EasyChair Preprint

№ 9827

Internet Threats and Ways to Protect Against Them: a Brief Review

Saken Mambetov, Yenlik Begimbayeva, Serik Joldasbayev and Gulnur Kazbekova

March 6, 2023

Internet threats and ways to protect against them: A brief review

1st Saken Mambetov
Doctoral student
Al-Farabi Kazakh National University
Almaty, Kazakhstan
https://orcid.org/0000-0002-7249-5378
mambetov.saken@gmail.com

2nd Yenlik Begimbayeva
Associate Professor of the Department of Information Systems
Al-Farabi Kazakh National University
Almaty, Kazakhstan
https://orcid.org/0000-0002-4907-3345
enlik_89@mail.ru

3rd Serik Joldasbayev
Senior Lecturer of the Department of Computer Engineering
International IT University
Almaty, Kazakhstan
https://orcid.org/0000-0002-8689-1822
serykjoldasbaev@mail.ru

4th Gulnur Kazbekova
Associate Professor of the Department of Computer Engineering
Khoja Akhmet Yassawi International Kazakh-Turkish University
Turkistan, Kazakhstan
https://orcid.org/0000-0002-2756-7926
gulnur.kazbekova@ayu.edu.kz

***Abstract*— Since the 21st century is the century of information technology, there are now a large number of Internet users, and the number of these users is steadily growing. Most people know the useful aspects of this network, but there are dangerous aspects of the Internet, which are reviewed in this paper. Today there are many types of threats such as phishing attacks, spam messages, malware, worms, spyware, Trojans, rooters, botnets. The main principle of protection against these threats is information and cyber literacy of using the Internet. The main purpose of this article is a summary of Internet threats based on a review of articles by previous researchers, as well as the definition of types of threats and attacks. As a result of the study, threats were filtered into two main types: threats of social engineering, where the main emphasis is on the information carrier, and technical threats, where various methods, algorithms and software implementations are used to hack directly into a computing device containing information.**

Keywords—Internet threats, social engineering, spam, phishing, ransomware, computer worm, trojan.

I. Introduction

At the early stages of the development of computer technology, in the 70-80-ies of the last century, the main goal of the creators of viruses was self-knowledge or to annoy the employer a little. And no one could have imagined that the end of this initiative would lead to the negative term «hacker». For example, «ILoveYou» is a virus launched in 2000 and left a mark in history and even got into the Guinness Book of Records as the most destructive computer virus in the world, the damage of which was estimated to be up to 15 billion US dollars [1]. Over the years, malware developers, being professional programmers, have focused on the commercial direction. This direction not only brings a good income, but also contributed to the creation of large communities such as Anonymous [2]. This means that attacks on a global scale and sophisticated attack scenarios allow many attackers to avoid prosecution by law enforcement agencies. If we reference to the data published on the portal Digital 2022: Global Overview Report – DataReportal [3] in January 2022, the number of people on the planet today is 7.91 billion people, and 67.1% of them are connected to the Internet, that is, 5.31 billion people are Internet users. According to the predictions of Cybersecurity Ventures analysts, by 2025, the global damage from cybercrime will amount to 10.5 trillion US dollars [4]. In this regard, protection from Internet threats is becoming more relevant every year, and the development of ways to reduce costs is assigned to qualified specialists.

II. Main Part

Threats from the Internet conditionally can be divided into two types.

1. ***Technical threats:*** malware, botnets, Dos and DDoS attacks, worms, Trojans, rootkits.
2. ***Social engineering:*** spam, phishing attacks, ransomware.

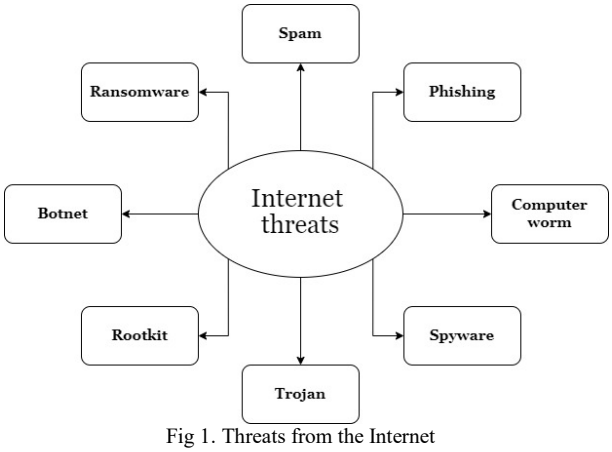

Fig 1. Threats from the Internet

Let's consider some of the types.

A. Review of works

One of the most common activities in the network is the distribution of unwanted messages, spam. Spam can be in the form of advertising or schemes for quick repayment of debt, quick enrichment, online dating or health-related

products. [5] Describes several methods for automatic detection of spam messages. The paper uses a pre-formed model of the BERT transformer (representations of bidirectional encoder converters). In addition, spam classification was carried out using machine learning and deep learning methods.

The article [6] describes another common way to combat spam — filtering. Mail filtering of incoming email is mainly based on the removal of incoming and outgoing mail, spam and computer viruses. Five of the most commonly used machine learning methods were used to filter unnecessary messages: Naive Bayes classifier, support vector machine, artificial neural network (ANN), KNN, classification of artificial immune system.

The paper [7] proposes a spam detection methodology based on the architecture of deep learning in the context of natural language processing (NLP). NLP uses a text representation. Various email representations are used to convert emails into email word vectors. In addition, based on the established hyperparameters, many deep learning architectures and optimal parameters for sending e-mail are determined.

The paper [8] presents a brief introduction to social spam, the process of sending spam and the taxonomy of social spam, as well as how spam in social networks causes polarization of moods, affects the time of user interaction on the network, reduces the quality of available information, network bandwidth, computing power, data exchange rate and other secondary indirect harmful effects spam.

When it comes to viewing and sharing video content, YouTube is ahead of the whole planet. However, there may be a hidden world behind the video content. The video may not match the title of the video or contain content unfavorable for viewing. The paper [9] presents an approach of the Markov decision-making process for modeling the problem of detecting video spam on YouTube.

Another threat is phishing attacks. Phishing is a type of Internet fraud, the main purpose of which is to gain access to personal information, usernames and passwords by sending electronic messages on behalf of government employees, bank employees, representatives of well—known brands, etc., by methods entering into the confidence of a potential victim. There are a large number of publications in this area, one of which is the work [10], which provides an overview of currently known phishing methods, as well as an overview of the current state of phishing. As a result of the study, the authors classified phishing attacks by the main mechanisms and by countermeasures, without taking into account the phishing lifecycle. The article also presents a new complete anatomy of phishing, including attack stages, types of attackers, vulnerabilities, threats, targets, attack tools and attack methods, and suggests new areas of security measures.

In [11], an approach is presented to overcome the «uncertainty» in the traditional risk assessment of phishing websites, and a model for detecting phishing websites based on Fuzzy Logic operators is also proposed, which are used to describe factors and indicators of phishing websites in the form of fuzzy variables with indicators of criteria for measuring phishing attacks of websites with a multi-level structure.

In [12], new features of the presented approach include sequences of URL characters without prior phishing information, various hyperlink information, and the text content of a federated web page for training the XGBoost classifier.

In [13] presents a systematic review of the literature, including all relevant studies of phishing and human influence on phishing attacks. The research shows how to take into account the human factor to protect against phishing attacks.

The article [14] presents a web architecture for predicting the presence of phishing for a given web address based on machine learning models such as Random Forest, classification trees and Support Vector Machine. In addition, 2 additional models based on web addresses with a previously processed symbol search module are used. Since everything is done in the API, any user can use any model that he considers correct to predict the presence of phishing.

Another well-established threat from the Internet is a type of software called Ransomware. The paper [15] describes in detail about this type of threat, and to detection such programs, the methods are given, optimization-based Deep Recurrent Neural Network (Deep RN) and Flame Water Moth (WMFO)

This article [16] presents the «susceptible-infected-recovered-dead» (SIRD) computer worm model. In addition, the mechanisms and features of the spread of helminths were analyzed. A model of simple differential equations was used to simulate the spread of computer worms. Unknown parameters of the SIRD model were estimated using the least squares method, Monte Carlo methods with Markov chains and Kalman ensemble filtering (ENKF).

B. Protection against threats from the Internet

In the course of this article, we saw many types and methods of attacks on the Internet. However, there are enough ways to protect against them. When working on the Internet, we recommend following the following requirements:

1. Use a password or passphrase. The password must be at least 8 characters long. When setting a password, it is best to use uppercase and lowercase characters, numbers and special characters. The password should not repeat previously set passwords, and also should not include information that can be guessed, such as date of birth, name, phone number. Today, in many places, passphrases are used instead of passwords.
2. Work under an account with limited rights on the computer. It is better to work with a user account for daily work in the operating system. The user account has all the same rights as the administrator account, except that it requires an administrator password when making changes to the operating system or installing new programs. This action reduces the risk of automatically installing malware on your computer, changing system settings, or accidentally deleting it.
3. Use data encryption. By encrypting your data, you can protect the information you need from other users. Today, special cryptographic programs encode in such a way that only a user who knows the key needed for decryption can read it. Operating systems also have built-in encryption tools. For example, in Windows 10, BitLocker

Drive Encryption is used to protect all files, stored on the operating system disk and internal hard drives, and BitLocker To Go is used to protect files stored on external hard drives, USB devices.

4. Keep your software up to date. The operating system and the programs you use should be updated regularly and on time. If you enable automatic update mode, all software will be installed in the background. If you are performing the update yourself, then you need to download the programs, which require updates from the official websites of software manufacturers.

5. Use antivirus programs and update them on time. It is necessary to protect the operating system from online threats coming from the Internet. In addition, antivirus is a key component of malware protection. Currently, antivirus programs automatically update antivirus databases. Such programs scan the necessary areas of the system and monitor all possible ways of virus penetration, such as mail applications and potentially dangerous websites, without interfering with the user's work. Antivirus should always be turned on: it is not recommended to turn it off. Regularly check all removable media (optical drives, flash memory, SSD drives) for viruses.

6. Use a firewall. The main function of the firewall is to monitor network packets passing through the firewall in accordance with the specified rules. In general, it monitors the processes of data exchange between the device and the Internet. It also checks all data received or sent from the network. If necessary, it blocks network attacks and prevents the secret sending of personal data over the Internet. The firewall does not allow suspicious information to pass through and does not release important information from the system.

C. Research results

As a result of the research, we found out the most common Internet threats in recent times. In particular, their statistics were kept. Information security statistics are maintained by many reputable resources. We received today's data from the «Kaspersky Laboratory» study [17]. Information on Kazakhstan is provided on the basis of KZ-CERT data [18]. But, even on the Kaspersky Lab website, it says that all the data are approximate calculations based on published cyberattacks. Many violations and extortion remain hidden because many companies want to protect their reputation. Therefore, the real numbers of attacks are available and much «sadder» than publicly available.

TABLE I. ATTACKS VIA WEB RESOURCES OVER THE PAST 5 YEARS

TITLE	ATTACKS VIA WEB RESOURCES	DISTRIBUTION OF WEB ATTACK SOURCES BY COUNTRY TOP-3	DISTRIBUTION OF WEB ATTACK SOURCES BY KAZAKHSTAN
2018	1876998691	United States-45.65% Netherlands-17.53% Germany-11.70%	19360
2019	975491360	United States-43,25% Netherlands-22,23% Germany-8,34%	
2020	666809967	United States-49,48% Netherlands-13,36% France-7,20%	17145
2021	687861449	Czech-30,87% United States-24,94% Germany-7,07%	17421
2022	505879385	United States-27,50% Germany-12,68% Netherlands-12,23%	16188

Table I shows the number of attacks on users' computers from the Internet. To determine the geographical source of web attacks, it is determined by comparing the domain name with the actual IP address on which this domain is located and determining the geographical location of this IP address. Thus, the main means of Internet security can be divided into 3 large groups (Fig. 2).

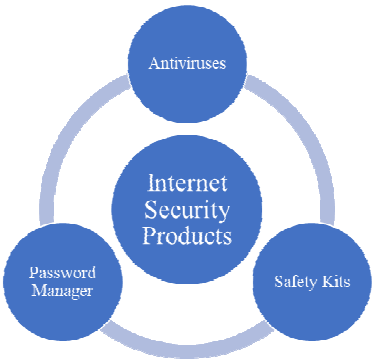

Fig 2. Internet Security Products

Antivirus software and Internet security programs help protect your device from attacks by detecting and removing malware.

Password Manager is a software that helps the user to store and organize their passwords. Password managers usually store passwords in encrypted form, requiring the user to create a master password that provides access to a database of all passwords. However, using a password manager may compromise the user's security.

Security packages were first offered by McAfee in 2003. Firewalls, antivirus programs, antispyware programs and much more. it consists of a set of programs. They also offer theft protection, portable storage security verification, private web browsing, cloud spam protection, file destruction, or security solutions (pop-up responses).

III. CONCLUSION

This review examines the main types of threats from the Internet, as well as developments of scientists and specific approaches, requiring minimal human intervention to ensure safety. A wide range of security ideas were also considered and the autonomy of security systems. In particular, this

work examines approaches to threat mitigation using an autonomous taxonomy and outlines the future direction of work.

REFERENCES

[1] P. Knight, "Iloveyou: Viruses, paranoia, and the environment of risk," The Sociological Review, vol. 48, no. 2_suppl, pp. 17–30, 2000. [Online]. Available: https://doi.org/10.1111/j.1467-954X.2000.tb03518.x

[2] Arruda, José PedroHackers, Hacktivistas e Whistleblowers: o caso português.. Perspectivas em Ciência da Informação [online]. 2021, v. 26, n. 02 [Acessado 20 Outubro 2022] , pp. 15-36. Disponível em: <https://doi.org/10.1590/1981-5344/3869>. Epub 23 Jul 2021. ISSN 1981-5344. https://doi.org/10.1590/1981-5344/3869

[3] Digital 2022: Global Overview Report – DataReportal, Internet resources [Data 10.02.2022] https://datareportal.com/reports/digital-2022-global-overview-report

[4] Cybersecurity Ventures Internet resources [Data 10.02.2022] https://cybersecurityventures.com/boardroom-cybersecurity-report/

[5] 2 Isra'a AbdulNabi, Qussai Yaseen, «Spam Email Detection Using Deep Learning Techniques», Procedia Computer Science, Volume 184, 2021, Pages 853-858, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2021.03.107

[6] 3 Omar Saad , Ashraf Darwish, Ramadan Faraj, «A survey of machine learning techniques for Spam filtering», IJCSNS International Journal of Computer Science and Network Security, VOL.12 No.2, February 2012.

[7] 4 Srinivasan, S., Ravi, V., Alazab, M., Ketha, S., Al-Zoubi, A.M., Kotti Padannayil, S. (2021). «Spam Emails Detection Based on Distributed Word Embedding with Deep Learning». In: Maleh, Y., Shojafar, M., Alazab, M., Baddi, Y. (eds) Machine Intelligence and Big Data Analytics for Cybersecurity Applications. Studies in Computational Intelligence, vol 919. Springer, Cham. https://doi.org/10.1007/978-3-030-57024-8_7

[8] 5 Sanjeev Rao, Anil Kumar Verma, Tarunpreet Bhatia, «A review on social spam detection: Challenges, open issues, and future directions» Expert Systems with Applications, Volume 186, 30 December 2021, 115742, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2021.115742

[9] 6 S. Kanodia, R. Sasheendran and V. Pathari, «A Novel Approach for Youtube Video Spam Detection using Markov Decision Process», 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 60-66, doi: 10.1109/ICACCI.2018.8554405.

[10] 7 Alkhalil Z, Hewage C, Nawaf L and Khan I (2021) «Phishing Attacks: A Recent Comprehensive Study and a New Anatomy». Front. Comput. Sci. 3:563060. doi: 10.3389/fcomp.2021.563060

[11] 8 M. Aburrous, M. A. Hossain, F. Thabatah and K. Dahal, «Intelligent Phishing Website Detection System using Fuzzy Techniques», 2008 3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008, pp. 1-6, doi: 10.1109/ICTTA.2008.4530019.

[12] 9 Aljofey, A., Jiang, Q., Rasool, A. *et al.* «An effective detection approach for phishing websites using URL and HTML features». *Sci Rep* **12**, 8842 (2022). https://doi.org/10.1038/s41598-022-10841-5

[13] 10 Giuseppe Desolda, Lauren S. Ferro, Andrea Marrella, Tiziana Catarci, and Maria Francesca Costabile. 2021. «Human Factors in Phishing Attacks: A Systematic Literature Review». ACM Comput. Surv. 54, 8, Article 173 (November 2022), 35 pages. https://doi.org/10.1145/3469886

[14] 11 J. Lamas Piñeiro and L. Wong Portillo, «Web architecture for URL-based phishing detection based on Random Forest, Classification Trees, and Support Vector Machine», *ia*, vol. 25, no. 69, pp. 107–121, May 2022. https://doi.org/10.4114/intartif.vol25iss69pp107-121.

[15] Nalinipriya, G, Balajee, M, Priya, C, Rajan, C. «Ransomware recognition in blockchain network using water moth flame optimization-aware DRNN». Concurrency Computat Pract Exper. 2022; Volume 34, Issue 19, Article number e7047. doi:10.1002/cpe.7047

[16] Deng, Yuea; Pei, Yongzhen; Li, Changguoc, "Parameter estimation of a susceptible–infected–recovered–dead computer worm model", Simulation, Volume 98, Issue 3, pp 209 – 220, March 2022, https://doi.org/10.1177/00375497211009576

[17] Kaspersky Security Bulletin 2018-2022. Statistics. SecureList by Kaspersky [Data 27.12.2022] https://securelist.com/ksb-2022-statistics/108129/

[18] KZ-CERT Incidents statistics [Data 27.12.2022] https://cert.gov.kz/press_club/infographics