Understanding contagion spreading process of cyber security threats through social networks

Terry Brett

Supervised by Assistant Prof. Nicol a Perra, Assistant Prof. George Loukas, Prof. Yamir Moreno (University of Zaragoza)

Abstract

The spreading of ideas, memes, norms, products and diseases are few examples of phenomena that can be studied and modelled as contagion process on networks. Similarly the propagation of computer virus can be modelled as a disease, which instead of humans infects machines.

Many studies have focused on the propagation of computer viruses on the web in terms of their features, yet failed to tackle the social aspect behind the spread. In other words the interactions of users on OSNs (online social networks) have been largely neglected.

This project will tackle this limitation, by introducing an experimental online platform to observe the spreading of simulated cyber threats in a population of connected users.

Current Issues

Despite using technology on day to day basis, humans still fail to spot unsafe cyberspace environment [D. C. Rowe, Lunt, and Ekstrom, 2011]. The average user cannot identify all security indicators [Heartfield, Loukas, and Gan, 2017].

Considering browsing the web; the experience is completely different on a desktop computer and a mobile device. For example the HTTPS security indicator visible in the URL address bar disappears on a mobile device. Lack of these indicators can lead to higher probability of cyber attack.

A new concept has been introduced to tackle this issue head on, in which it is the human that is to detect a potential threat [Heartfield, Loukas, and Gan, 2016]. The idea is that the reports from a user will yield a stronger indications to a potential cyber threat compared to something detected by software, which may be a mistake, or it might not detect the threat at all [Sukwong, H. Kim, and Hoe, 2010].

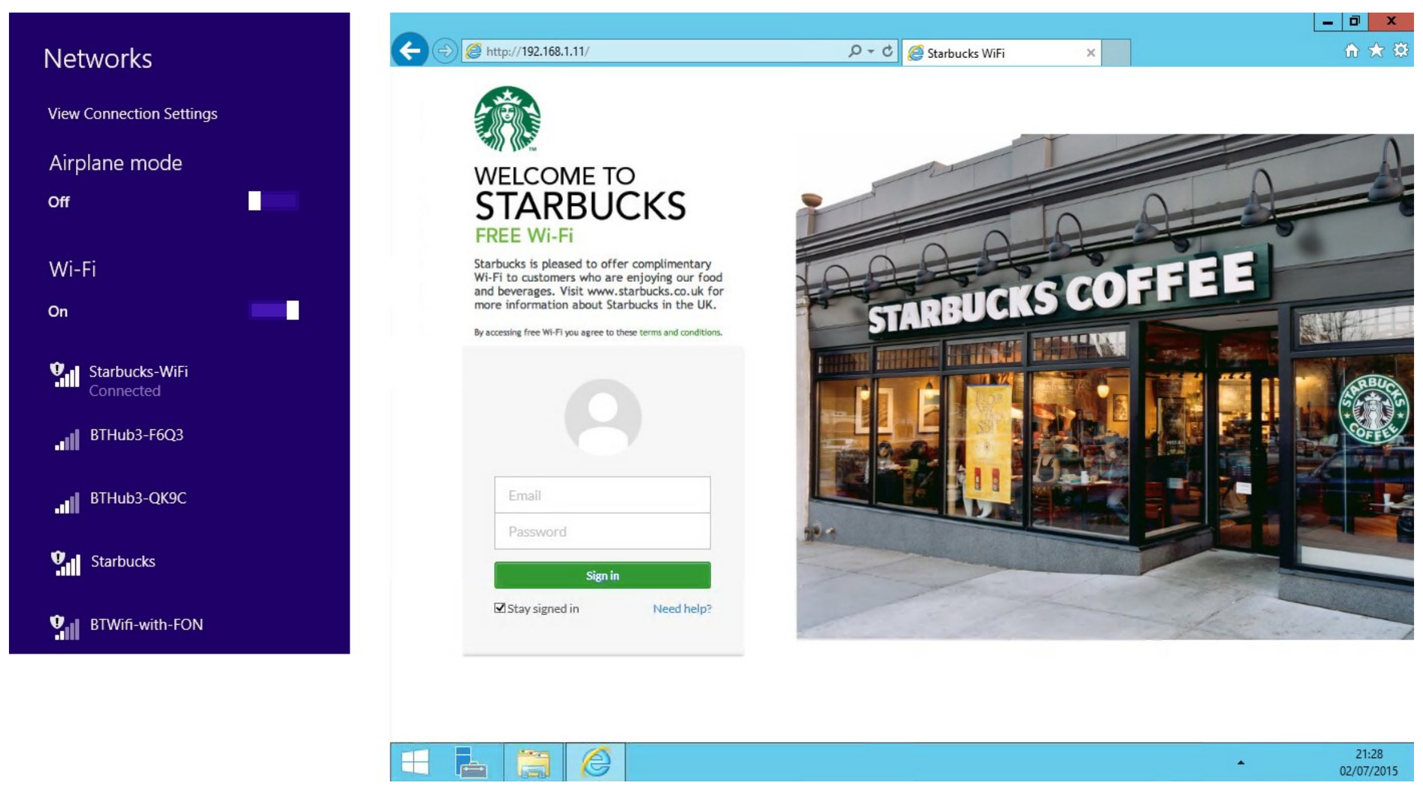
Since cyber threats are similar to their biological counterparts [Kephart and White,1992], we can use the compartmental spreading models to study the behaviour of a virus on a social network.

Although number of studies have been carried out modelling the contagion spread on social networks, their main focus is on the features of the virus [Guillén, Rey, andEncinas, 2017, T. M. Chen and Robert, 2004]. This research follows a novel approach, in which we are describing the largely neglected, unsupervised social interactions on these networks.

Cyber Attacks Examples

We can identify three types of deception attacks, cosmetic, behavioural and hybrid [Heartfield and Loukas, 2016].

1. Cosmetic type of attacks focus on the user interface, for example a file can have a right icon association with user expectation, such as adobe reader for pdf files, yet be a .pdf.exe file, which is an executable.
2. Behavioural attacks mimic the behaviour of a system. Following certain standards and conventions users are tricked to believe that what they are using is indeed legitimate, such as seeing an open WiFi connection in the list of available networks.
3. Finally hybrid exception combines the two. Not only the application in use looks legitimate, but it also behaves in the same way the original would. This is because some attackers would copy the code from the original [Dhamija, Tygar, and Hearst, 2006]. The combination of two therefore creates a convincing attack, which is more likely to be successful.



WELCOME TO STARBUCKS FREE Wi-Fi

Starbucks is pleased to offer complimentary Wi-Fi to customers who are enjoying our food and beverages. Visit www.starbucks.co.uk for more information about Starbucks in the UK.

By accessing free Wi-Fi you agree to these terms and conditions.

Fig. 1. Man-in-the-middle attack

In this cyber attack, have a hybrid approach, which combines both the visuals and behaviour of a system. This is an attack since we can see that there is multiple open Starbucks networks available within the computers range. More importantly the URL is the giveaway. Here we have https://192.168.1.1, this is a reserved IP for private network use, indicating that this Starbucks login page is hosted on someone's laptop within the coffee shop.

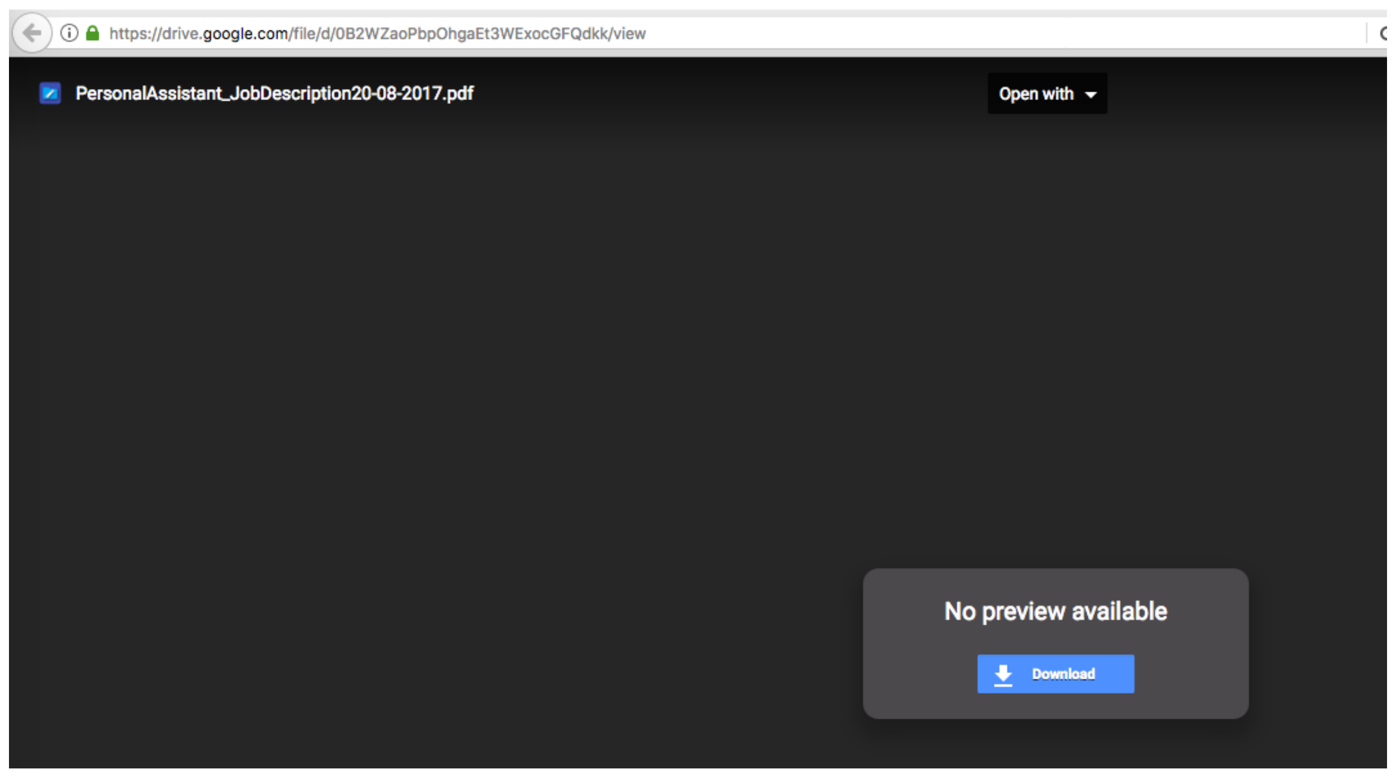
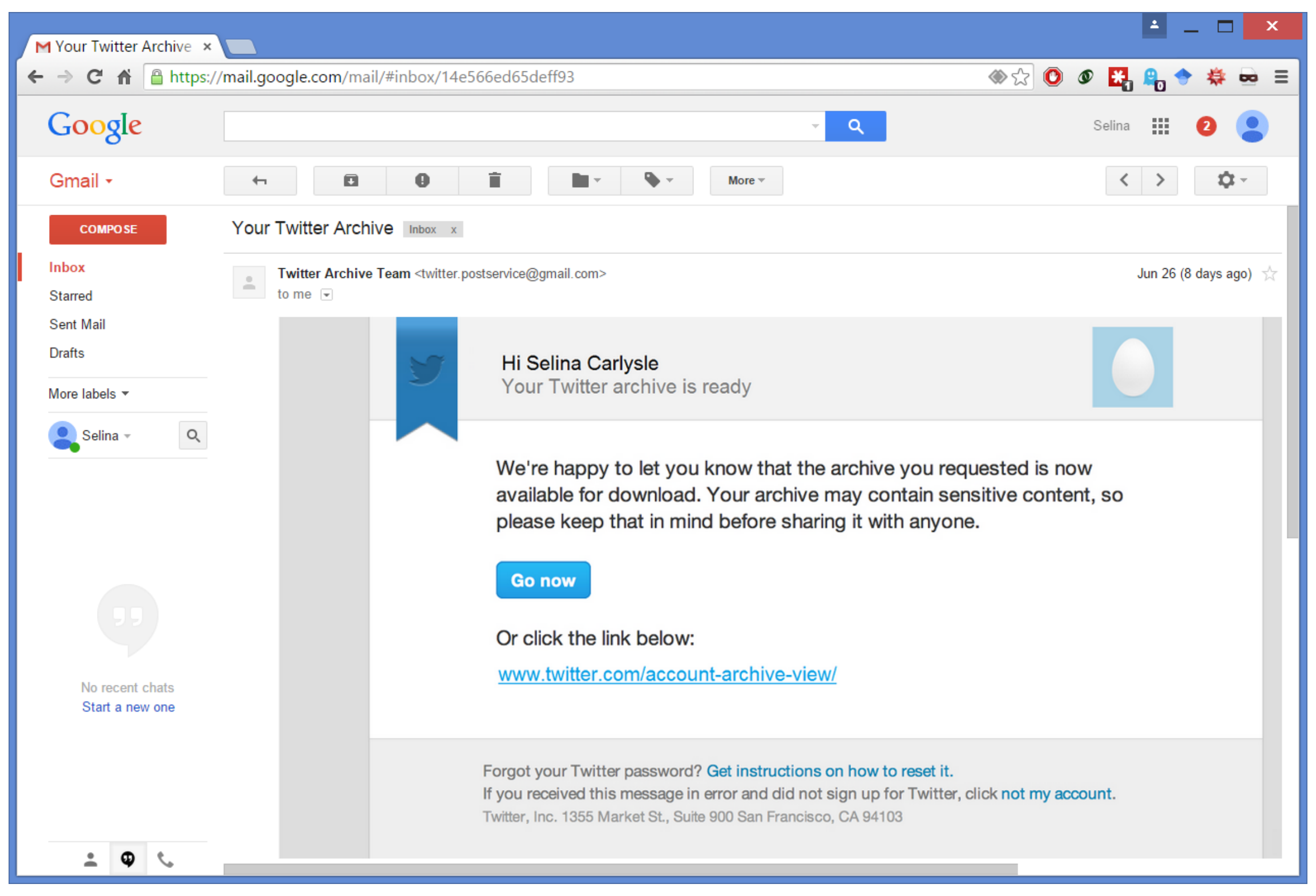

https://drive.google.com/file/d/0B2WZaoPbpOhgaEt3WExocGFQdkk/view
PersonalAssistant_JobDescription20-08-2017.pdf
Open with
No preview available
Download

Fig. 2. Double extension attack

Here we have an example of a double extension attack. The victim is supposed to be lead into believing that the files shared with them is a PDF. This at first might seem legitimate, however many online storage providers (Google Drive in this case) are able to provide a preview of commonly shared files such as PDFs, word documents, PowerPoints etc. Here we can see that although the file extension is .pdf, there is no preview available indicating a malicious file.



Your Twitter Archive Inbox

Twitter Archive Team <twitter.postservice@gmail.com> to me — Jun 26 (8 days ago)

Hi Selina Carlysle
Your Twitter archive is ready

We're happy to let you know that the archive you requested is now available for download. Your archive may contain sensitive content, so please keep that in mind before sharing it with anyone.

Or click the link below:
www.twitter.com/account-archive-view/

Forgot your Twitter password? Get instructions on how to reset it.
If you received this message in error and did not sign up for Twitter, click not my account.
Twitter, Inc. 1355 Market St., Suite 900 San Francisco, CA 94103

Fig. 3. Phishing attacks

Phishing attacks are the most common attacks, without trying to compromise the systems security itself [Cisco, 2019], they use social engineering techniques instead. Phishing attacks are a hybrid type attacks, since the user will not only have to be convinced that the content looks legitimate, but it will also have to behave in the way that the original would.

In this example the email looks like a legitimate confirmation of Twitter archive request, with the link provided even using twitter.com URL. However if we look at who the email was sent from, we can see that the email address comes from a Gmail account, which gives away this message as being a cyber attack.

Research

Social networks are the prime channel for the spreading of computer viruses. Yet the study of their propagation neglects the temporal nature of social interactions and the heterogeneity of users' susceptibility. Here, we introduce a theoretical framework that captures both properties. We study two realistic types of viruses propagating on temporal networks featuring Q categories describing their susceptibility to cyber threats measured in terms of their gullibility and time needed to recover from successful attacks. and derive analytically the invasion threshold. We found that the temporal coupling of categories might increase the fragility of the system to cyber threats. Our results show that networks' dynamics and their interplay with users features are crucial for the spreading of computer viruses.

Results

We model users' interactions using a time-varying network model [Perra et al., 2012] and consider two types of viruses. The first mimics threats that can propagate only via connections activated during the infection period. The second instead, considers viruses able to access also information about past connections. We investigate the impact of different classes of susceptibility considering that they might also influence the link formation process. In all cases, we analytically derive (see Fig. 4 A-C) the conditions regulating the spreading of the virus. Interestingly, these are defined by the interplay between the features of the cyber threats, the categories of susceptibility and their time-varying connectivity. Furthermore, in some scenarios the coupling between categories creates a complex phenomenology that increases the fragility of the system. Our results have the potential to initiate future efforts aimed at describing more realistically the spreading of computer viruses on online social networks.

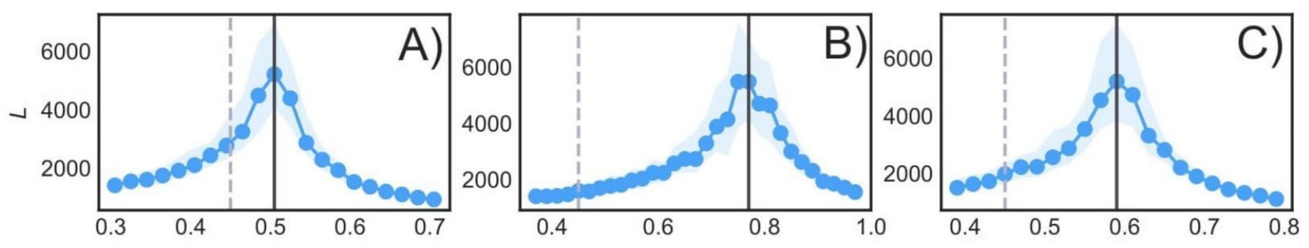
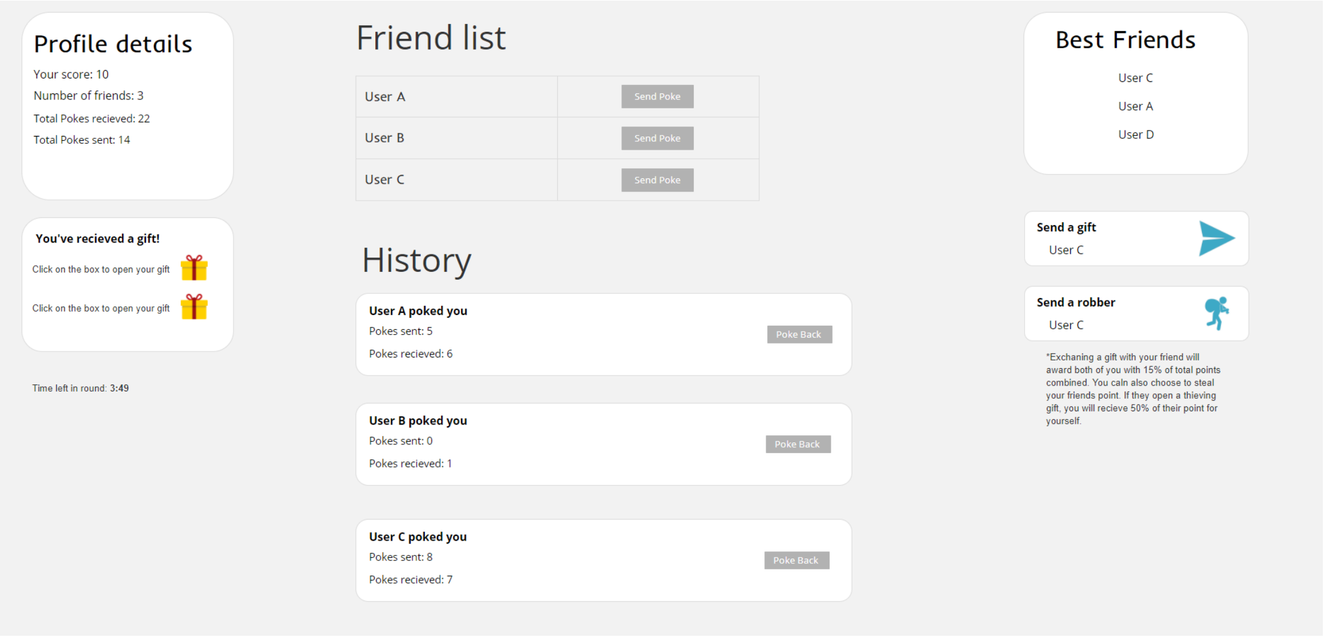


Fig 4: Lifetime of the SIS process (A-C). In A-B nodes are randomly assigned to two categories, in C instead decreasing order of activity. We set $p = 0.9$ (A), $p = 0.4$ (B-C). In A-C we fix $N = 2x10^5, m = 4, \alpha = -2.1, \mu_1 = \mu_2 = 10^{-2}, \lambda_2 = 0.2, Y = 0.3$ and 0.5% of random initial seeds. We plot the median and 50% confidence intervals in 10^2 simulations per point. The solid line represents R_0, the basic reproductive number, and the dashed lines are the analytical threshold in case of a single category.

Platform

The proposed platform will be used to collect the unsupervised interactions between users. We will especially focus on how the users share links and files. Some of the content that will be shared on the platform will be "infected", in such way that it has a "tracker" attached to it, so we can see how it propagates on the network.

Self-adaptive software is able to change its behaviour based on the environmental changes such as user input, hardware, sensors etc. [Oreizy et al., 1999]. The platform can take advantage of this, since we are going to isolate the social factor responsible for spread of malicious content, using techniques of machine learning and a distributed system, we can have the platform perform multiple analysis, such as link unpacking, visual recognition, pattern recognition etc. With these techniques we will try to train the platform to be able adapt to cyber threats, and block them from users, drastically reducing the propagation of malware on OSNs.



Profile details
Your score: 10
Number of friends: 3
Total Pokes recieved: 22
Total Pokes sent: 14

You've recieved a gift!
Click on the box to open your gift
Click on the box to open your gift

Time left in round: 3:49

Friend list	
User A	Send Poke
User B	Send Poke
User C	Send Poke

History

User A poked you
Pokes sent: 5
Pokes recieved: 6

User B poked you
Pokes sent: 0
Pokes recieved: 1

User C poked you
Pokes sent: 8
Pokes recieved: 7

Best Friends
User C
User A
User D

Send a gift
User C

Send a robber
User C

*Exchaning a gift with your friend will award both of you with 15% of total points combined. You caln also choose to steal your friends point. If they open a thieving gift, you will recieve 50% of their point for yourself.

Fig. 5 Early platform prototype screenshot.

Users will be able to exchange the information on the platform, this will include files, images, videos, links. Each of those will be individually tracked, and small number of those will be "infected", so that we can track how the infected content propagated on the social network. From the user perspective, their goal is to collect number of points, the person with highest number of points at each round has a higher chance of winning the monetary award for participating in the experiments.

References

Rowe, Dale C., Barry M. Lunt, and Joseph J. Ekstrom (2011). "The Role of Cyber-security in Information Technology Education". In:Proceedings of the 2011 Conference on Information Technology Education. SIGITE '11. West Point, New York, USA:ACM, pp. 113-122.ISBN: 978-1-4503-1017-8.DOI:10.1145/2047594.2047628.URL:http://doi.acm.org/10.1145/2047594.2047628.

Heartfield, Ryan and George Loukas (2016). "A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks". In:ACM Computing Surveys (CSUR)48.3, p. 37.

Dhamija, Rachna, J Doug Tygar, and Marti Hearst (2006). "Why phishing works". In:Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM,pp. 581-590.

Oreizy, Peyman et al. (1999). "An architecture-based approach to self-adaptive soft-ware". In:IEEE Intelligent Systems and Their Applications14.3, pp. 54-62.

Heartfield, Ryan, George Loukas, and Diane Gan (2017). "An eye for deception: A case study in utilizing the human-as-a-security-sensor paradigm to detect zero-day semantic social engineering attacks". In:Software Engineering Research, Management and Applications (SERA), 2017 IEEE 15thInternational Conference on. IEEE, pp. 371-378.

Newman, Mark EJ, Stephanie Forrest, and Justin Balthrop (2002). "Email networks and the spread of computer viruses". In:Physical Review E66.3, p. 035101.

Guillén, JD Hernández, A Martın del Rey, and L Hernández Encinas (2017). "Study of the stability of a SEIRS model for computer worm propagation". In:Physica A:Statistical Mechanics and its Applications479, pp. 411-421.

Liu, Chang et al. (2005). "Beyond concern–a privacy-trust-behavioral intention model of electronic commerce". In:Information & Management42.2, pp. 289-304.

Kephart, Jeffrey O and Steve R White (1992). "Directed-graph epidemiological mod-els of computer viruses". In:Computation: the micro and the macro view. WorldScientific, pp. 71-102.

Zhang, Xi and Krishna Chaitanya Tadi (2007). "Modelling virus and antivirus spread-ing over hybrid wireless ad hoc and wired networks". In:Global Telecommunica-tions Conference, 2007. GLOBECOM'07. IEEE. IEEE, pp. 951-955.

Chen, Thomas M and Jean-Marc Robert (2004). "The evolution of viruses and worms".In:Statistical methods in computer security1.

Heartfield, Ryan, George Loukas, and Diane Gan (2016). "You are probably not the weakest link: Towards practical prediction of susceptibility to semantic social engineering attacks". In:IEEE Access4, pp. 6910-6928

Sukwong, Orathai, H Kim, and J Hoe (2010). "An empirical study of commercial antivirus software effectiveness". In:Computer44.3, pp. 63-70.

Brett, Terry et al. (2019). "The spreading of computer viruses on time-varying net-works". In:arXiv preprint arXiv:1901.02801.

Perra, Nicola et al. (2012). "Activity driven modeling of time varying networks". In:Scientific reports2, p. 469.

Cisco. 2019. What Are the Most Common Cyberattacks?. [ONLINE] Available at: https://www.cisco.com/c/en/us/products/security/common-cyberattacks.html. [Accessed 11 January 2019].