

[3] Tippett PS. The kinetics of computer virus replication: A theory and preliminary survey. In: Safe Computing: Proceedings of the Fourth Annual Computer Virus and Security Conference. 1991. pp. 66-87

[4] Yang LX, Yang X. A new epidemic model of computer viruses. Communications in Nonlinear Science and Numerical Simulation. 2014;**19**(6): 1935-1944

[5] Cohen FB. A Short Course on Computer Viruses. Pittsburgh, PA, USA: John Wiley & Sons, Inc; 1994

[6] Botta A, De Donato W, Persico V, Pescapé A. Integration of cloud computing and internet of things: A survey. Future Generation Computer Systems. 2016;**56**:684-700

[7] Kephart JO. Direct-graph epidemiological models of computer virus. In: Proceedings of IEEE Symposium on Security and Privacy. Oakland, CA, USA: IEEE; 1991

[8] Kephart JO, White SR. Measuring and modeling computer virus prevalence. In: Proceedings 1993 IEEE Computer Society Symposium on Research in Security and Privacy. Oakland, CA, USA: IEEE; 1993. pp. 2-15

[9] Kermack WO, McKendrick AG. Contributions of mathematical theory to epidemics. Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character. 1927;**115**(772):700-721

[10] Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character. 1932;**138**(834):55-83

[11] Anderson RM, Anderson B, May RM. Infectious Diseases of Humans: Dynamics and Control. New York, USA: Oxford University Press; 1992

[12] Murray JD. Mathematical Biology. 3rd ed. New York: Springer-Verlag; 2002

[13] Clancy D. Optimal intervention for epidemic models with general infection and removal rate functions. Journal of Mathematical Biology. 1999;**39**(4): 309-331

[14] Allen LJ, Brauer F, Van den Driessche P, Wu J. Mathematical Epidemiology. Vol. 1945. Berlin: Springer; 2008. p. 2008

[15] Brauer F, Castillo-Chavez C, Castillo-Chavez C. Mathematical Models in Population Biology and Epidemiology. Vol. 2. New York: Springer; 2012. p. 2012

[16] Amador J, Artalejo JR. Modeling computer virus with the BSDE approach. Computer Networks. 2012;**57**(1):302-316

[17] Sanders J, Noble B, Van Gorder RA, Riggs C. Mobility matrix evolution for an SIS epidemic patch model. Physica A; Statistical Mechanics and its Applications. 2012;**391**(24):6256-6267

[18] Wang Y, Cao J, Jin Z, Zhang H, Sun GQ. Impact of media coverage on

epidemic spreading in complex networks. Physica A; Statistical Mechanics and its Applications. 2013; **392**(23):5824-5835

[19] Tomovski I, Trpevski I, Kocarev L. Topology independent SIS process: An engineering viewpoint. Communications in Nonlinear Science. 2014;**19**(3):627-637

[20] Piqueira JRC, De Vasconcelos AA, Gabriel CE, Araujo VO. Dynamic models for computer viruses. Computers and Security. 2008;**27**(7–8):355-359

[21] Piqueira JRC, Araujo VO. A modified epidemiological model for computer viruses. Applied Mathematics and Computation. 2009;**213**(2):355-360

[22] Ren J, Yang X, Yang LX, Xu Y, Yang F. A delayed computer virus propagation model and its dynamics. Chaos Soliton & Fractals. 2012;**45**(1): 74-79

[23] Wierman JC, Marchette DJ. Modeling computer virus prevalence with a susceptible-infected-susceptible model with reintroduction. Computational Statistics & Data Analysis. 2004;**45**(1):3-23

[24] Zhu Q, Yang X, Ren J. Modeling and analysis of the spread of computer virus. Communications in Nonlinear Science. 2012;**17**(12):5117-5124

[25] Shukla JB, Singh G, Shukla P, Tripathi A. Modeling and analysis of the effects of antivirus software on an infected computer network. Applied Mathemaics and Computation. 2014; **227**:11-18

[26] Mishra BK, Saini DK. SEIRS epidemic model with delay for transmission of malicious objects in computer network. Applied Mathematics and Computation. 2007; **188**(2):1476-1482

[27] Wang F, Zhang Y, Wang C, Ma J. Stability analysis of an e-SEIAR model with point-to-group worm propagation. Communications in Nonlinear Science. 2015;**20**(3):897-904

[28] Mishra BK, Pandey SK. Dynamic model of worms with vertical transmission in computer network. Applied Mathematics and Computation. 2011;**217**(21):8438-8446

[29] Mishra BK, Jha N. SEIQS model for the transmission of malicious objects in computer network. Applied Mathematical Modelling. 2010;**34**(3): 710-715

[30] Mishra BK, Keshri N. Mathematical model on the transmission of worms in wireless sensor network. Applied Mathematical Modelling. 2013;**37**(6): 4103-4111

[31] Mishra BK, Pandey SK. Dynamic model of worm propagation in computer network. Applied Mathematical Modelling. 2014;**38**(7–8):2173-2179

[32] Radha M, Balamuralitharan S, Geethamalini S, Geetha V, Rathinasamy A. Analytic solutions of the stochastic SEIA worm model by homotopy perturbation method. AIP Conference Proceedings. 2019;**2112**(1): 020050

[33] Amador J, Artalejo JR. Stochastic modeling of computer virus spreading with warning signals. Journal of the Franklin Institute. 2013;**350**(5):1112-1138

[34] Amador J. The stochastic SIRA model for computer viruses. Applied Mathematics and Computation. 2014; **232**:1112-1124

[35] Zhang C, Zhao Y, Wu Y, Deng S. A stochastic dynamic model of computer viruses. Discrete Dynamics in Nature and Society. 2012;**2012**:1-16

[36] Geethamalini S, Balamuralitharan S, Radha M, Geetha V, Rathinasamy A. Stability analysis of deterministic SEIA worm model by reproductive number. AIP Conference Proceedings. 2019; **2112**(1):020044

[37] Geetha V, Balamuralitharan S, Geethamalini S, Radha M, Rathinasamy A. Analytic solutions of the deterministic SEIA worm model by homotopy perturbation method. AIP Conference Proceedings. 2019;**2112**(1): 020100

[38] Batistela CM, Piqueira JRC. SIRA computer viruses propagation model: Mortality and robustness. International Journal of Applied and Computational Mathematics. 2018;**4**(5):128

[39] Martcheva M. An Introduction to Mathematical Epidemiology. Vol. 61. New York: Springer; 2015. p. 2015

[40] Wang F, Zhang Y, Wang C, Ma J. Stability analysis of an e-SEIAR model with point-to-group worm propagatio. Communications in Nonlinear Science. 2015;**20**(3):897-904

[41] Li P, Yang X, Xiong Q, Wen QJ, Tang YY. Defending against the Advanced Persistent Threat: An Optimal Control Approach. Security and Communication Networks. 2018; **2018**:1-14

[42] Piqueira JRC, Batistela CM. Considering quarantine in the SIRA malware propagation model. Mathematical Problems in Engineering. 2019

[43] Kermack WO, McKendrick AG. Contributions of mathematical theory to epidemics. Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character. 1932;**138**(834):55-83

[44] Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics - further studies of the problem of endemicity. Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character. 1933;**141**(843): 94-122

[45] Piqueira JRC, Cabrera MA, Batistela CM. Malware propagation in clustered computer networks. Physica A: Statistical Mechanics and its Applications. 2021;**573**:125958