Romanian Association for
Information Security Assurance

PROCEEDINGS
OF
THE INTERNATIONAL CONFERENCE ON CYBERSECURITY AND CYBERCRIME

Volume X
eISSN 2393-0837

CyberCon Romania
2023

THE INTERNATIONAL CONFERENCE ON CYBERSECURITY AND CYBERCRIME

PROCEEDINGS
OF
THE INTERNATIONAL CONFERENCE ON CYBERSECURITY AND CYBERCRIME

Volume X

A scientific conference organized by the
Romanian Association for Information Security Assurance

CyberCon Romania
2023

THE INTERNATIONAL CONFERENCE ON CYBERSECURITY AND CYBERCRIME

The International Conference on Cybersecurity and Cybercrime (IC3) is an annual scientific conference, with the purpose to encourage the exchange of ideas about the evolution of cyberspace, information security challenges, and new facets of the phenomenon of cybercrime. The event provides the appropriate framework for experts to present their research in this field.

The Proceedings of the International Conference on Cybersecurity and Cybercrime includes scientific papers reviewed by the *Editorial Board* that consists of experts from the academic field, their work taking place under the guidance of the *Advisory Board*, composed of internationally recognized personalities from the academic field.

Proceedings of the International Conference on Cybersecurity and Cybercrime
Online ISSN: 2393-0837
Print ISSN: 2393-0772
DOI: 10.19107/CYBERCON
URL: proceedings.cybercon.ro

The International Conference on Cybersecurity and Cybercrime is part of the **CyberCon Romania** event, organized by the Romanian Association for Information Security Assurance.

CyberCon Romania brings together experts from public institutions, private companies, and universities, for raising the level of awareness and embodies the cybersecurity culture.

Website: www.cybercon.ro

The Romanian Association for Information Security Assurance (RAISA) is a professional, non-governmental, non-partisan political, nonprofit, and public benefit association.

Founded in 2012, the association started as an initiative with the aim of promoting and supporting information security activities in compliance with applicable laws and creating a community for the exchange of knowledge between the experts from the public, private, and academic environment. Its vision is to encourage the cybersecurity research and education, and to contribute to the creation and dissemination of knowledge and technology in this domain.

Website: www.raisa.org

CONFERENCE COMMITTEES

CONFERENCE CHAIRMAN

Professor **Ioan C. BACIVAROV**, PhD
University Politehnica of Bucharest, Romania
Faculty of Electronics, Telecommunications and Information Technology

INTERNATIONAL SCIENTIFIC COMMITTEE

Professor Emeritus **Alessandro BIROLINI**, PhD
ETH Zurich, Switzerland

Professor **Angelica BACIVAROV**, PhD
University Politehnica of Bucharest, Romania

Irina BAKHAYA, PhD
"Al. I. Cuza" Police Academy, Romania

Natalia BELL, D.Sc.
Marymount University, United State of America

Professor **Răzvan BOLOGA**, PhD
University of Economic Studies, Romania

Viorel GAFTEA, PhD
Romanian Academy, Romania

Alexandru GEORGESCU, PhD
National Institute for R&D in Informatics, Romania

Angela IONIȚĂ, PhD
Research Institute for Artificial Intelligence, Romania

Assoc. Prof. **Dumitru-Iulian NĂSTAC**, PhD
University Politehnica of Bucharest, Romania

Cristian PAȚACHIA-SULTĂNOIU
Orange, Romania

Pierluigi PERRONE, PhD
LUISS University, Rome, Italy

Professor **Florin POPESCU**, PhD
"Carol I" National Defence University, Romania

Assoc. Prof. **Eduard-Cristian POPOVICI**, PhD
University Politehnica of Bucharest, Romania

Professor **Răzvan RUGHINIȘ**, PhD
University Politehnica of Bucharest, Romania

Professor **Anurag SHARMA**, PhD
GNA University, India

Assoc. Prof. **Emil SIMION**, PhD
University Politehnica of Bucharest, Romania

Professor **Pradeep Kumar SINGH**, PhD
University of Information Technology, India

Professor **Sandeep TIWARI**, PhD
Amity University, India

Professor **Dănuț TURCU**, PhD
"Carol I" National Defence University, Romania

Vlad-Alexandru VOICESCU, PhD
"Al. I. Cuza" Police Academy, Romania

ORGANIZING COMMITTEE

Sabina-Daniela AXINTE, PhD
University Politehnica of Bucharest, Romania

Larisa GĂBUDEANU, PhD
Babeș-Bolyai University, Romania

Ioan-Cosmin MIHAI, PhD
"Al. I. Cuza" Police Academy, Romania

Gabriel PETRICĂ, PhD
University Politehnica of Bucharest, Romania

TABLE OF CONTENTS

Exploring a Diplomatic System of Cooperation in the Cyber Space through a Proposed Cyber Diplomacy Cooperation Framework

Natalia BELL, Alex MBAZIIRA
School Technology and Innovation,
Marymount University, Arlington, Virginia 22207, United States of America
nbell@marymount.edu, ambaziir@marymount.edu

Abstract

Cyberattacks are on the rise, and cyber weapons are the main tools used in modern warfare. All these occurrences are changing the nature of traditional diplomacy, contributing to developing new avenues for Cyber Diplomacy. The world's leading nations have realized the importance of establishing a diplomatic system of collaboration in the cyber sphere to facilitate bilateral relationships between nations and cooperation in cyberspace in already-established alliances such as NATO, the United Nations, and regional trade associations. Multiple studies have discussed and detailed the concept of "cyber diplomacy" and the diplomatic behavior associated with it; however, few of these analyses have sought to distinguish the "cyber diplomacy" concept from the more traditional and well-known concept of "diplomacy." The scope of this proposal is to create a Cyber Diplomacy Cooperation Framework which will bring together conventional elements of diplomacy and cutting-edge cybersecurity mechanisms. As cyber warfare concerns are growing, nations need a normative cyber diplomacy framework that can be adapted by countries to prevent cyber-crises and engage more nations in the discussion.

Index terms: cybersecurity, cyber diplomacy, framework

1. Background

In September 2022, the US Senate unanimously confirmed Mr. Nathaniel Fick to serve as the first-ever cyber ambassador-at-large for cyberspace and digital policy. Mr. Fick will run the newly established State Department’s Bureau of Cyberspace and Digital Policy. According to National Geographic (2022), records of diplomatic letters date back to the 14th century B.C. The novelty of the notion of "cyber diplomacy" relates to relationships in the digital environment as opposed to the physical space. As cyber-attacks continue to increase and cyber weapons transform modern warfare, all these events are transforming diplomacy by creating new avenues for Cyber Diplomacy (Halpern, 2019). For example, the Russian and Ukrainian conflict of 2022, which is involving both state and non-state actors participating as proxies in this hybrid war has rendered existing diplomatic tools ineffective and caused a need for a new framework of cyber diplomacy which can be effective in averting future cyber geopolitical crises (EU-Cyber Direct, 2022). Some of the key issues in modern cyber hybrid warfare which a cyber diplomacy framework needs to address include information operations, cyber-attacks, political manipulation, engagement of involved non-state actors, among others. There seems to be no agreeable term to define diplomatic tools and processes for cyberspace despite a growing need in national states to extend diplomacy to cyberspace. According to Diplo, a Swiss-Maltese nonprofit organization that focuses on capacity development in the area of digital

policy and Internet governance, the terms “cyber diplomacy” and “digital diplomacy” are often interchanged. Cyber diplomacy is associated with diplomatic efforts to address cyber security challenges while digital diplomacy is used to define the adoption of new tools in diplomatic practice, such as social media, websites, and online meeting platforms, as well as the implementation of digital foreign policy, including recent issues on the diplomatic agenda (diplomacy.edu). Furthermore, the State Department uses the terms “digital diplomacy” and “digital policy,” to refer to “responsible state behavior in cyberspace and advance policies that protect the integrity and security of the infrastructure of the Internet” (state.gov). While the European Commission widely uses the term “Digital Agenda" in its essential internet-related documentation, the Council of the European Union clearly articulates “Cyber Diplomacy” in its Outcome of Proceedings, “Council Conclusions on Cyber Diplomacy,” coming from the General Secretariat of the Council to the Delegations (consilium.europa.eu, 2015). Other adjectives and prefixes like "tech," "net," "virtual," and "e-" diplomacy are informally used; however, Diplo suggests that such trends appear to be confusing discussions and policies surrounding this topic (diplomacy.edu). Several studies have talked about and outlined the concept of "cyber diplomacy" and the diplomatic behavior that goes along with it, yet few of them have attempted to differentiate the “cyber diplomacy” notion from the more traditional and well-known concept of "diplomacy." As Attatfa et al. (2020) note, there is a considerable gap in the literature - a subject for future research.

2. Introduction

According to Attatfa et. al. (2020), cyber diplomacy began in 2007, a year that will always be recognized because of a large-scale cyberattack on Estonia. Since then, Europe has actively sought cyberspace security. In 2017, the European Council of the European Union agreed to develop a framework for a “joint EU diplomatic response to malicious cyber activities”: the cyber diplomacy toolbox (consilium.europa.eu, 2017). The Paris Call for Trust and Security in Cyberspace, launched at the 2018 Paris Peace Forum, has emerged as the multi-actor framework of reference for promoting core principles for the safety of cyberspace. NATO did not remain behind. In 2019, Jens Stoltenberg, NATO Secretary General, commented that “a serious cyberattack could trigger Article 5, where an attack against one ally is treated as an attack against all” (NATO, 2019). Due to the issue's importance, European countries started to make and use national cybersecurity strategies based on European institutions' cyber initiatives. The US also created several programs, such as the White House Office of the National Cyber Directorate, the 2021 President's Executive Order on Improving the Nation's Cybersecurity, the State and Local Government Cybersecurity Act of 2021, and agencies like the Cybersecurity and Infrastructure Security Agency, and, most recently, the State Department's Bureau of Cyberspace and Digital Policy and its first ambassador-at-large.

Alongside Europe and the United States, several other regions and countries have recognized the urgency to govern cyberspace. Moreover, the world’s great powers have come to recognize the need to establish a diplomatic system of collaboration in the cyber realm, to serve both individual countries' cooperation as well as through already established alliances such as NATO, the UN, etc.

This paper aims to introduce a Cyber Diplomacy Cooperation Framework that will incorporate traditional diplomacy and cooperation aspects as well as cybersecurity components in a novel manner. The objective is not to create a brand new framework but rather to adapt existing frameworks and extend them to address cybersecurity elements that have not been previously examined.

3. Proposed Cyber Diplomacy Cooperation Framework

The proposed Cyber Diplomacy Framework will follow the UN Sustainable Development Cooperation Framework's adapted key objectives:

1) Address national priorities and gaps in their pathway towards meeting their cybersecurity goals;
2) Must embody the spirit of partnership;
3) Collective promise to leave no one behind;
4) Responses to a countries' specific needs and realities;

To develop the framework, we will assess the short-comings of existing diplomatic tools in averting cyber warfare, which involves both state and non-state actors, information operations and political manipulation targeting civilians, loop-holes in international law exploited by aggressor nation-state and non-nation state actors to engage in cyberwar among others.

Traditional diplomacy covers various forms of cooperation, such as bilateral and multilateral agreements, regional pacts, international organizations that facilitate collaboration between nations, and strategic partnerships. Nonetheless, the emergence of cyber diplomacy requires the incorporation of new components. For example, addressing global cybercrime issues and fostering international cooperation for cyber security offense and defense are essential factors to consider. In addition, there is an increasing need to impose sanctions for criminal activities on a global scale, while the concept of cyber attribution has acquired significant importance in modern times, particularly in the context of cyber-attacks by foreign entities. Furthermore, creating new avenues for cybersecurity research and collaboration on a global scale could be a crucial element of the cyber diplomacy framework.

The US and NATO are ahead of most countries around the world on cyber-diplomacy. For example, the US House of Representatives recently passed the bill for Cyber Diplomacy Act of 2021, which is awaiting to be presented and passed by the Senate before being signed into law by the US President (McCaul, 2021). Given the resources and talent of the US and NATO in cybersecurity and information technology, the objective of our study is to develop a generalizable framework that can be adapted by countries desiring to develop cyber diplomacy programs.

The proposed Cyber Diplomacy Farmwork is based on three main pillars: Capacity Building, Cooperation and Trust. The Capacity Building refers to the process of countries working together to share resources, provide assistance to one another in the form of training and education, and provide technical assistance. The nations that participate in diplomatic cyber cooperation would pool their resources, provide one another with technological help, and collaborate on educational and training initiatives.

Cooperation is the second pillar, and it refers to the countries' willingness to build and adhere to a common agenda in terms of the many areas of cybersecurity. It also refers to the countries' willingness to develop various partnerships between governmental institutions, as well as private-public partnerships. In addition, the cybersecurity industry frequently depends on the community for the sharing of cyber threat intelligence as well as open-source intelligence regarding cyber dangers. Countries will coordinate their efforts to collaborate on a variety of fronts under the overarching concept of collaboration, including research and development.

The final pillar that we propose to incorporate into this structure is Trust or confidence. The concept of transparency, which refers to being open and honest with cybersecurity policies, procedures, and operations, is one of the three factors that would be considered to come under the Trust pillar. Within the context of a cyber diplomacy cooperation environment, it is anticipated that the countries will construct and continue to maintain trust and confidence in each other's practices. The second part of the cyber diplomacy collaboration is the diplomatic engagement. Just as in conventional diplomacy, the diplomatic engagement in the cyber diplomacy cooperation would strive for appropriate conflict resolution methods and bilateral and multilateral dialogues regarding a variety of topics pertaining to cybersecurity. And last but not least is Accountability, which is the concept under which we propose shared sanctions norms to ensure proper attribution of cyber-attacks.

Fig. 1. Proposed Cyber Diplomacy Framework

4. Conclusion

As the cyber threat landscape continues to evolve and the attack surface continues to widen to integrate more globally interconnected devices over the Internet. There is a need for a normative cyber diplomacy framework that can be adapted by countries to avert cyber-crises and also involve more countries in dialogue as cyberwar continues to evolve. In this paper we proposed a Cyber Diplomacy Farmwork founded upon three pillars: Capacity Building, Cooperation, and Trust. The process of countries working together to share resources, providing aid to one another in the form of

training and education, and provide technical assistance is referred to as capacity building. Cooperation is the second pillar, and it refers to the countries' commitment to construct and adhere to a shared agenda in terms of the many areas of cybersecurity. Some examples of these areas include knowledge exchange in research and development, open-source intelligence and threat intelligence partnerships, and collaborative agendas. Trust, which includes dispute resolution, bilateral and multilateral discussion, transparency, and accountability, is the third and final pillar that we propose to add into this system.

Despite the fact that we are aware that this framework is capable of significant advancement, our primary objective was to present an exploratory viewpoint on a topic that is now under development. Furthermore, the intention behind this proposal is to make the topic accessible to possible recommendations, which will then be incorporated into subsequent works.

References

[1]. Attatfa A., Renaud K., and Paoli S., "Cyber Diplomacy: A Systematic Literature Review," Procedia Comput Sci. 2020;176:60-69. doi: 10.1016/j.procs.2020.08.007. Epub 2020 Oct 2. PMID: 33042293; PMCID: PMC7531992.

[2]. consilium.europa.eu (2015, February 11). Cyberattacks: The EU ready to respond with a range of measures, including sanctions. Retrieved from https://www.consilium.europa.eu/en/press/press-releases/2017/06/19/cyber-diplomacy-toolbox/

[3]. consilium.europa.eu (2017, June 19). Council Conclusions on Cyber Diplomacy Retrieved from https://data.consilium.europa.eu/doc/document/ST-6122-2015-INIT/en/pdf

[4]. diplomacy.edu. (2022). Digital diplomacy. Retrieved from www.diplomacy.edu: https://www.diplomacy.edu/topics/digital-diplomacy/

[5]. diplomacy.edu. (2022). FAQ about Diplomacy. Retrieved from https://www.diplomacy.edu/: https://www.diplomacy.edu/

[6]. europa.eu. (2022). Shaping Europe's digital future. Retrieved from https://digital-strategy.ec.europa.eu/en

[7]. National Geographic. (2022). Diplomacy. Retrieved from https://education.nationalgeographic.org/resource/diplomacy

[8]. NATO. (2019, August 29). Article by NATO Secretary General Jens Stoltenberg published in Prospect's new cyber resilience supplement. Retrieved from https://www.nato.int/cps/en/natohq/news_168435.htm?selectedLocale=en

[9]. Paris Peace Forum. (2018). The Paris Call for Trust and Security in Cyberspace. Retrieved from https://parispeaceforum.org/en/initiatives/the-paris-call-for-trust-and-security-in-cyberspace/

[10]. state.gov. (2022). Bureau of Cyberspace and Digital Policy. Retrieved from https://www.state.gov/:https://www.state.gov/bureaus-offices/deputy-secretary-of-state/bureau-of-cyberspace-and-digital-policy/

[11]. EU-Cyber Direct. (2022). Is War in Ukraine the End of Cyber Diplomacy? » directions blog. https://directionsblog.eu/is-war-in-ukraine-the-end-of-cyber-diplomacy/

[12]. Halpern, S. (2019, July 18). How Cyber Weapons Are Changing the Landscape of Modern Warfare. The New Yorker. https://www.newyorker.com/tech/annals-of-technology/how-cyber-weapons-are-changing-the-landscape-of-modern-warfare

[13]. McCaul, M. T. (2021, April 22). Text - H.R.1251 - 117th Congress (2021-2022): Cyber Diplomacy Act of 2021 (2021/2022) [Legislation]. http://www.congress.gov/

Artificial Intelligence to Counter Cyber-Terrorism

Serena BIANCHI[1], Marina MANCUSO[2], Caterina PATERNOSTER[3], George KALPAKIS[4], Theodora TSIKRIKA[5], Stefanos VROCHIDIS[6], Denitsa KOZHUHAROVA[7], Bernhard JAEGER[8]

[1] Research Department, SYNYO GmbH, Vienna, Austria
serena.bianchi@hotmail.it
[2] Transcrime Università Cattolica del Sacro Cuore, Milan, Italy
marina.mancuso@unicatt.it
[3] Transcrime Università Cattolica del Sacro Cuore, Milan, Italy
caterina.paternoster@unicatt.it
[4] Information Technologies Institute, Center for Research and Technology Hellas, Thermi-Thessaloniki, Greece
kalpakis@iti.gr
[5] Information Technologies Institute, Center for Research and Technology Hellas, Thermi-Thessaloniki, Greece
theodora.tsikrika@iti.gr
[6] Information Technologies Institute, Center for Research and Technology Hellas, Thermi-Thessaloniki, Greece
stefanos@iti.gr
[7] Law and Internet Foundation, Sofia, Bulgaria
denitsa.kozhuharova@netlaw.bg
[8] Research Department, SYNYO GmbH, Vienna, Austria
bernhard.jaeger@synyo.com

Abstract

This paper discusses the role of disruptive and innovative technologies for countering the spread of terrorist online content (TCO). In particular, it focuses on the use of Artificial Intelligence (AI) in support to Host Service Providers (HSPs) and Law and Enforcement Agencies (LEAs). The violent and terrorist content is more and more disseminated online taking advantages of the opportunities offered by Internet. The diffusion of terrorist propaganda has a negative impact on the civil society and poses several risks. For this reason, the European institutions published in 2021 the Regulation (EU) 2021/784 to address the misuse of hosting services for the dissemination to the public of TOC. It regulates the measures to be applied by HSPs and Member States' authorities in order to identify and ensure the quick TOC removal and to facilitate cooperation with each other and Europol. In order to be compliant with these dispositions, AI-based disruptive technologies can provide LEAs and HSPs, especially the small and micro-ones, a concreate support. The implementation of the Regulation and the use of AI technologies have legal and ethical implications that have to be considered. The paper is based on the work and preliminary research conducted in the framework of the European funded project ALLIES, "AI based framework for supporting micro and small Hosting Service Providers (HSPs) on the report and removal of online terrorist content", Grant Number 101080090.

Index terms: Artificial Intelligence, Counter Extremism, Cyberterrorism, Ethical and Legal Framework, Online Radicalisation

1. Introduction

Online media channels deeply changed the way how people communicate, work, interact and live. Together with innovation and unprecedented potentialities, they also introduced new threats into our society, changing the modus operandi and the structures of terrorist organisations.

As highlighted by the European Parliament (2021), the dissemination of terrorist content is one of the most widespread and most dangerous forms of misuse of online services in the field of internal security [1]. Online and social media channels have been broadly used in the past years by terrorist organisations in order to spread violent content, to train, recruit and motivate individuals, and last but not least, to finance terrorist organisations [2, 3, 4].

Similarly as for "terrorism", the definition of "cyber-terrorism" is still very controversial. Some authors define cyber-terrorism as the premeditated use of disruptive activities, or the threat thereof, against computers and/or networks, with the intention to cause harm or further social, ideological, religious, political or similar objectives, or to intimidate any person in furtherance of such objectives. [5].

In this paper, the authors will consider cyber-terrorism in broader terms, including also the capacity of carrying out cyber-related crimes, such as the **spread of online violent content**, training and recruiting of individuals and terrorist financing (following the Directive (EU) 2017/541) [6].

In particular, the focus is given to the spread of online terrorist generated content, having as scope one or more of the following actions (according to Art.2(7) of the Regulation (EU) 784/2021):

a) inciting or advocating, including by glorifying, the commission of terrorist offences, thereby causing a danger that such acts be committed.
b) encouraging the contribution to terrorist offences.
c) promoting the activities of a terrorist group, in particular by encouraging the participation in or support to a terrorist group within the meaning of Article 2(3) of Directive (EU) 2017/541.
d) instructing on methods or techniques for the purpose of committing terrorist offences.

The paper is structured in four main sections. In the first one, the authors discuss the threat posed by terrorism online content (TCO) mainly for the civil society. In the second one, the TCO EU Regulation is presented in order to understand what are the main new dispositions foreseen at EU level for preventing and fighting against TCO. After that, and considering this initial framework, in Sections 3 and 4 the authors analyse two main components related to TCO: (a) the technical capabilities that innovative technologies can provide as support to the Law Enforcement Agencies (LEAs), as well to private service providers, and (b) the ethical and legal component to be considered when dealing with these topics.

2. The Threat posed by Terrorist Online Generated Content

Before diving into how innovative technologies can support the detection and removal of violent and extremist content online, it is of importance to highlight why terrorist online generated content poses such a big threat to our civil society and what are the risks related to it.

On the one hand, the establishment and vast evolution of new communication channels via the Internet opened new possibilities for the overall distribution of terrorist-generated online content. While there was previously almost no alternative to dissemination via text, radio or television, the Internet offered countless possibilities for dissemination. More specifically, livestreaming, pictures and videos are widely used as a format of disseminating terrorist-generated content. Several extremists and members of terrorist organisations adapted to the new communication channels and, starting already from the early 2000s, numerous photos and videos were distributed praising the hijackers and attackers of terrorist attacks - see the most recent attacks in Christchurch; Buffalo; or

the synagogue shooting in Halle, Germany, along with another attack on a synagogue in Poway, California; a racist attack at a Walmart in El Paso, Texas; and at a mosque in Bærum, Norway; or again the several violent videos shared by ISIS members starting from 2014.

On the other hand, the online world has opened up new opportunities for radicalisation, even for those who lack physical connections to radicalised individuals or environments, increasing the chance of self-radicalisation [7, 8]. Social media can provide an outlet for vulnerabilities that stem from offline sources, helping to compensate them. For example, online spaces are able to draw individuals who are feeling socially alienated or isolated, who see in the online world an alternative social setting where they can express their frustrations. In an environment where an abundance of information and propaganda is present, connecting with people who share similar extreme views can reinforce radical thinking by providing validation, and also legitimising the use of violence [9, 10, 11].

The spread of violent content online therefore implies the **facilitation** of **inciting terrorist activities** and **glorifying** such proceedings, by distributing various formats, such as photos, videos or texts, and consequentially, promoting subtle **ideological indoctrination** to compel others to commit terrorist acts [1]. Similarly to the commercial brand e-communication strategies, also terrorist organisations and ideologies make use of marketing tools to spread their ideas more effectively. This phenomenon is enabled and facilitated by creating the highest possible reach of a particular message through the social media channels and websites. That way the range of indirect online recruitment, which is clearly expressed through indoctrination, is increased and the pool of audience drastically enlarged. Besides social media channels extremists increasingly utilised another powerful tool for their activities during the last decade. Gaming and related platforms have grown to become some of the world's largest entertainment industries, providing extremist groups with significant opportunities for recruitment and organisation [12]. The video game Salil alSawarem (The Clanging of Swords) can be considered for instance as modern example of high-quality propaganda. It is a "first-person shooter" game that imitated the popular Grand Theft Auto franchise and was designed to gain attention for ISIS. The game's trailers were released on numerous websites and platforms, including YouTube, which had 3.5 billion views per month on gaming channels alone when the game was released in 2014. The terrorists' use of cinematic productions, social media, and the appeal of a videogame demonstrates their tactic of exploiting popular culture to make their propaganda go viral and reach their target audience [13]. Also, far-right extremists are increasingly present in online gaming, while the industry's lack of content moderation, hidden metrics, and avoidance of the issue are impeding efforts to assess and combat the problem [12]. Extremist groups are spreading abusive messages and forming relationships in games ranging from military shooters like Call of Duty to open creative environments like Roblox. According to a 2019 report by the Anti-Defamation League, 23%, reported exposure to discussions of white supremacist ideology [14].

In the last years, the amount, and consequently the detection, of TCO has sharply increased. Between July 2015 and 2018, the EU Internet Referral Unit[1] (EU IRU) received over 50,000 decisions of referrals to service provides about terrorist content on their platforms. In 2021, Twitter reported over 1.8 million accounts for violations of their terms of service in relation to the promotion of violence and extremism [15], while in 2022 Facebook removed more than 56 million pieces of content containing terrorist propaganda [16].

In this content, rapid actions using disruptive and innovative technologies become fundamental means for countering the spread of TCO.

[1] Europol established the EU IRU in 2015 to actively scan the Internet for terrorist content and then refer it to the hosting service providers, accompanied by an assessment.

3. TCO Regulation & the role of Host Service Providers (HSPs)

The European Commission carried out different consultations as part of the impact assessment SWD (2018) 409 [17], to understand the stakeholders' view on TCO matter. Results show that HSPs are mostly and noticeably damaged by the dissemination of TCO. This affects not only their reputation, but also the relations with their users, and all the relevant stakeholders, who support them in their business – among others, also payment processors.

In this continuously evolving context, where the spread of TCO is an actual emerging threat, negatively impacting on several HSPs, including micro, medium, and small enterprises, what is the role of the hosting service providers (HSPs) and who is responsible for the identification, detection and removal of TCO?

The European Commission implemented several measures for tackling the dissemination of TCO, such as the EU Internet Forum, Radicalisation Awareness Network, the European, Strategic Communications Network, and most importantly, the above-mentioned EU Internet Referral Unit (IRU).

However, after noticing that *referrals alone will not be able to achieve the necessary impact – particularly with regards [sic] to volume and speed of response*, the Commission proposed the TCO (Terrorist Content Online) Regulation, which envisages the transformation of extant co-regulation and voluntary self-policing by HSPs **into a framework of obligations** (to respond to referrals; to comply with removal orders; to implement proactive measures) bolstered by sit-up-and-take-notice sanctions. [18]

As highlighted by the EU-funded project ALLIES, the fact that these new requirements set out one hour timeframe for reaction leads to the need for a prompt adaptation and implementation of measures by the HSPs. Additionally, any HSP exposed to terrorist content as per Art. 5, para 4 of the TCO Regulation should implement the so-called specific measures, making the balance between avoiding fines and respecting the legal principles a very complicated and demanding matter for the HSPs. Also, as outlined in the points above, with the new Regulation alongside hopes for better results in the removal of TOC, significant challenges likewise arise for HSPs in meeting the new obligations and expectations. The adaptation for micro & small HSPs will be far more challenging than for larger HSPs, due to the limited capacity of capital and human resources. In this regard, in the next paragraph, the authors will analyse how Artificial Intelligence can practically support the HSPs in identifying, tackling, and removing online terrorist content.

Additionally, it should be considered, that even before the Regulation entered into force, 61 human rights organisations have officially stated their position against the new piece of legislation due to their concerns about how it will affect the freedom of expression, the rights to information and privacy, and the rule of law in general.[2] Similar concerns have been outlined by the European Data Protection Supervisor in 2019 that was mainly focused on the need for safeguards and prevention measures for conducting profiling of any kind, thus recommending human oversight and verification mechanisms as the best options.[3] These questions, together with the delicate topics on the potential misuse of AI tools applied for the removal of TOC will be analysed in paragraph 4.

4. Artificial Intelligence for countering cyber-terrorism

Given the extensive subversive use of the Internet and the various underlying available online channels by terrorist and extremist groups for the dissemination and propagation of terrorism-related

[2] Joint letter of 61 human rights organisations addressed to the European Parliament, (2021) accessed 03.01.2022.

[3] European Data Protection Supervisor, Formal comments of the EDPS on the Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online (2019).

multilingual and multimodal content using mainly the capabilities offered by HSPs, the adoption of AI-based disruptive technologies by hosting providers and law enforcement has become ever more significant to counter such threats. On the one hand, modern LEAs can leverage the available innovative solutions to collect open-source intelligence and process large amounts of data rapidly, hence, to analyse online content relevant to an ongoing investigation in a timely manner, whilst also allocating the available resources to cover additional operational needs. In this context, the LEAs are equipped with additional means to detect and identify subversive content online, including material that poses an imminent threat to life, and request its removal from the respective hosting providers under the TCO Regulation.

On the other hand, the HSPs can benefit by enhancing their arsenal with tools permitting the continuous automatic monitoring and analysis of the big amounts of data uploaded by the numerous users of their platforms, resulting in the detection and removal of such content in a more efficient manner, requiring fewer human resources. This is of particular interest for the ALLIES project, which focuses on HSPs of small and medium size characterised by limited capacity to respond to the abuse of their online spaces. In this context, disruptive technologies can also contribute to the capacity of HSPs to respond to removal orders issued by law enforcement within the timeframe set by the one-hour rule; AI-based tools can be used to automatically analyse removal orders acting as a decision-support mechanism facilitated by the HSP content moderators.

Several domains of AI including (but not limited to) deep neural networks and machine learning based on supervised or unsupervised learning approaches, can facilitate the fight towards countering the multimodal terrorist content online. Natural Language Processing (NLP) solutions provide useful insights by analysing the multilingual online textual content and extracting, disambiguating, linking, and semantically enhancing concepts and named entities [19] associated with terrorism- and extremism-related content. Automatic speech recognition tools process audio and visual files to produce accurate transcriptions for languages of interest that can be subsequently analysed by NLP technologies [20]. Computer vision models employing deep learning algorithms analyse videos and images to detect and recognise objects [21], concepts, human behaviour and activities [22] relevant to the domain of terrorism and extremism.

The indicators produced by the separate analysis of the textual, audio, and visual modality is leveraged by multimodal classifiers that automatically categorise the online data into a set of predefined categories relevant to the terrorism and extremism domain [23]. Powerful explainable AI techniques [24] leveraging the outputs produced by the aforementioned multimodal AI models support human interpretation, providing HSP moderators and LEA analysts with reasoning and explanation on the results delivered by the AI solutions, thus helping towards making informed decisions related to the removal of abusive content from the online space. Finally, threat assessment tools [25] help HSP moderators and LEA analysts towards assessing the severity of threats posted online, thus enforcing the appropriate measures including the removal of such material.

5. The Legal and Ethical Framework

Based on the above-mentioned considerations, it should be also underlined, that regulating the digital realm can oftentimes lead to interference with fundamental rights and presents a difficult conundrum of striking the balance between fundamental rights respect and guaranteeing public order and security. This is applicable also to the case of the TCO Regulation practical implementation. The latter outlines a number of obligations to HSPs leaving them with a little margin for independent decision-making, most notably in the case of removal order receipt. According to Art. 3 para 3 of the TCO Regulation, HSPs need to act promptly upon a removal order receipt, taking down the respective content in an hour, or in case of removal order issued by a competent authority from another EU Member State – 72 hours (Art. 4, para 3, TCO Regulation). In this case HSPs are not provided with

a discretion whether or not to comply with the removal order, the Regulation provides quite limited circumstances in which removal could be refused – if the removal order itself is erroneous or does not provide sufficient information (Art. 3, para 8, TCO Regulation), where sufficient information is defined as "*location of that content, by indicating the exact URL*" [26].

At the same time, the new TCO Regulation obliges HSPs to establish a complaint mechanism where content providers could contest the removal of their content or the disabling the access to it (Art. 10 para 1, TCO Regulation). This being said, it should be noted that any specific action a HSP would take in connection to applying measures countering the misuse of the services it provides is intrinsically linked to receiving removal orders from the competent authorities, as per Art. 5 of the TCO Regulation. Thus, the situation where HSPs are put in requires them to deal with complaints and objection to their actions, while the latter are being prompted by the assessment made by the competent authorities. The Regulation also establishes in Art. 9 the availability of legal remedies for both HSPs and content providers challenging a removal order before the competent courts in the respective EU Member State. It could be argued that going to court for contesting a removal order on behalf of a HPS or a content provider might render the rights holders unwilling due to lengthy and costly procedures, still it remains to be seen how often and to what extent the available legal remedy would be used in practice.

Considering the fundamental rights and the practical implementation of the TCO Regulation, a few remarks need to be made with respect to the enforcement of the right to freedom of expression. The text of the Regulation explicitly states that it should be applied in a sense that changes the postulates of EU law with respect to freedom of expression and information [26]. This is further elaborated in the preamble of the Regulation, where it is noted that "*expression of radical, polemic or controversial views in the public debate on sensitive political questions*" can never be qualified as terrorist content. What is more, the preamble calls for due consideration of the freedom of expression, alongside the right to information, the freedom and pluralism of media and the freedom of arts and sciences, whenever an assessment is being made whether certain types of content should be classified as terrorist or not.

Such an assessment becomes more complex once AI-powered tools are involved in the flagging of terrorist content online, which is the case of the ALLIES project striving the equip micro and small HSPs with the necessary knowledge, skills, and tools to ensure the practical implementation of the TCO Regulation. One of the risks that the ALLIES project team would address is how to prevent the planned tools of bias and discrimination, which may hamper freedom of expression. The latter could occur in case information leading to discrimination is being processed by the AI algorithm and contributes to the determination of whether a certain piece of content is in fact terrorism-related [27]. Which type of information could lead to unfavourable treatment is prescribed by EU law (Directive 2000/43/EC and Directive 2000/78/EC), namely information of:

- racial or ethnic origin,
- religion or belief,
- disability,
- age,
- sexual orientation.

An example of such a case would be if a video is removed from a video hosting platform due to the fact that the AI tool has labelled it as terrorism-related solely because the author of the video is of particular ethnic origin. However, the workings of an AI tool are not that simple, and one may fall victim of discrimination and unfair treatment due to "*profiling identities based on a combination of behavioural and demographic characteristics*" [27].

The application of AI in the field of tackling TOC is not that gloomy despite the risks outlined above. Although at the moment there is no legal framework at EU level regulating the use and application of AI per se, there is robust ethical guidance available postulating best practice and golden

standards which, when followed, provide high guarantees to the rights and freedoms of the individuals. One of the main directions EU policy takes is that of trustworthy AI [28]. The latter is achieved through the mainstreaming of the following key requirements [28]:

- human agency and oversight,
- technical robustness and safety,
- privacy and data governance,
- transparency,
- diversity, non-discrimination and fairness,
- societal and environmental wellbeing, and
- accountability

Aiming to minimise the risk of bias and discrimination in the case of the AI-supported TCO Regulation implementation, the key requirements presented above should be transformed into practical parameters. For example, the requirement of 'diversity, non-discrimination and fairness' might be translated into a practice where robust, diverse, and representative data sets are used for the training and testing of AI algorithms [28]. This could be further enhanced by putting into practice the requirements 'human agency and oversight' where personnel of diverse backgrounds are involved in the training, testing and application of the AI tools [29]. The latter requirement is of high importance in dealing with terrorism content online, as identifying potential criminal behaviour is a decision-making process that cannot be entirely allocated to machines considering the risk of potential harm. To this end, it is recommended that the human-in-command approach is utilised [29], meaning that a person of appropriate knowledge and experience is put in a position to supervise, closely monitor and decide how the results yielded by the AI algorithm should be interpreted, and what kind of actions should follow-up. Additional oversight mechanisms in place could further the trustworthiness of the AI system and serve as an additional guarantee against bias and discrimination occurring [29].

6. Conclusion

Based on the initial findings of the EU-funded project ALLIES, this paper described the strengths and weaknesses of disruptive technologies, with particular focus on AI used by LEAs and HSPs for countering cyber-terrorism.

Technologies can, on the one side, contribute to the capacity of HSPs to respond to removal orders issued by law enforcement within the timeframe set by the one-hour rule (following the TCO Regulation). Thus, HSPs can benefit by enhancing their arsenal with tools permitting the continuous automatic monitoring and analysis of the big amounts of data uploaded by the numerous users of their platforms, resulting in the detection and removal of such content in a more efficient manner, requiring fewer human resources. On the other side, they allow to equip LEAs with additional means to detect and identify subversive content online and request its removal from the respective hosting providers under the TCO Regulation. Overall, AI models support by providing HSP moderators and LEA analysts with reasoning and explanation on the results delivered by the AI solutions, thus helping towards making informed decisions related to the removal of abusive content from the online space. Finally, threat assessment tools help HSP moderators and LEA analysts towards assessing the severity of threats posted online, thus enforcing the appropriate measures including the removal of such material.

However, the support provided from innovative technologies, should be properly framed in a well-grounded legal framework, which allow to avoid bias and discrimination. The authors provide a way forward in order to strengthen this issue, by putting into practice the requirements 'human agency and oversight' where personnel of diverse backgrounds are involved in the training, testing and application of the AI tools.

References

[1] K. Luyten, "Addressing the dissemination of terrorist content online," European Parliamentary Research Agency, 2021.

[2] M. Mengu and S. Mengu, "Violence and Social Media," Athens Journal of Mass Media and Communications, vol. 1, no. 3, pp. 211-228, 2015.

[3] A. Tsesis, "Social Media Accountability for Terrorist Propaganda," Fordham Law Review, vol. 86, no. 2, pp. 605-632, 2017.

[4] T. Keatinge and F. Keen, "Social Media and (Counter) Terrorist Finance: A Fund-Raising and Disruption Tool," Studies in Conflict & Terrorism, vol. 42, no. 1-2, pp. 178-205, 2019.

[5] F. D. University., "Cybersecurity and Cyber Terrorism," [Online]. Available: https://online.fdu.edu/program-resources/cybersecurity-and-cyber-terrorism/. [Accessed 2023 April 2023].

[6] EUROPEAN PARLIAMENT AND COUNCIL, "eur-lex.europa.eu/," 17 March 2017. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017L0541&rid=6. [Accessed 15 March 2023].

[7] A. Meleagrou-Hitchens and N. Kaderbhai, "Research Perspectives on Online Radicalization. A Literature Review, 2006-2016," International Centre for the Study of Radicalization (ICSR), London, 2017.

[8] E. Charlie and L. Gribbon, "Pathways to Violent Extremism in the Digital Era," The RUSI Journal, vol. 158, no. 5, pp. 40-47, 2013.

[9] G. N. Mølmen and J. A. Ravndal, "Mechanisms of online radicalisation: how the internet affects the radicalisation of extreme-right lone actor terrorists," Behavioral Sciences of Terrorism and Political Aggression, pp. 1-25, 2021.

[10] L. S. Neo, "An Internet-Mediated Pathway for Online Radicalisation: RECRO," in Combating Violent Extremism and Radicalization in the Digital Era, Hershey, PA, USA, IGI Global, 2016, pp. 197-224.

[11] P. R. Neumann, "The trouble with radicalization," International Affairs, vol. 89, no. 4, pp. 873-893, 2013.

[12] L. B. Galen Lamphere-Englund, "State of Play on Gaming & Extremism – Reviewing the literature on gaming & extremism," 6 October 2021. [Online]. Available: https://drive.google.com/file/d/1WEq4OjtqZYdltAB0SK46M88gFF863jWs/view. [Accessed 28 April 2023].

[13] A. Al-Rawi, "Video games, terrorism, and ISIS's Jihad 3.0," Terrorism and Political Violence, vol. 30, no. 4, pp. 740-760, 2018.

[14] C. Ingersoll, "Free to Play? Hate, Harassment and Positive Social Experiences in Online Games 2020," Anti-Defamation League, 2020.

[15] Twitter, "Transparency - Rules Enforcement," Twitter, July-December 2021. [Online]. Available: https://transparency.twitter.com/en/reports/rules-enforcement.html#2021-jul-dec. [Accessed 27 04 2023].

[16] Meta, "Meta - Transparency Center," Meta, 2023. [Online]. Available: https://transparency.fb.com/data/community-standards-enforcement/dangerous-organizations/facebook/. [Accessed 27 04 2023].

[17] European Commission, "https://eur-lex.europa.eu/," 12 September 2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=SWD:2018:408:FIN. [Accessed 14 March 2023].

[18] G. Robinson, "The European Commission's Proposal for a Regulation on Preventing the Dissemination of Terrorist Content Online," 31 January 2019. [Online]. Available:

https://eucrim.eu/articles/commission-proposal-regulation-preventing-dissemination-terrorist-content-online/. [Accessed 2023 April 28].

[19] "Autoregressive Structured Prediction with Language Models," AutorarXiv preprint arXiv:2210.14698, 2022.

[20] K. Veselý, A. Ghoshal, L. Burget, and D. Povey, "Sequence-discriminative training of deep neural networks," Interspeech, vol. 2013, pp. 2345-2349, 2013.

[21] D. Touska, K. Gkountakos, K. Ioannidis, T. Tsikrika, S. Vrochidis, I. Kompatsiaris, "Graph-Based Data Association in Multiple Object Tracking: A Survey," in In Proceed of the 29th International Conference on Multimedia Modeling, 2023.

[22] K. Gkountakos, D. Touska, K. Ioannidis, T. Tsikrika, S. Vrochidis, and I. Kompatsiaris, "Spatio-temporal activity detection and recognition in untrimmed surveillance videos.," in In Proceedings of the 2021 International Conference on Multimedia Retrieval, 2021.

[23] "Domain-aligned Data Augmentation for Low-resource and Imbalanced Text Classification," in Proceedings of the 45th European Conference on Information Retrieval (ECIR'23), Dublin, 2023.

[24] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.Z. Yang, Explainable artificial intelligence. Science robotics, vol. 4, no. 37, 2019.

[25] O. Theodosiadou, D. Chatzakou, T. Tsikrika, S. Vrochidis and I. Kompatsiaris, "Real-time Threat Assessment based on Hidden Markov Models.," 2023.

[26] European Parliament and the Council, Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online (Text with EEA relevance), Official Journal of the European Union.

[27] European Union Agency for Fundamental Rights, Bias in Algorithms – Artificial Intelligence and Discrimination, Vienna: Luxembourg: Publications Office of the European Union, 2022.

[28] European Commission, "WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust," 2020.

[29] High-Level Expert Group on Artificial Intelligence, "ETHICS GUIDELINES FOR TRUSTWORTHY AI," European Commission, Brussels, 2019.

[30] K. Luyten, "Addressing the dissemination of terrorist content online," European Parliamentary Research Service, 2021.

[31] N. Stylianou, D. Chatzakou, T. Tsikrika, S. Vrochidis, I. Kompatsiaris, "Domain-aligned Data Augmentation for Low-resource and Imbalanced Text Classification," in ECIR'23, Dublin, 2023.

Innovation in the Financial Sector (FinTech): Paradigms, Causes, Effects and Perspectives

Ruxandra RÎMNICEANU, PhD
National Bank of Romania, Bucharest, Romania
ruxandra.rimniceanu@bnro.ro

Abstract

The changes and evolution of the international and domestic financial-banking system, in the context of globalization, after the financial crisis of 2008, determined the emergence of global, virtual banks, megabanks, financial groups that use disruptive technologies and technological innovations. The first FinTech Action Plan (technology-based innovation in the field of financial services or financial technological innovations) of the European Union mark, as well, the first step circumscribed to the EU Digital Finance Strategy, in order to allow the expansion of innovative business models, but without forgetting to strengthen cyber security and to increase the degree of integrity of the financial system. In this context, however, it is important to take into account the variety of the institutions and the technologies in the countries participating in the Single Supervisory Mechanism (SSM), because the FinTech banks capture the different activities of the credit institutions in different jurisdictions to be closer to the customers and the investors and, in the same time, to expand the area of supervision of the problems related to the emergence of FinTech, because they exceed a sector of the economy or a geographical area and involves multiple financial-banking supervisory and regulatory institutions, belonging to various sectors.

Index terms: cyber security, Digital Operational Resilience, disruptive technologies and technological innovations, innovative business models, systemic risks

1. Introduction

The changes and evolution of the international and domestic financial-banking system, in the context of globalization [1], after the financial crisis of 2008, determined the emergence of global, virtual banks, megabanks, financial groups that use ***disruptive technologies and technological innovations***.

According to the experts' opinion [2], the trends of that period were "***decentralization***"[1] and "***integration***"[2] (The **World Economic Forum**, 2018). To these is added, more recently, the "***digitalization***" trend, accelerated during the pandemic, so that "***digital finance***" (*FinTech*) offered new solutions, based on technology and adapted to the needs of the market and the requirements of the final consumer. Thus, **the dynamics of technological revolutions and the international and national regulatory framework**, quite vast and only for the traditional system**,** ***laid new foundations for the functioning of the financial-banking system*** and allowed ***the development of companies with a new type of approach*** to the financial sector, in parallel, which, first of all, ***changed the perspective*** on how financial-banking institutions interact with the consumer and ***facilitated the automation of processes, increasing the speed of transaction processing*** etc.

[1] Decentralized Finance (DeFi)
[2] Financial, banking and capital markets integration

2. Innovation in the financial sector – paradigms and general context

In March 2018, the European Commission presented the first FinTech Action Plan [3] (*technology-based innovation in the field of financial services or financial technological innovations* [4]) of the EU, accepted by the Chamber of Deputies of the Romanian Parliament through Decision no. 33/30.05.2018[3] for the adoption of the *Opinion on the Communication from the Commission to the European Parliament, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions - Action Plan on FinTech: for a more competitive and innovative European financial sector - COM(2018) 109* [5].

The objective of this *Plan* was to enable the EU financial sector to benefit from the new technologies that have transformed the industry, according to the measures listed below [6], aimed at:

- ***allowing the expansion of innovative business models;***
- ***strengthening cyber security and increasing the degree of integrity of the financial system.***

Supporting the assimilation of technological innovation in the financial sector	
Technology neutrality adequacy review	The Commission will set up an expert group to assess whether there are unjustified regulatory obstacles to financial innovation in the financial services regulatory framework.
Removing barriers to using the cloud services	The COM invites the European Supervisory Authorities to explore the need for guidelines on outsourcing to cloud service providers.
	In the context of the Communication on Building the European Data Economy, the Commission invites cloud stakeholders to develop cross-sectoral self-regulatory codes of conduct to facilitate switching between cloud service providers. The Commission will also invite representatives from the financial sector to enable easier data porting for financial institutions.
	In this context, the Commission shall encourage and facilitate the development of standard contractual clauses for cloud outsourcing by financial institutions, building on the cross-sectorial cloud stakeholder efforts already facilitated by the Commission, and ensuring the involvement of the financial sector in this process. This work should be undertaken by a balanced mix of companies from the financial sector and cloud service providers, and should address in particular audit requirements, reporting requirements or the determination of materiality of the activities to be outsourced.
EU public initiative on blockchain	The Commission will consult publicly on further digitisation of regulated information about companies listed on EU regulated markets, including the possible implementation of a European Financial Transparency Gateway based on distributed ledger technology.
	The Commission will continue to work on a comprehensive strategy, considering all relevant legal implications, on distributed ledger technology and blockchain addressing all sectors of the economy, including enabling FinTech and RegTech applications in the EU.
	The Commission launched an EU Blockchain Observatory and Forum in February 2018, as well as a study on the feasibility of an EU public blockchain infrastructure to develop cross-border services. It will be assessed whether blockchain can be deployed as a digital services infrastructure under the Connecting Europe Facility. With the support of the EU Observatory and Forum and the European Standardisation Organisations, the Commission will continue to appraise legal, governance and scalability issues and support interoperability and standardisation efforts, including further evaluating cases of blockchain use and its applications in the context of the Next Generation Internet.
Building capacity and knowledge in an EU FinTech lab	The Commission will host an EU FinTech Lab where European and national authorities will be invited to engage with technology solution providers in a neutral, non-commercial space during targeted sessions on specific innovations.

[3] Transposed into Romanian legislation by *Decision of the Parliament of Romania no. 33 of May 30, 2018 regarding the adoption of the opinion regarding the Communication from the Commission to the European Parliament, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions - Action Plan on FinTech: for a sector more competitive and innovative European finance - COM(2018) 109*

Improving the security and resilience of the financial sector	
Strengthening the cyber resilience of the EU financial sector	The Commission will organise a public-private workshop in Q2 2018 to explore and assess barriers limiting information sharing on cyber threats between financial market participants and to identify potential solutions while ensuring data protection standards are met.
	The Commission invites the European Supervisory Authorities to map, by Q1 2019, the existing supervisory practices across financial sectors around ICT security and governance requirements, and where appropriate: a) to consider issuing guidelines aimed at supervisory convergence and enforcement of ICT risk management and mitigation requirements in the EU financial sector and, b) if necessary, provide the Commission with technical advice on the need for legislative improvements.
	The Commission invites the European Supervisory Authorities to evaluate, by Q4 2018, the costs and benefits of developing a coherent cyber resilience testing framework for significant market participants and infrastructures within the whole EU financial sector.

At the same time, ***the European supervisory authorities*** were to carry out, by the first quarter of 2019, a review of the current **authorization and licensing approaches for innovative FinTech business models**, analyzing in particular **the national legislation in the field of financial services**, from the perspective of proportionality and flexibility, on the basis of which guidelines must be issued **regarding the approaches and procedures implemented** by the European supervisory authorities or **recommendations presented to the Commission regarding the need to adapt the EU legislation** in that matter. Further, on **24 September 2020**, the Commission adopted **a new digital finance package** [7][4], which includes:

- ***EU Digital Finance Strategy*** [8], with a focus on **fostering responsible innovation in the European Union's financial sector**, especially **for highly innovative digital start-ups (FinTech)**, with a view ***to developing safer and more user-friendly rules***, from an IT point of view, by consumers, but also ***the reduction***, at the same time, ***of possible risks related to investor protection, money laundering and IT crime***;
- *The EU Strategy on retail payments* [9], adopted by the European Council on 22.03.2021;
- legislative proposals **on crypto-assets** [10] and **digital operational resilience**.

The draft of the **Regulation on markets for crypto-assets (MiCA)** for which, on **October 5, 2022**, the Committee of Permanent Representatives (Coreper) approved the provisional agreement, thus, starting the formal adoption process:

- will apply to ***the persons involved in the issuance of crypto-assets***, as well as in ***the services related to the provision of crypto-assets in the European Union (EU)***, which are not regulated by other European legislation, but ***will not apply in the situations where it is already applicable, as well as, to the current EU financial legislation on financial instruments and structured deposits (MiFID), electronic money (EMD), bank deposits (D.49/2014), securitized products (R.2402/2017)***;
- establishes rules on transparency and publication requirements for the issuance and admission to trading of crypto-assets, the authorization and supervision of crypto-assets service providers and of their own issuers, for the operation, the organization and the governance of asset-backed token issuers – asset-referenced tokens (ART) and for the issuers of tokens assimilated to electronic currencies [electronic money tokens (EMT)] and for the crypto-asset service providers, the consumer protection rules, as well as the measures to prevent market abuse and to ensure the integrity of markets in crypto-assets.

[4] This package covers:
- the stimulation of European **competitiveness and innovation** in the financial sector;
- offering consumers and businesses **more choices in financial services** and modern payment solutions;
- ways to ensure **consumer protection** and **financial stability.**

In the same time, the complete package of ***legislative proposals on Digital Operational Resilience*** [11], adopted at the level of the European Union, is aimed at **preventing cyber attacks and improving the supervision of outsourced services**, which, in recent years, have been continuously growing[5], disadvantaging their own staff to being specialized and to accumulate superior know-how or to dismantling their own specialized departments, on the one hand, respectively, to favoring or to maintaining risks and threats to the information and communication systems (SIC), by facilitating an uncontrolled access to their own equipment and data (such, the own databases), to third parties, on the other hand.

Thus, **the technology companies** are increasing their importance in finance, both as **IT services providers** for the financial firms and as **financial services providers** themselves. From this perspective, ***the digital operational resilience legislation*** [12] ***aims to ensure that all the participants in the financial system have the necessary safeguards to mitigate the cyber-attacks and other risks***. Through the established measures, all businesses are required to:

- ensure that they ***can cope with all types of disruptions and threats related to information and communication technology (ICT)***;
- introduce ***a supervisory framework for the ICT providers***, in the cases of *cloud computing services providers*, for example.

As well, summarizing, the draft of the **Regulation on Digital Operational Resilience for the Financial Sector (DORA)** aims to be applied by a very wide range of entities, such as ***credit institutions, payment institutions, electronic currency institutions, investment firms, crypto-asset service providers, central securities depositories, central counterparties, trading venues, trade repositories, alternative fund managers, management companies, data reporting service providers, insurance and reinsurance companies, insurance and reinsurance intermediaries, occupational pension institutions, credit rating agencies, audit and statutory audit firms, administrators of critical benchmarks, crowdfunding service providers, central securitization registries and third party information and communication technology (IT&C) service providers***, and:

- establishes **the applicable requirements to the financial entities** in terms of IT&C risk management, the contractual clauses between the financial entities and the external IT&C service providers, the supervisory framework for external critical services providers and the rules carried on to the cooperation between the competent authorities;
- it is incumbent upon financial entities to have in place **an extensive internal governance and control framework for IT&C risk management**, as well as **a robust, comprehensive and well-documented IT&C risk management framework**. In this regard, the financial entities will need to implement **a robust and comprehensive digital operational resilience testing program** which includes ***the IT&C testing tools, systems and methodologies***;
- imposes upon the financial entities to establish and implement **a specific IT&C incident management process** to identify, track, classify and categorize IT&C incidents.

The draft of ***the Regulation*** includes **a separate set of provisions applicable to the critical third-party providers of IT&C services**, which will be designated by a committee of the European supervisory authorities, based on the criteria set out in the Regulation.

[5] According to the latest Gartner analyst forecast, global IT spending is expected to reach $4.6 trillion in 2023, up 5.1% year-over-year. In this context, the demand for IT in 2023 is expected to be consistent, as companies will continue to roll out their business digitization initiatives in response to economic developments, [Online]. Available: https://www.economica.net/consultanti-gartner-cheltuielile-it-vor-creste-cu-51-pana-la-46-trilioane-de-dolari-in-2023-la-nivel-mondial_621319.html

3. "Innovative technology" and "technological innovations" - causes and effects

In the last three years, the investments in the new technologies have increased substantially and ***the pace of innovation is exponential***.

Currently, a significant percentage of the population interacts with the banks using the mobile technology. At the same time, using a variety of the new tools that didn't exist a few years ago, ***the consumers are making payments, they are transferring money and they are making investments remotely***. **The Artificial Intelligence, the social networks, the machine learning, the mobile applications, the distributed ledger technology, the cloud computing and the Big Data analysis** have contributed to the development of ***new services and business models*** by the established financial institutions and by the new market entrants (Fintech).

As a result, the European Commission's proposal to create **a regulatory framework for crypto-assets**[6] aims to enable ***the innovation in a way that preserves the financial stability and that protects the investors***.

From this perspective, the Commission distinguishes between ***the crypto-assets already regulated by the EU law and the other crypto-assets*** and, additionally, on **30 May 2022**, the European Parliament and the European Council adopted *the Regulation (EU) 2022/858 of the European Parliament and of the Council regarding a pilot regime for market infrastructures based on distributed ledger technology, as well as amending the Regulations (EU) no. 600/2014 and (EU) no. 909/2014 and the Directive 2014/65/EU* [13] (hereinafter referred to as the **DLT Regulation**).

*The **DLT Regulation*** is limited to concerns for **improving the digital resilience** that can be achieved through ***the creation of the pilot units and the experimentation spaces***, as the main objective of the European Union leaders established at October 1-2, 2020 who debated **the digital transformation** during *the Extraordinary Meeting of the European Council*[7] and invited the European Commission to present, by March 2021, **a comprehensive Digital Compass** [14], the goal of which is mainly:

- to establish **the EU's concrete digital ambitions for 2030**, in which context it was agreed that **at least 20% of the funds from *the Recovery and Resilience Mechanism***[8] **should be made available for the digital transition**, including for SMEs;
- to strengthen the public-private partnerships and the joint digital capabilities to implement appropriate measures in **the fields of artificial intelligence, quantum computing and blockchain**, with a focus on developing the cutting-edge strategic capabilities ***to design and deploy the digital solutions at scale***, with interoperability, in the digital infrastructures.

These funds should contribute to making progress towards **achieving objectives** such as:

- **promoting the European development of the next generation of digital technologies**, including the supercomputers, the quantum computing, the blockchain technology and the artificial intelligence centered by the human factor;
- **developing the capabilities within strategic digital value chains**, especially the microprocessors;

[6] Crypto-assets are ***digital representations of values or rights that are transferred and stored electronically***. They may serve as a key to access a service, to facilitate payments or to be designed as financial instruments.

[7] Extraordinary European Council, October 1-2, 2020, [Online]. Available: https://www.consilium.europa.eu/ro/meetings/european-council/2020/10/01-02/

[8] The German Presidency and the European Parliament negotiators, on **18 December 2020** reached a provisional agreement on *the Recovery and the Resilience Mechanism*. With a financial envelope of EUR 672.5 billion, the facility is the centerpiece of the ***EU-next-generation*** recovery tool. This instrument will support the investments and the public reforms in the Member States, helping them face the economic and the social impact of the COVID-19 pandemic, as well as the challenges of the green and the digital transition.

- **accelerating the implementation of the secure and very high-capacity network infrastructures**, including those of optical fiber and 5G, throughout the E.U.
- **increasing the EU's ability to protect itself against the cyber threats**;
- **harnessing the full potential of the digital technologies** in such a manner as to achieve the ambitious objectives of the EU **in the environmental and climate policies**;
- **modernizing the digital capabilities of the education system**.

Thus, *the DTL Regulation* sets up the framework for **the establishment of the pilot units for market infrastructures** that wish to attempt ***to trade and to settle transactions with financial instruments in the form of crypto-assets***, in order:

- to allow the market participants and the regulatory authorities to gain experience in **the use of DLT exchanges** that would ***trade or register shares or bonds*** in *the dedicated digital ledger*;
- **to develop the secondary markets** for ***tokenized financial instruments***[9] by promoting the adoption of the distributed ledger technology (DLT) in the sphere of the trading and the post-trade services.

On **August 23, 2022**, the Financial Supervisory Authority (ASF) in Romania announced that the *Law no. 244/20.07.2022 on establishing measures to implement the Regulation (EU) 2020/1.503*[10] *of the European Parliament and of the Council of October 7, 2020 regarding European providers of crowdfunding services for businesses and amending the Regulation (EU) 2017/1.129 and Directive (EU) 2019/1.937* [15] was published in the Official Gazette Part I, no. 754 of 27.07.2022 [16] and provides express provisions regarding **the designation of the ASF as the competent authority responsible for the authorization and supervision of crowdfunding services providers**, as well as the fact that ***the provision of crowdfunding services***[11] ***does not fall into the category of the lending activities regulated by the legislation specific to non-banking financial institutions***, except for the situation in which they are granted by a non-banking financial institution.

Based on article 6 paragraphs (4)-(6) of the *Law no. 244/2022,* the Financial Supervisory Authority collaborates with the National Bank of Romania, according to the provisions of article 33 of the Regulation (EU) 2020/1.503, if:

- the requests for authorization come from ***the entities that already hold a valid authorization issued by the National Bank of Romania*** as a credit institution, an electronic money issuing institution or a payment institution or have been registered as specialized providers of information services regarding accounts, legal entities or non-banking financial institutions in accordance with the provisions of Government Emergency Ordinance no.99/2006, approved with amendments and additions by the Law no.227/2007, with its subsequent amendments and additions, of the Law no.

[9] It is expected that the so-called "tokenization" of financial instruments (i.e. the digital representation of financial instruments on the distributed ledgers or the issuance of traditional asset classes in tokenized forms to enable their issuance, storage and transfer using a distributed ledger), to open the new possibilities for improving efficiency in the trading and the post-trading process, [Online]. Available: https://www.ilegis.ro/eurolegis/ro/index/act/85145

[10] The main objective of the Regulation (EU) 2020/1.503 is to create a regulatory framework that allows crowdfunding platforms to access to the single market, through an authorization based on a single set of rules

[11] Crowdfunding:

- represents a type of activity in which a crowdfunding services provider, **without assuming any risk itself, manages a digital platform**, opened to the public, in order to connect or to facilitate the connection of the potential investors or of the potential business lenders seeking financing. Such financing could consist of ***loans, the purchase of securities or other instruments admitted*** for the purpose of crowdfunding;
- it provides **an alternative source of financing**, including ***the venture capital***, but it can also provide other benefits for businesses: it can validate a business idea, it can give entrepreneurs access to a large number of people providing insights and information, while also being ***a marketing tool***.

209/2019 regarding payment services and for the modification of some normative acts, and of the Law no. 210/2019 regarding the activity of issuing electronic money or of the Law no. 93/2009, with its subsequent amendments and additions;

- ***a credit institution authorized by the National Bank of Romania*** according to the provisions of the Government Emergency Ordinance no. 99/2006, approved with amendments and additions by the Law no. 227/2007, with its subsequent amendments and additions, intends to provide crowdfunding services and the request for authorization as a crowdfunding services provider is submitted to the Financial Supervisory Authority only after obtaining prior approval from the National Bank of Romania regarding the filling out of the activity object, according to the national regulatory framework applicable to credit institutions;
- ***the credit institutions authorized by the National Bank of Romania*** according to the provisions of the Government Emergency Ordinance no. 99/2006, approved with amendments and additions by the Law no. 227/2007, with its subsequent amendments and additions, can provide crowdfunding services only if they have obtained authorization from the ASF.

As for **the FinTech** banks (the banks that have "***a business model*** *in which the production and the delivery of banking products and the services is* ***based on innovative technology***", thus, their scope of activity being delineated from **the concept of the FinTech** defined in **June 2017**, by the Financial Stability Board (FSB)[12] which categorized it as "***the innovative technology used in financial services*** *that could lead to new business models, applications, processes or products with an associated material effect on the provision of financial services*" [17]) the act that is applicable to them is *the Guide for evaluating the license applications for FinTech credit institutions* [18], from **March 2018**, developed by the European Central Bank (ECB), such as the entities that fall within the definition of "*a credit institution*" provided for by *the Capital Requirements Regulation (CRR)*[13].

In this context, however, **it is important to take into account the variety of the institutions and the technologies in the countries participating in the Single Supervisory Mechanism (SSM)**, because **the FinTech banks** capture the different activities of the credit institutions in different jurisdictions and **include**:

- ***new FinTech branches*** of existing authorized banks[14];
- ***new market entrants*** that adopt technological innovation to compete with established traditional banks along the value chain, as well as the existing financial services providers (e.g. the payment institutions, the investment firms, the electronic money institutions, etc.) that are expanding their scope to include banking activities and therefore may be considered new market entrants that **require a banking license from the central bank**.

[12] **The Financial Stability Board (FSB)** is an international body that monitors and makes recommendations on the global financial system, established after the G20 Summit in London, in April 2009, as a successor to the Financial Stability Forum (FSF). It was established to coordinate the work of the national financial authorities and the global standards bodies at international level with the aim to develop an effective regulatory policy and a policy on supervisory legislation in the financial markets and to promote its implementation. The Council includes all major G20 Economies, FSF members and the European Commission. It is hosted and funded by the Bank for International Settlements (BIS) and the Board members are based in Basel, Switzerland.

[13] Article 4 paragraph (1) point (1) of the CRR defines a credit institution as "*an enterprise whose activity consists in attracting deposits or other repayable funds from the public and granting loans on its own account*", [Online]. Available: https://www.eba.europa.eu/regulation-and-policy/single-rulebook/interactive-single-rulebook/504

[14] For example, a licensed institution might decide to establish a new legal entity to apply fintech solutions that were previously developed in-house.

From this perspective, the FSB and the ECB considered it useful to classify the FinTech developments according to the main existing economic functions they provide, **highlighting their double impact on the financial-banking system, meaning that the common concerns aim at**:

- *identifying the implications of the FinTech on the financial stability*, **compared to the existing financial market**;
- *analyzing the activities and the results of the FinTech banks*, **as well as that of** *the FinTech services providers or the underlying technologies*.

However, **the area of the direct taxation** should not be neglected either. In certain Member States (Bulgaria, Denmark, Finland, France, Latvia, Lithuania, Poland, Slovakia, Slovenia) and in the United Kingdom, the activities related to ***the virtual currencies are subject to tax legislation***.

In Romania, in **December 2018**, *the Law no. 30 of January 10, 2019*[15] which subjects to income tax the gain from the transfer of virtual currency, the Ministry of Public Finances (MFP) specifying that, with regard to the specific case of operations regarding the virtual currency *Bitcoin*, from the point of view of the VAT regime, it must be taken into account to see the Decision of the Court of Justice of the European Union (CJEU) in case C-264/14 - Hedqvist, in which the CJEU ruled on the application of the VAT exemption provided for in article 135 paragraph (1) letter (e) from the Directive 2006/112/EC regarding the common system of VAT, transposed into national legislation at article 292 paragraph (2) letter (a) point 4 of the Law no. 227/2015 on the Fiscal Code, with its subsequent amendments and additions. Thus, the provision of services consisting of the exchange of traditional currencies with units of the virtual currency "Bitcoin" and vice versa, performed against the payment of an amount corresponding to the margin constituted by the difference between the purchase price of the coins and their sale price, are exempt operations of VAT, within the meaning of this provision.

4. The FinTech - perspectives

The rapid progress registered by **the FinTechs** (whose map, at the European level, was made in 2019 by WallStreet.ro [19]), **determines structural changes in the financial sector, generating, equally, systemic risks**, among which we list:

- overly prescriptive and hasty regulation risks leading *to undesirable results, from a business point of view*;
- by avoiding the updating policies and the regulatory frameworks, *the financial services providers in the EU may be put at a disadvantage*, in the context of an increasingly globalized market;
- in the case of cyber security, *the main risks remain unresolved*.

The first map of the Romanian FinTechs (Figure 1), displayed during *the Future Banking FinTech Edition event*[16] (held on October 1st, 2019, in Bucharest), contained the profiles of over 40 businesses in this industry, based in Romania or with Romanian founders.

Thus, forced by the technology, the regulations and the customer expectations, **the banks will have to introduce the sharing economy** (collaborative/shared economy) **and to be prepared for partnerships** to offer the customer a truly unique service.

[15] By the Law no. 30 of January 10, 2019 for the approval of the Government Emergency Ordinance no. 25/2018 regarding the modification and completion of some normative acts, as well as for the approval of some fiscal-budgetary measures, in article 116 paragraph (2) of **the new Fiscal Code**, after letter (b), a new letter was introduced, letter (c), with the following content: "(c) the gain from the transfer of virtual currency in the case of the income provided for in article 114 paragraph (2) letter (m), determined as the positive difference between the sale price and the purchase price, including the direct costs related to the transaction. Earnings below the level of 200 lei/transaction are not taxed, provided that the total earnings in a fiscal year do not exceed the level of 600 lei."

[16] [Online]. Available: https://www.wall-street.ro/articol/Finante-Banci/270851/fintech-map-by-future-banking-primeste-un-update-ce-jucatori-intra-pe-harta.html#gref

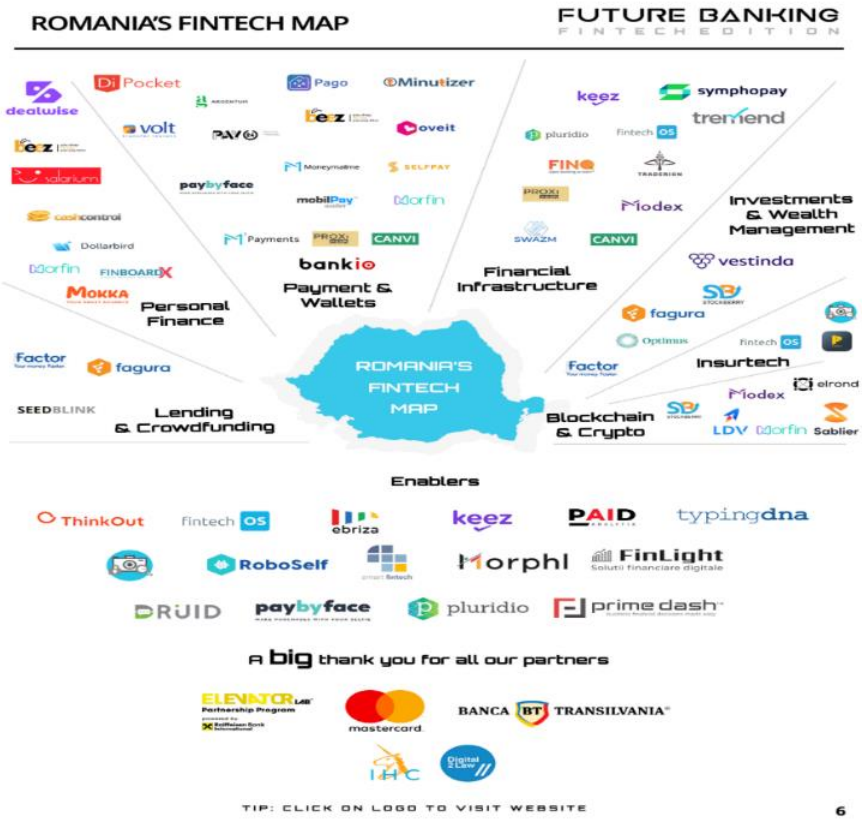


Fig. 1. Map of the Romanian FinTechs in 2019

This is also the conclusion expressed in the report of **October 10, 2022**, developed by BNY Mellon in collaboration with Aite-Novarica Group and entitled *The Forces Disrupting Payments* [20], according to which the financial institutions (FIs) serving the business customers are being disintermediated by FinTech payment providers, while the banks are bucking the trend by ***partnering with larger banks that have already built connections with the FinTechs***.

Moreover, some specialists in the field, since the beginning of 2019, based on some ***scenarios related to the evolution of the banking industry under the influence of the FinTech*** [21] have assessed that it is necessary **to expand the area of supervision of the problems related to the emergence of FinTech**, because **they exceed a sector of the economy or a geographical area and involves multiple financial-banking supervisory and regulatory institutions**, belonging to various sectors.

Table 1. Scenarios [20] related to the evolution of the banking industry under the influence of the FinTech

Scenario	Content
The better bank	***The modernization and the digitization of the existing players on the market.*** In this scenario, the existing banks are digitizing and modernizing to preserve their relationship with the customers and the main services offered to them, by using the new technologies and by changing their current business models.
The new bank	***Replacing the existing players with the new ones.*** In the future, the existing players cannot survive the wave of destruction produced by the technology and they will be replaced by the new "*technological*" banks, constituted, for example, by large technology companies, such as Amazon, Apple (hereinafter referred to as **BigTech**), which will, in time, fully digitize banking platforms. These new players will provide lower cost banking services. The new banks can obtain banking licenses from the supervisory authorities.
The distributed bank	***The division of the financial services between the specialized FinTech companies and the traditional banks.*** In this scenario, the financial services can be provided by anyone, old or new players, who can connect to the digital customer interface.

Scenario	Content
The downgraded bank	***The traditional banks will become commoditized services providers and will yield the direct relationship with the client to other financial services providers, such as The FinTechs and the BigTechs.*** Those will use platforms of direct communication with the customers in order to offer a lot of basic financial services, such as lending, saving, etc. The risk of the assets thus created can be transferred to the old bank or even to the owner of the platform, depending on the contract concluded between the two parties.
The disintermediated bank	***The banks will become irrelevant, the customers interacting directly between themselves with the help of individual financial services providers.*** The need for intermediation through a trusted third party disappears, the banks being replaced by more agile platforms and technologies, which will ensure the direct relationship between the customers, depending on their financial needs (a loan, making a payment, attracting capital etc.). The customers will choose their services and providers, but will assume more direct some responsibilities within the transactions, thus increasing the risks to which they are exposed.

Additionally, there is **a need for the banking supervision authorities to work with the public authorities**, which are responsible for ***data protection, the consumer protection, the competition and the national security***, **to ensure that the banks that use innovative technologies comply with the laws and the regulations in force**.

On the other hand, **the continuous training of personnel is necessary to face new challenges and to be able to exploit the potential of new technologies in order to improve their own methods and processes**, and **the regulatory framework requires periodic review and improvement**, to be adapted according to the emergence of the new risks generated by the technological progress.

From this perspective, in the European Union, 13 (thirteen) Member States have created the so-called "***the FinTech facilitators***"[17] (also called "*innovation centers*" or "*regulatory testing grounds*" [22]) **to provide the companies general guidelines during the process of authorization**. This allows such businesses to gain faster access to the market and better understand of the rules and more expectations regarding supervision.

"*The facilitators*" can also **provide guidance to established financial institutions**. From the perspective of the supervisors, ***such approaches are an important source of information***, helping them to gain a better understanding of the innovative business models and of the market developments at an early stage. "*The regulatory testing grounds*" takes away the idea of "*the innovation hubs*", further by creating an environment where the supervision is tailored to the innovative businesses or services.

The competent national authorities must apply the relevant EU legislation in the field of the financial services. However, the relevant rules include a margin of appreciation in applying the principles of proportionality and flexibility that are integrated in these rules. This issue can be particularly useful in the context of the technological innovation.

In order **to create a picture of these new types of financial institutions**, depending on their main field of activity, in correlation with the products and the services of traditional financial institutions, a classification of them, as it was made and published [23], in July 2018, by Sorin Mititelu, President-General Manager of Optimus, would integrate:

1. Platforms for money transfers
2. Platforms intended for online brokerage operations on the capital market
3. Robo-advisors
4. Platforms for financial planning services (personal finance management)

[17] "Innovation Hub" means an institutional mechanism where the regulated or the unregulated entities (i.e. unauthorized undertakings) collaborate with the competent authority to discuss the FinTech matters (exchange of information and views etc.) and to request clarification on the compliance of business models with the regulatory framework or on the regulatory requirements/licensing requirements (e.g. individual guidance provided to a business on the interpretation of the applicable rules)

5. Aggregation platforms for financial product offers for retail customers (current accounts, deposits, loans, insurances, etc.)
6. Alternative financing platforms for retail customers - peer-to-peer lending
7. Alternative financing platforms for companies - equity crowdfunding
8. Platforms intended for trading operations for institutional clients
9. Alternative insurance platforms - peer-to-peer insurance; mixed (traditional + P2P)
10. Big data management and analytics services, including through the use of IoT (Internet of Things)
11. Distributed ledger technologies (blockchains)
12. Platforms for market research and networking
13. Fraud and security risk management solutions
14. Solutions for smart cities
15. Trading platforms with mixed products and services (marketplaces).

5. Instead of conclusions

The European Systemic Risk Committee's warning of **September 22, 2022** regarding the vulnerabilities of the Union's financial system (CERS/2022/7), sent against *the backdrop of the increased systemic risks to financial stability*, urges "*(...) that the private sector institutions, the market participants and the relevant authorities to continue to prepare for* ***the materialization of the extreme risk scenarios***. *Maintaining or increasing the resilience of the Union's financial sector remains essential so that* ***the financial system can continue to support the real economy if and when the risks to the financial stability should materialize.***" [24]

In other words, the recommendations addressed to the European Commission regarding the digital finance sector are kept up to date, as **regulatory and supervisory challenges in the field of the financial services, institutions and markets**, so that the measures adopted at the European level [24][18] guarantee that the market participants, from the small to the large ones, have ***the necessary regulatory space in which to innovate***, and the legislation, new or updated, and the supervision in the field to be based on the following principles:

- the identical activities and services and related similar risks should be subject to the same rules;
- proportionality and technological neutrality;
- an approach based on risk, transparency and responsibility;
- respect for the fundamental rights, in particular to the protection of private life and personal data, guaranteed by the articles 7 and 8 of the Charter of Fundamental Rights of the European Union;
- high level of consumer and investor protection;
- fair competition;
- a friendly approach to innovation.

Circumstance of the above-mentioned approaches, we appreciate that ***the academic training in the financial field*** with the latest trends in the area of sustainable development, ***with an emphasis on digitization, innovation, new technologies and standards associated with the concept of sustainability*** on all its specific levels - environmental, social and governance (ESG), constitutes a **major objective** that is the basis for obtaining concrete and sustainable results.

[18] The European Parliament Resolution of **8 October 2020** containing recommendations to the Commission on the digital finance sector: emerging risks in cryptoassets - regulatory and supervisory challenges in the field of services, institutions and financial markets

References

[1]. Emanuela Mihaela Savu, "Trends regarding the globalization of the international banking system", Romanian Banking Institute Bucharest, [Online]. Available: http://store.ectap.ro/articole/210.pdf.

[2]. coord.: Eugen Dijmărescu, "Sistemul financiar internaţional în derivă?" – Bucureşti, Editura Centrului de Informare şi Documentare Economică, 2019, [Online]. Available: http://www.cide.ro/Sistemul%20financiar%20international.pdf.

[3]. "The first FinTech Action Plan", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/DOC/?uri=CELEX:52018DC0109&from=RO.

[4]. Financial Stability Board, "Report on the implications of FinTech on financial stability, June 27, 2017, [Online]. Available: https://www.fsb.org/wp-content/uploads/R270617.pdf.

[5]. Opinion on the Communication from the Commission to the European Parliament, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions - Action Plan on FinTech: for a more competitive and innovative European financial sector - COM(2018) 109, [Online]. Available: https://legislatie.just.ro/Public/DetaliiDocument/201301.

[6]. Extract from the Annex to the Action Plan on FinTech: for a more competitive and innovative European financial sector, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52018DC0109.

[7]. "The digital financial package", [Online]. Available: https://finance.ec.europa.eu/publications/digital-finance-package_en.

[8]. "The first progress report on the EU Strategy on a security union", [Online]. Available: https://ec.europa.eu/transparency/documents-register/api/files/COM(2020)797_0/de00000000009518?rendition=false.

[9]. "The Council conclusions on the Commission Communication on a European Union strategy on retail payments", 22.03.2021, [Online]. Available: https://data.consilium.europa.eu/doc/document/ST-7225-2021-INIT/ro/pdf.

[10]. "The Regulation on markets for crypto-assets (MiCA)", [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0593.

[11]. The entire package include:
1) EU Cyber Security Strategy for the Digital Decade, JOIN(2020) 18 final, 16.12.2020, [Online]. Available: http://www.cdep.ro/afaceri_europene/CE/2020/JOIN_2020_18_RO_ACTE_f.pdf.
2) Directive related the Critical Infrastructure Resilience, 16.12.2020, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:52020PC0829&from=EN.
3) Directive (EU) 2016/1148 on the Security of Networks and Information Systems, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32016L1148&from=IT.

[12]. "The Regulation on Digital Operational Resilience for the Financial Sector (DORA)", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52020PC0595.

[13]. "The Regulation (EU) 2022/858 of the European Parliament and of the Council regarding a pilot regime for market infrastructures based on the distributed ledger technology, as well as amending the Regulations (EU) no. 600/2014 and (EU) no. 909/2014 and the Directive 2014/65/EU", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:32022R0858&from=RO.

[14]. "Compass for the Digital Dimension 2030: The European Model for the Digital Decade", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/DOC/?uri=CELEX:52021DC0118&from=ro.

[15]. "The Regulation (EU) 2020/1.503 of the European Parliament and of the Council of October 7, 2020 regarding European providers of crowdfunding services for businesses and amending the Regulation (EU) 2017/1.129 and Directive (EU) 2019/1.937", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:32020R1503.

[16]. Law no. 244/20.07.2022 on establishing measures to implement the Regulation (EU) 2020/1.503 of the European Parliament and of the Council of October 7, 2020 regarding European providers of crowdfunding services for businesses and amending the Regulation (EU) 2017/1.129 and Directive (EU) 2019/1.937, [Online]. Available: https://legislatie.just.ro/Public/DetaliiDocument/257861.

[17]. Financial Stability Board, "Financial Stability Implications from FinTech", p.7, 27 June 2017, [Online]. Available: http://www.fsb.org/wp-content/uploads/R270617.pdf

[18]. „The Guide for evaluating the license applications for FinTech credit institutions", [Online]. Available: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.201803_guide_assessment_fintech_credit_inst_licensing.en.pdf?b12c4827939c8f29ad1c83ad7271b96d.

[19]. On 01.10.2019, Wall-Street.ro launched the first European map of the FinTechs, [Online]. Available: https://storage.icorp.ro/storage/html5/200909_6954_r1599681417/FBHE2020-FintechMap-v3.2.16_75790b54f9f673ccdb963697f86fab85.pdf.

[20]. "The Forces Disrupting Payments", [Online]. Available: https://www.bnymellon.com/content/dam/bnymellon/documents/pdf/insights/the-forces-disrupting-payments.pdf.

[21]. New banking industry configurations in the context of FinTech development, Ilinca GOROBEŢ, [Online]. Available: http://oaji.net/articles/2019/1425-1575536067.pdf, published in the Journal „ECONOMICA" nr.3 (109) 2019, p.103, following the analysis carried out by the Basel Committee on Banking Supervision.

[22]. The approach to financial technology (Fintech), EBA/DP/2017/02, [Online]. Available: https://www.eba.europa.eu/regulation-and-policy/other-topics/approach-to-financial-technology-fintech-.

[23]. The Financial Market Magazine, [Online]. Available: http://www.piatafinanciara.ro/wp-content/uploads/2018/08/BANCHERI-2018.pdf, p.90-94.

[24]. The warning of the European Systemic Risk Board of 22 September 2022 on vulnerabilities in the financial system of the Union (CERS/2022/7), [Online]. Available: https://www.esrb.europa.eu/pub/pdf/warnings/esrb.warning220929_on_vulnerabilities_union_financial_system~6ae5572939.en.pdf.

[25]. Digital finance: emerging risks in crypto-assets - regulatory and supervisory challenges in the financial services, institutions and markets, (2020/2034(INL)) (2021/C 395/10), [Online]. Available: https://www.ilegis.ro/eurolegis/ro/index /act/80440.

An Overview of RPL Networks from the Viewpoint of Cybersecurity

Cosmina STALIDI, **Eduard-Cristian POPOVICI, George SUCIU**
Telecommunications Department & Research & Development Department, Faculty of Electronics, Telecommunications and Information Technology & Beia Consult International, Bucharest, Romania
cosmina.stalidi@beia.ro, eduard.popovici@upb.ro, george@beia.ro

Abstract

In the past decade, the Internet of Things (IoT) has had a significant impact on a global scale. The Internet of Things (IoT) has facilitated the interconnection of a vast number of devices in contemporary times. The proliferation of Internet of Things (IoT) devices underscores the importance of ensuring robust security measures to safeguard against potential threats. The RPL protocol has been specifically designed for routing purposes within the context of IoT devices, operating at the network layer. The exploitation of the RPL protocol poses a threat to IoT networks and has the potential to substantially affect network performance. This article introduces the STACK project, which aims to improve IoT transmission capabilities, identify and mitigate attacks using performance and interference monitoring, and use methods tightly integrated with an intelligent edge.

Index terms: RPL, Contiki, COOJA, Security challenges, IoT

1. Introduction

The Internet of Things, also known as IoT, is an expansive field of technology and study, a portion of which is made up of Low-power and Lossy Networks, or LLNs for short [1]. Due to the fact that the nodes that make up such networks are vulnerable to a variety of restrictions and problems, the currently used routing protocols are inadequate [2].

The maturity of RPL has been demonstrated in its ability to establish connectivity between IPv6 devices, while exhibiting an acceptable amount of control overhead, even under demanding circumstances such as lossy links, heterogeneous and constrained devices and new security risks [1]. The RPL routing standard has been developed with an elevated level of adaptability, necessitating its customization to meet the particular demands of various applications. Current research endeavors pertaining to RPL are focused on enhancing its energy efficiency through various techniques. Numerous objective functions and metrics have been proposed in previous scholarly works, as a result of this.The parent nodes that are selected are experiencing an excessive burden as a consequence of multiple child nodes being linked with each parent node, which consequently leads to a compromising of these particular nodes [3].

The purpose of this article is to introduce the STACK project (Smart and Resilient IoT Networks against Attacks), which aims to enhance the transmission capabilities of IoT, identify and mitigate attacks through the use of performance and interference monitoring, and employ algorithms that are tightly integrated with an intelligent edge. This manuscript outlines the procedures involved in conducting a series of simulations using the COOJA and Contiki simulator. Additionally, it provides an analysis of the network topology parameters and potential cybersecurity obstacles that may arise.

2. Literature review

Contemporary Internet of Things (IoT) gadgets are typically equipped with low power and lossy networks, rendering conventional routing protocols such as RIP, DSR, and OSPF inapplicable. The RPL protocol is utilized for routing method in intelligent gadgets that are subject to constraints such as restricted memory and energy, as well as reduced processing power [4]. The development of destination-oriented DAG (DODAG) was aimed at achieving a loop-free system and merging towards one destination [5]. Each node in the DODAG graph is assigned a rank number that indicates its location. The rank number additionally serves to calculate the proximity between a node and adjacent ones as well as its location from the root node [6].

The RPL protocol (Fig.1) allows for the categorization of nodes through three distinct methods. The first category of nodes demonstrates host-like behavior; these are referred to as end devices or leaf nodes. The subsequent classification is denoted as router and bears the responsibility of executing the tasks of traffic generation and message transmission. The aforementioned group of nodes may be regarded as a border router, commonly referred to as a downstream node or DODAG root.

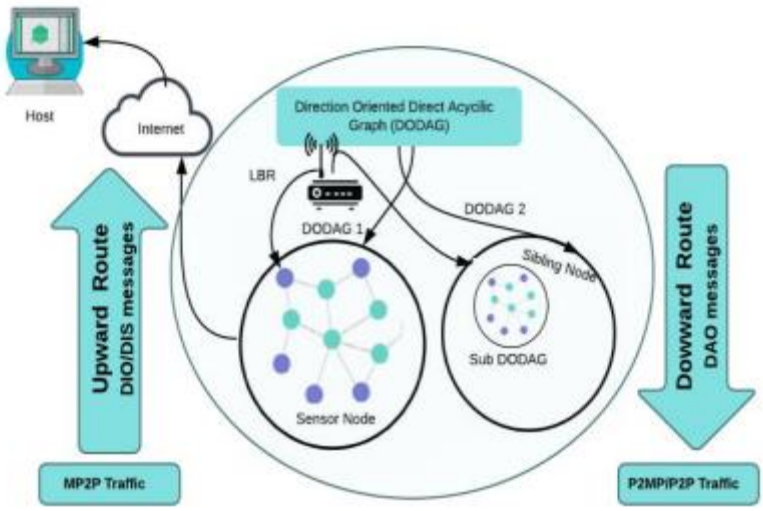


Fig. 1. Concept of RPL Protocol (source: [7])

Numerous attacks on networks have been observed on the RPL protocol, including but not limited to connection failures, processing power limitations, accessibility, network switching, and network structure. The classification of the network layer attacks primarily consists of external attacks and internal attacks [8].

a. Clone ID Attack

It is possible for an assailant to replicate the characteristics of additional nodes, which are commonly referred to as hooked nodes. Consequently, the attacker is able to generate several copies of packets by obtaining the encryption secret and ID of the node, leading to the misrouting of said packets. In general, nodes that act as attackers gather information such as rank ID and other relevant data pertaining to the nodes they target [9].

b. Selective Forwarding Attack

The present assault is initiated through the discriminatory transmission of packets, resulting in significant disturbances within the routing trajectory. This attack has the potential to facilitate the activation of a DDoS attack. The perpetrators discard all network traffic except for control communications, according to the statement [10].

c. Blackhole Attack

The blockhole attack (Fig.2) is characterized by the intentional dropping or blocking of data packets that carry legitimate messages by a malicious node. Instead, the attacker intentionally forwards messages containing misleading data, resulting in increased control overhead and packet delay. A node that seeks to reach its final location may be misled into following the shortest path towards said destination, thereby falling prey to deception. Upon receipt of data packets of data, a loss of service may occur, resulting in failure to deliver the packet to its intended destination. This may also lead to location exploitation. Consequently, there is a deficiency of interaction between the authentic source and destination nodes. The blockhole node is not visually discernible within the network, thus necessitating meticulous monitoring of network traffic. The occurrence of a blackhole attack results in a decline in network performance, manifesting as decreased throughput and routing complications [11].

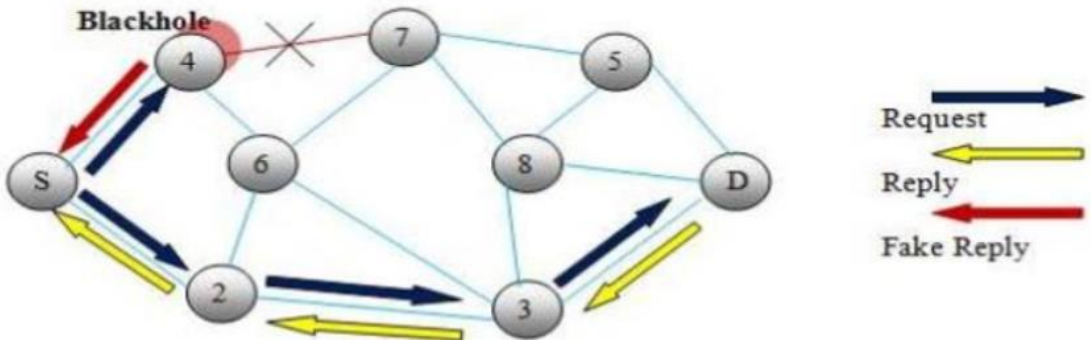


Fig. 2. Black Hole Attack (source: [12])

d. Wormhole Attack

The RPL protocol may be adversely affected by wormhole attacks, which can cause disturbances in both traffic patterns and the structure of networks. The present assault involves the rerouting of all information packets and traffic through a tunnel that has been established by two assailants. There exist two distinct mechanisms that determine how a wormhole may manifest. The encapsulation process involves the reception of a packet in a defined form by the connected or neighboring node, whereby the packet is detached from the payload. Packet relay is a method by which a malevolent node transmits packets to remote nodes that are considered as neighboring nodes [13].

3. Experimentation in Contiki

a. Performance metrics of the RPL network

The Packet Delivery Ratio (PDR) is a metric that quantifies the proportion of messages that are successfully transmitted from the origin node to be received at the root node of the network. It is calculated by dividing the quantity of acquired messages at the root node by the total amount of packets of data delivered through the origin node. The mean Packet Delivery Ratio (PDR) is calculated by aggregating all received and processed transmissions at the base node in the network.

The power consumption of a node is the aggregate amount of power utilized for transmission, acceptance, low power/sleep mode, and data analyzing by the microcontroller for processing.

Latency refers to the amount of duration required for a packet of data to obtain the root network node from its origin node.

b. Contiki and Cooja simulation environment

Contiki is an operating system designed for the Internet of Things (IoT) that is tailored to cater to the requirements of tiny IoT gadgets that have limited memory, power, bandwidth, and processing capabilities.

Contiki facilitates both conventional and contemporary enabling protocols for the Internet of Things (IoT). The uIP protocol is designed for use with IPv4. The present implementation of TCP/IP has the capability to provide support for microcontrollers of both 8-bit and 16-bit. The uIPv6 extension to uIP is a fully compliant implementation of IPv6. The Rime stack serves as a viable option in situations where the utilization of IPv4 or IPv6 is not feasible. The system provides a collection of fundamental components suitable for energy-efficient devices. The acronym 6LoWPAN denotes the implementation of Internet Protocol version 6 over wireless personal area networks with low power consumption. The technology offers compression capabilities to facilitate the utilization of low data rate wireless communication, which is essential for devices that possess restricted resources.

The Routing Protocol for Low-Power and Lossy Networks (RPL) is an IPv6 distance vector protocol that enables the identification of the optimal path in a network of devices with diverse capabilities, particularly those operating in low-power and lossy networks (LLNs). The Constrained Application Protocol (CoAP) facilitates communication for low-power and resource-constrained devices, typically those necessitating extensive remote monitoring [14].

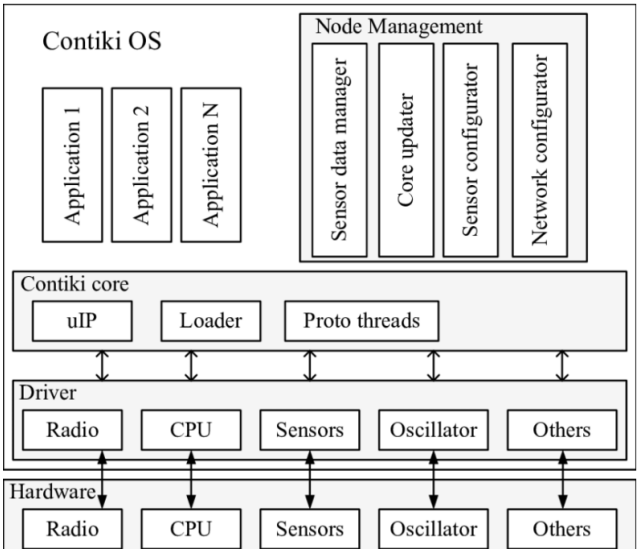


Fig. 3. Architecture of the OS Contiki (source: [15])

As previously discussed, Contiki is a minimalistic operating system that has been primarily designed for wireless nodes. Contiki's developed algorithms provide numerous benefits. The software platform known as Contiki offers a simulator named Cooja, which is based on the Java programming language and is utilized for the purpose of simulating wireless sensors. The Cooja simulator exhibits a higher degree of flexibility, as numerous components of the simulator are amenable to replacement and extension. The replaceability of certain components within the simulator, such as the simulated node hardware, extensions and broadcasting circumstances, represents a notable feature. Cooja is characterized by its scalability, efficiency, extensibility, and flexibility. The Contiki Cooja Wireless Sensor Network Simulator is primarily utilized for simulating numerous wireless scenarios [16].

c. The implementation and setup of RPL

Upon completion of the software technology installation procedure, the user inputs a command `cd contiki/tools/cooja/ant run` into the terminal, which triggers the emergence of an initial window, thereby enabling the commencement of the simulation process in Cooja.

```
robert@Ubuntu:~/contiki/tools/mspsim$ ant run
Buildfile: /home/robert/contiki/tools/mspsim/build.xml

init:
    [mkdir] Created dir: /home/robert/contiki/tools/mspsim/build

compile:
    [javac] Compiling 242 source files to /home/robert/contiki/tools/mspsim/build
    [javac] warning: [options] bootstrap class path not set in conjunction with -so
urce 7
    [javac] warning: [options] source value 7 is obsolete and will be removed in a
future release
    [javac] warning: [options] target value 7 is obsolete and will be removed in a
future release
    [javac] warning: [options] To suppress warnings about obsolete options, use -Xl
int:-options.
    [javac] /home/robert/contiki/tools/mspsim/se/sics/mspsim/Main.java:52: warning:
 [deprecation] newInstance() in Class has been deprecated
    [javac]       return nodeClass.newInstance();
    [javac]                       ^
```

Fig. 4. The command prompt following the installation of Contiki

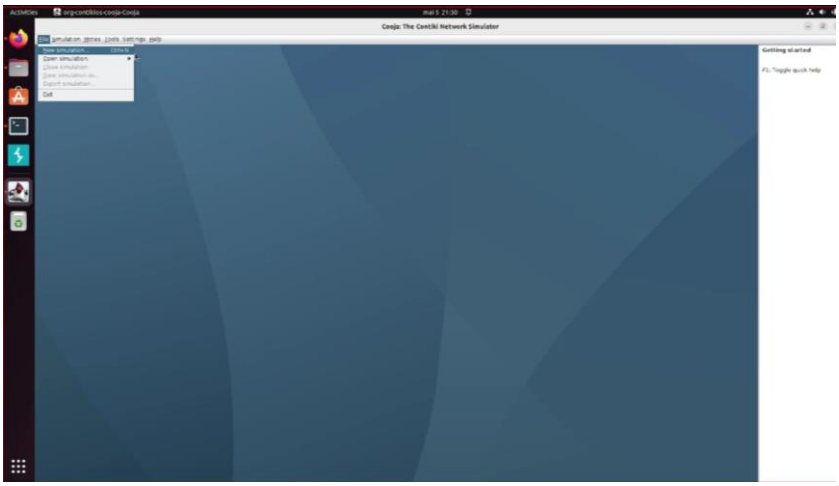


Fig. 5. A new Cooja simulation window

To initiate a new simulation, please proceed to the STACK simulation and select the option to create it.

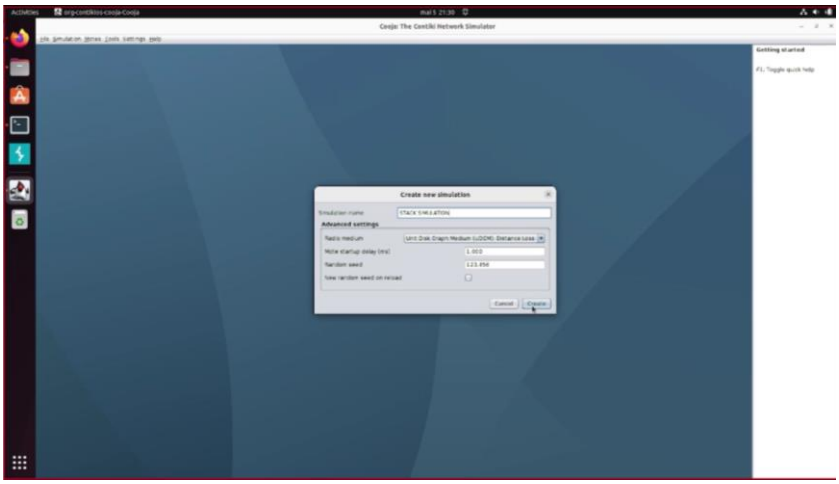


Fig. 6. The initial step in commencing a Cooja simulation

In the event that the Network Window lacks Grids, one may navigate to the view menu and select the 10m background grid from the options presented. Now to add a Sky mote, access the Motes menu, select the option to Add motes, proceed to click on Create new mote type, and ultimately opt for the Sky mote alternative. The Sky Mote option was selected for the purpose of emulating Tmote Sky motes.

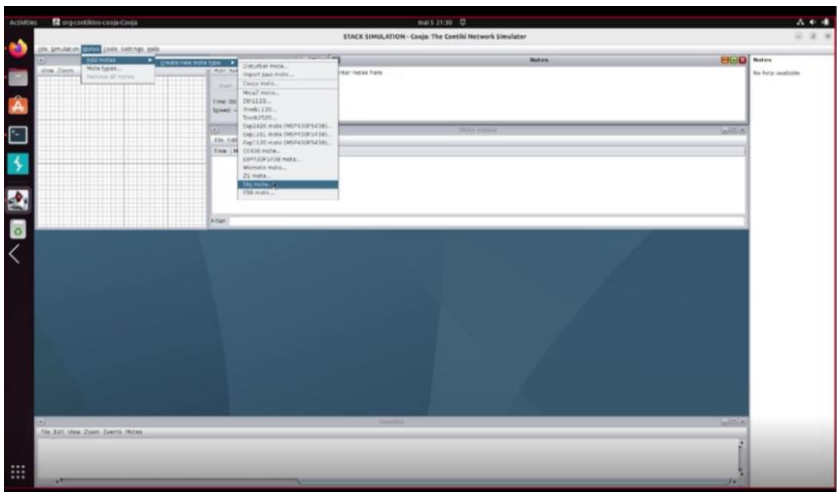


Fig. 7. Configuring parameters

To initiate the creation of a Mote type window, please select the option to browse. After this step, navigate to the directory *path /home/user/contiki/examples/ipv6/simple-udp-rpl*.

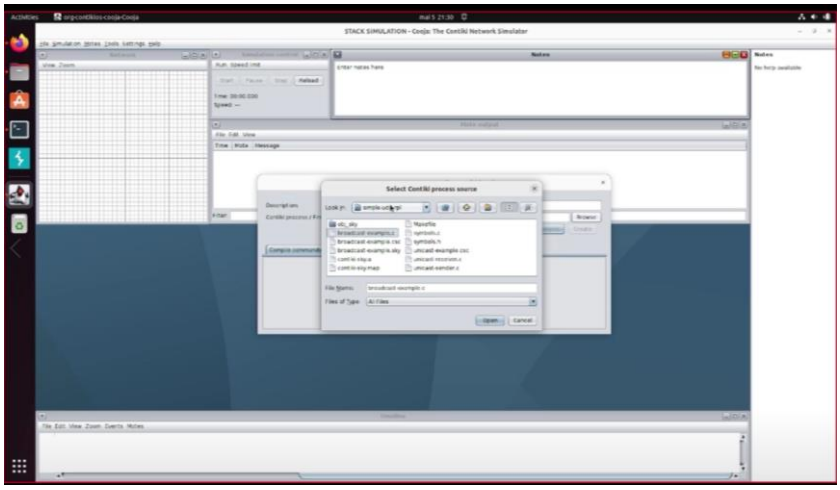


Fig. 8. The process of preparing for compilation

Select the file named *broadcast-example.c* and click on the Open button. To proceed, navigate back to the previously accessed Menu, select the option to compile, allow for a period of time to elapse, and ultimately initiate the process by selecting the Create button.

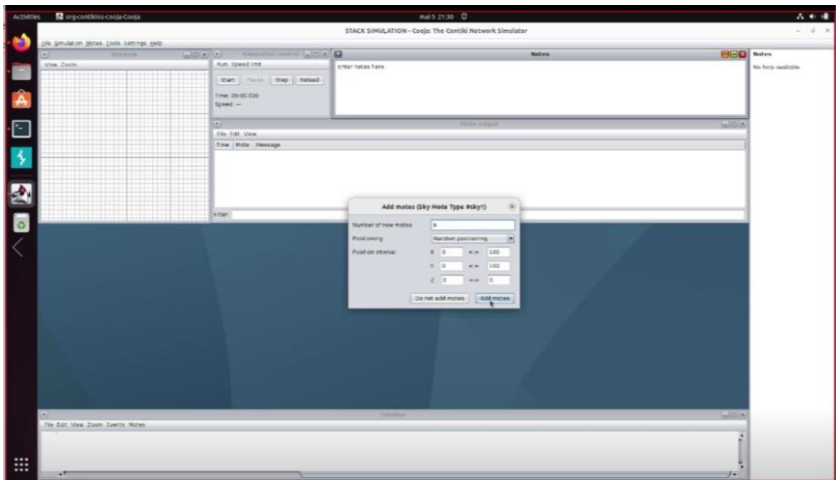


Fig. 9. Determining the quantity of nodes

A prompt will be displayed to add notes in a new window. The next step is to choose multiple options and click on the "Add Notes" button. To distinguish among these motes, access the View Menu located in the Network Window and select the Mote IDs option. Each mote will be assigned a numerical value.

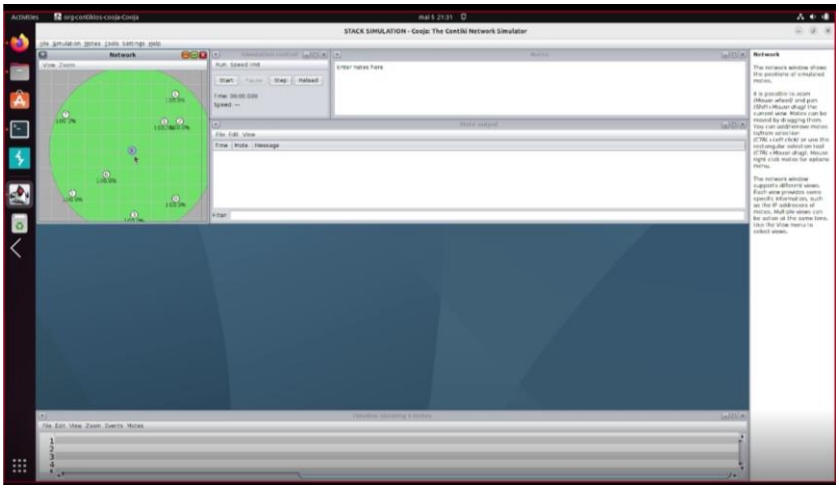


Fig. 10. The practice of monitoring network radio traffic

To monitor the radio traffic, navigate to the View Menu and select the option for Radio Traffic. Depress the Start button to initiate the simulation. By pressing the "Pause" button, the user can observe the interactions within the network through the Network Monitoring Window. The Mote results display will display a printout of the simulated motes. The conclusion of our simulation has been reached.

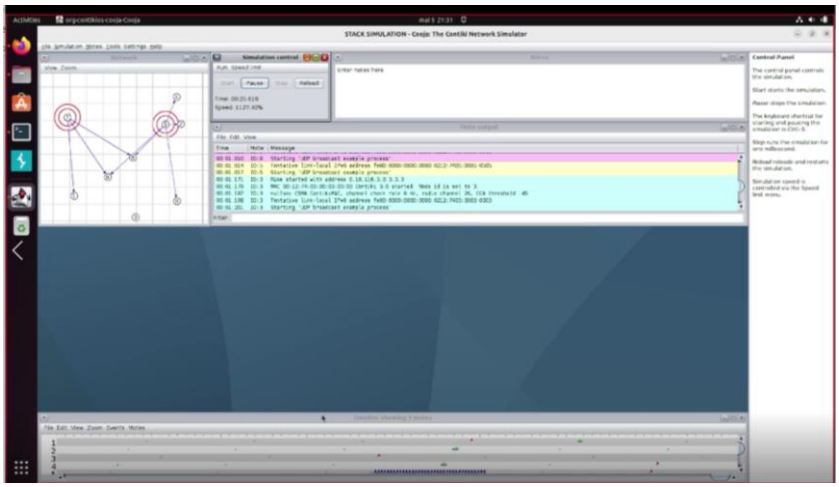


Fig. 11.

4. STACK Project

The STACK project, part of the ITEA initiative, aims to facilitate the provision of high Quality of Service (QoS) for Internet of Things (IoT) applications, even in situations that are not benign, by making them resistant to attacks. The objectives encompass enhancing the transmission capabilities of IoT, detecting and mitigating attacks through accomplishment and interference tracking, and employing algorithms that leverage a closely integrated smart edge [17].

4.1. Approach methodology for addressing the challenge

The absence of assured reliability, delay, and privacy in numerous IoT devices is a significant apprehension, particularly in light of the escalating incidence of security breaches. The vulnerability of IoT mesh networks comprising devices that are embedded is primarily attributed to their wireless communication and relatively low output power. Despite these limitations, their impact on critical domains such as autonomous vehicles and healthcare is increasingly significant, thereby posing a significant threat to our security and livelihood [18]. The issue in ensuring the performance of Internet of Things (IoT) networks during challenging circumstances, such as incidents and cross-technology interference, lies in the limitations imposed by resource constraints that avoid the implementation of advanced defenses on machines.

4.2. Possible project outcomes and impact

STACK provides a variety of innovative solutions to address this difficulty. Given the need of both recognition of attacks and prevention in non-benign circumstances, the initiative aims to leverage the computational capabilities of the smart edge.

The acquisition of requisite training data and the development of novel defenses will be facilitated through the utilization of deployments and testbeds. The data can be utilized to develop compressed models that can be deployed on IoT devices or gateways [18]. The proposed approach for mitigating new attacks involves leveraging frequency, data rate, and protocol variety. This strategy aims to ensure quality of service (QoS) levels and communication prioritization in the event of an

5. Conclusion

The current investigation introduces the routing protocol for low power consumption and lossy networks (RPL) specifically developed for wireless sensor networks. This study aims to provide a comprehensive understanding of the operational mechanisms of RPL, that lacks a predetermined standard for its security operations, necessitating the standardization of security operation protocols by researchers. It outlines the step-by-step configuration of RPL in the Cooja simulation environment, while also discussing potential cybersecurity-related obstacles that could impact the routing protocol for low-power wireless networks that are prone to packet loss. lacks a predetermined standard for its security operations, necessitating the standardization of security operation protocols by researchers.

References

[1]. George Simoglou, George Violettas, Sophia Petridou, Lefteris Mamatas, Intrusion detection systems for RPL security: A comparative analysis, Computers & Security, Volume 104, 2021,102219, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2021.102219.

[2]. Olfa Gaddour, Anis Koubâa, Mohamed Abid,Quality-of-service aware routing for static and mobile IPv6-based low-power and lossy sensor networks using RPL, Ad Hoc Networks, Volume 33, 2015, Pages 233-256, ISSN 1570-8705, https://doi.org/10.1016/j.adhoc.2015.05.009.

[3]. Qasem M, Al-Dubai A, Romdhani I, Ghaleb B, Gharibi W, "A new efficient objective function for routing in internet of things paradigm", in Standards for Communications and Networking (CSCN), 2016 IEEE Conference on 2016 Oct 31 (pp. 1-6). IEEE.

[4]. S. Y. Hashemi and F. Shams Aliee, "Dynamic and comprehensive trust model for IoT and its integration into RPL," J. Supercomput., vol. 75, no. 7, pp. 3555–3584, 2019.

[5]. A. E. Hassani, A. Sahel, A. Badri, and E. M. Ilham, “A hybrid objective function with empirical stability aware to improve RPL for IoT applications,” Int. J. Electr. Comput. Eng., vol. 11, no. 3, pp. 2350– 2359, 2021.

[6]. Z. A. Almusaylim, N. Z. Jhanjhi, and A. Alhumam, “Detection and mitigation of RPL rank and version number attacks in the internet of things: SRPL-RP,” Sensors (Switzerland), vol. 20, no. 21, pp. 1–25, 2020.

[7]. A. Jahangeer, S. U. Bazai, S. Aslam, S. Marjan, M. Anas and S. H. Hashemi, "A Review on the Security of IoT Networks: From Network Layer’s Perspective," in IEEE Access, doi: 10.1109/ACCESS.2023.3246180.

[8]. W. Fang, W. Zhang, W. Chen, T. Pan, Y. Ni, and Y. Yang, “TrustBased Attack and Defense in Wireless Sensor Networks: A Survey,” Wirel. Commun. Mob. Comput., vol. 2020, 2020.

[9]. S. R. Taghanaki, S. B. Arzandeh, and A. Bohlooli, “A Decentralized Method for Detecting Clone ID Attacks on the Internet of Things,” Proc. 2021 5th Int. Conf. Internet Things Appl. IoT 2021, 2021.

[10]. Q. Zhang and W. Zhang, “Accurate detection of selective forwarding attack in wireless sensor networks,” Int. J. Distrib. Sens. Networks, vol. 15, no. 1, 2019.

[11]. Muneer Bani Yassein, I. Hmeidi, Y. Khamayseh, M. Al-Rousan, and D. Arrabi, “Black Hole Attack Security Issues, Challenges & Solution in Manet,” no. April 2019, pp. 199–207, 2018.

[12]. Gulzar, C.M., Kurnool, & Kashyap, R. (2015). Prevention of Black Hole Attack in MANET.

[13]. O. R. Ahutu and H. El-Ocla, “Centralized Routing Protocol for Detecting Wormhole Attacks in Wireless Sensor Networks,” IEEE Access, vol. 8, pp. 63270–63282, 2020.

[14]. https://www.tutorialspoint.com/internet_of_things/internet_of_things_contiki.htm accessed at March 20th, 2023.

[15]. Derogarian, Fardin. (2015). Design of a Body Sensor Network Embedded in Textiles for Biomedical Applications. 10.13140/RG.2.1.1192.3920.

[16]. https://phdinfo.org/contiki%20cooja%20wsn%20simulator.html accessed at April 3rd, 2023.

[17]. File STACK%20Project%20profile%20leaflet.pdf accessed at April 23rd, 2023.

[18]. https://agile.ro/stack/about/ accessed at April 23rd, 2023.

Vulnerability Scanner: Web-based Security Testing

Andrei-Daniel ANDRONESCU, Ioana-Ilona BRĂSLAȘU, Dumitru-Iulian NĂSTAC
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
andronescu.andreidaniel@gmail.com, ioanabraslasu2000@gmail.com, iulian.nastac@upb.ro

Abstract

As the use of internet-based software increased, cybersecurity has emerged as a major issue in the current world. The fast-paced technology innovations allowed most companies to scale their business, consumers to access easier their favorite products, thus increasing the reliance on web-based software. The importance of web security cannot be emphasized given the increase in cybercrime and the damage it poses to businesses, people, and governments. This paper proposes an automated solution capable of detecting and exploiting common vulnerabilities found on web-based software, this being done without performing any malicious intended operations. By using software capable of automatically detecting the means a client could communicate with a server, users can ensure that a thorough verification is done on their web-applications, revealing the blind spots that developers may have overlook.

Index terms: Chromium, File Inclusion Attacks, NodeJS, Puppeteer, SQL injection, vulnerability scanner, web application security, testing

1. Introduction

In today's age, the internet is responsible of most business operations. It provides the companies the ability of communicating and collaborating with customers and partners instantly. It offers the companies the access to a huge amount of data, that, if understood correctly, can improve its revenue and its overall performance. As a result, every modern company is in some shape or form, depended on the internet. But, as the saying goes, "with great power, comes great responsibility". The responsibility in their case is ensuring the security of the company's data as well as customer's sensitive data. One simple easily accessible vulnerability in the system could cause the loss of customer's trust, the loss of competitive edge or even the company itself.

Malicious actors are always seeking for those opportunities. It is well known that there are armies of bots (botnets [1]), constantly scanning the internet for newly deployed servers. Everything that you expose on the internet should be protected, and security measures must be taken way before the service becomes public. Otherwise, malicious actors can easily plant malicious content on your server, plant backdoors allowing them later access, thus exposing your own internal network.

It's always easier to prevent, than treat. Ensuring that a website is secure before making it public can save you from lots of unwanted attention. Therefore, in our paper, we're trying to propose a testing software, capable of detecting common attacks, that bots or even real people will use. Those common attacks consist of SQL injections, cross-site scripting, and buffer overflows.

2. Design

Generally, websites offer clients the possibility of interacting with the server through login forms, search forms, etc. This, in turn, allows the users to have access (although limited) to the back-end systems of a website through specific APIs or HTTP requests. Web applications are becoming more and more of a popular target for attackers, and they will try to exploit the previously mentioned methods in order to gain information and control. Ensuring that web applications are secure is a critical step that needs to be taken by organizations to keep their client's trust. One way that this security can be ensured is by using a vulnerability scanner that can identify potential safety issues in web applications.

In this paper, we propose one such vulnerability scanner that is capable of detecting potential security flaws in web-based applications. Our scanner utilizes a five-step process, which includes a website crawling module, the processing of input forms, a malicious payload generator, a payload injection phase and response analysis. The diagram of the system can be observed in Fig 1.

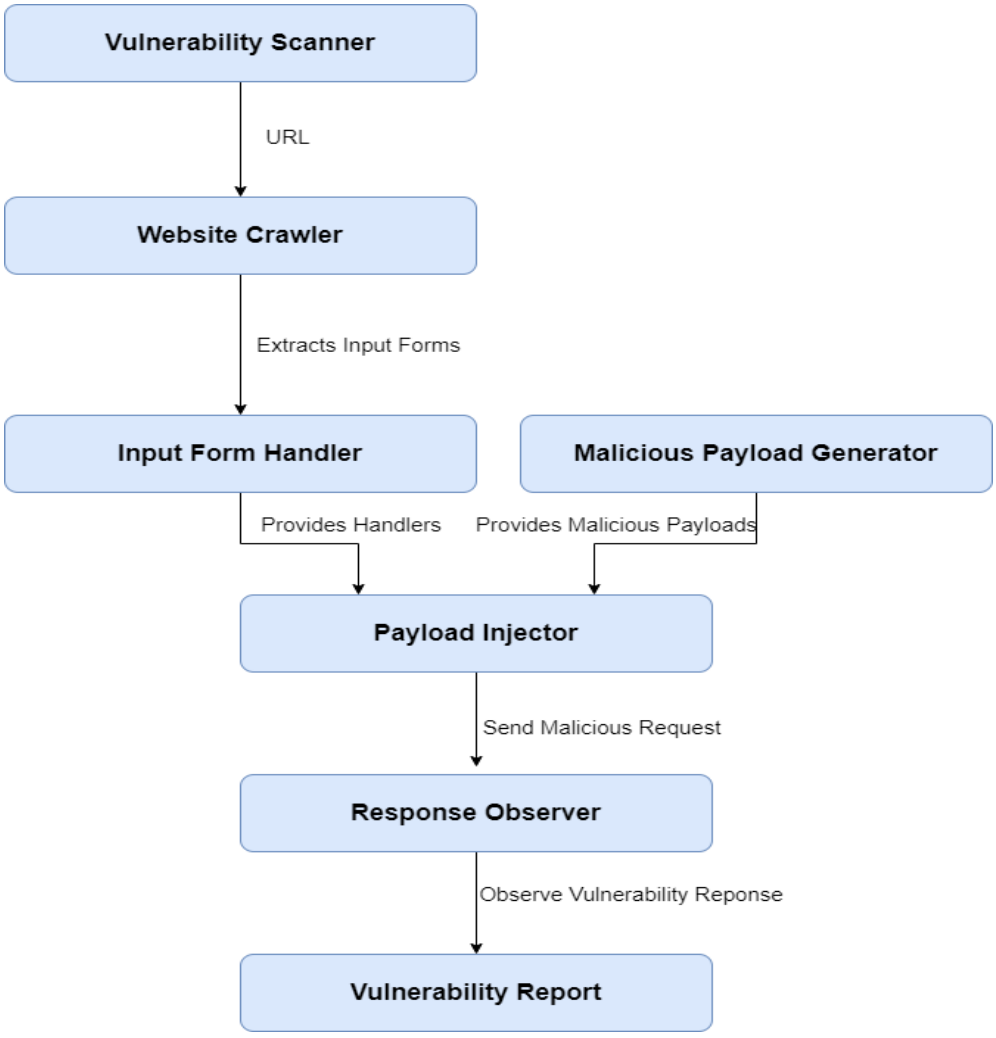


Fig. 1. Vulnerability Scanner diagram for a single web page

2.1. Web Crawling Module

In this article, we propose an exhaustive method of scanning the vulnerabilities of a web server. For this reason, it is important to be able to verify each easily accessible communication bridge between client and server. Thus, a web crawler is used in order to extract all the input forms available on the specified web page. It's also important to learn that a web domain might contain many resources, and attackers always try to find the weak link in a domain to gain access. One old server

running on an outdated version of Apache, one server running on a Java backend taking use of an outdated Log4j module [2] will be the primary target of an attack. Therefore, a scanning process is necessary when using a vulnerability scanner. Our implementation of resource scanning is performed not only on the given web page, but rather in a recursive manner. The web page represents only the starting point in the reconnaissance mission. Links to other resources are scraped out of each web page, and if the link is part of the same domain, it will be followed.

It is well known that modern websites contain some sort of protection against automated software that try to access it (Captcha2, ads, honeypots, automated bot prevention solutions, hidden fields, IP monitors). Bypassing those security measures is a complex task, and it does not represent the scope of our paper. However, in the development process of the web crawler, it was observed that a good website accessing rate can be achieved by removing ads present on most web sites, by using an Ad Blocking browser extension. One other important factor that impacted the accessing rate of websites using automated software was a security feature hosted on the websites that inspects the client's web browser profile. To mitigate this issue, a real web browser profile was used, profile that was manually created and exported for later use in each crawling process.

The crawling process was built on top of the Chromium web browser. A Chromium instance was launched by the NodeJS back-end and controlled with the help of the *puppeteer* module in order to simulate human-like behavior. By using this configuration, we ensured that we had access to a broader range of websites, mostly because it supports scraping from dynamic web pages. Once accessing a web page, the extraction process occurs by identifying the elements of interest such as input forms and links with the use of selectors. Links are stored for later use, when the testing process is fulfilled for the current web page. From the input forms, event handlers are extracted, handlers that can be used to populate the input fields, with the specified malicious payload and to submit those payloads.

2.2. Malicious Payload Generator

The Malicious Payload Generator represents a separate, highly scalable, part our system. It is critical component in our Vulnerability Scanner, being responsible of generating HTTP-based attack payloads. The malicious load must be carefully designed when used against real life websites. Only payloads that can validate the presence of a vulnerability in a system must be used in this case such that no malicious actions are performed. This type of attacks is also known as "nospoilt".

This module generates a variety of attacks which are determined by the web application under scan and the type of input fields. Numerous attacks can be produced, including the following most typical ones:

- **SQL Injection Attacks**. Significant data breaches and confidential information loss can result from those types of attacks [3]. They mostly consist in the injection of SQL code in the input fields that interact with a database. Those codes are then executed as an SQL query and the result returned to the malicious actor. This allows him to gain access to secret information such as user credentials, to update the prices on an e-commerce website, or to obtain Business Intelligence regarding the company's financial situation. This is why, servers must parse queries sent by the users, in order to be sure that those type of attacks cannot be executed.
- **XSS attacks**. Modern web servers are using dynamic web pages, using JavaScript that provides most of the robust functionality [4]. XSS attacks (also known as cross-site scripting attacks) are executed when unsuspecting users access the affected page. The attack consists of injection of malicious script in a victim's browser [5], resulting in the tracking of user's activity, or performing other malicious tasks.

- **Command Injection Attacks**. The input fields are the primary targets for those type of attacks. They inject malicious commands that the server executes due to the insufficient input validation [6], using the privileges of the web application. A successful attack can lead to a full system compromise and data leaks, thus posing serious threats to a web application.
- **File Inclusion Attacks**. Local File Inclusion (LFI) attacks generally occur when a web server allows users to access files on a web server. Once an attacker gains the means of accessing files on the server directly from the browser [7] (example: finding a php script that can open a file on the system), the system is compromised. Remote file inclusion (RFI) attacks on the other hand, are performed by exploiting web server functions that can reference different resources, outside of the domain. This allows attackers to upload malware on the server and generally result in information theft and site takeovers [8].

2.3. Payload injector

The payload injector is a module that is responsible for receiving the input handler from the web crawler, identify the type of input type being used. With this obtained information, it can perform a request on the Malicious Payload Generator module to retrieve a suitable payload, that can be used upon the handler.

The payload injector sends the HTTP request to the server and awaits for a response. One important thing that must be taken into account is that for a successful and thorough attack performed on the server, the payload injector needs to be find tuned on the specific website, with thorough understanding of the web application under test. Our paper proposes a general approach, more oriented towards a general testing scenario. With this in mind, if the testing of a specific domain needs to be performed, a fine tuning on the payload injector needs to be done. Additionally, it is extremely important to use a custom implementation of the payload injector in a responsible manner, by requiring permission from the owner. In the development process of this project, no servers were used without proper clearance, and with explicit malicious intent.

2.4. Response Observer and Vulnerability Report

The Response observer module is responsible for detecting if an attack was successful. This is done by analyzing the server's response for the HTTP request sent by the Payload Injector. Each type of attack would have its characteristic response type. This is an important challenge to overcome when designing an automated vulnerability scanner, since each type of attack has a different outcome, and even the same attack can have different outcomes when tried against different systems. For this reason, detection heuristics fine-tuned for each attack were added. A general overview of the detection heuristics that can be used for detecting attacks will be presented in the following lines.

In the case of an SQL injection attack, the module can analyze the response data received, and identify if any anomalies are present. It will look for specific signatures that represent a successful SQL attack, such as error messages, unexpected status codes, database information that would contain specific keywords in relation to the provided malicious input. If any of those signals occur, the response observer will treat the SQL injection as successful.

In the case of XSS attacks, it is important to note that automatically detecting if an XSS attack successful is a serious challenge since the behavior of those attacks is different. One common scenario is that the attacker tries to insert a script code in order to get some unauthorized data. This XSS attack is known to be present if an unexpected <script> tag is present in an input location. Another method for checking if the XSS attack is present is by verifying with a network traffic monitor if the requests are sent to a location outside the website's domain.

Detection of a successful command injection attack depends mostly on system that has been targeted and on the type of command itself. Verifications if data has been leaked are observed, or verifications if errors are observed in the case of intentionally erroneous injected commands.

File inclusion attacks are also hard to detect if they were executed successfully. In general, insertion of some malicious files that could trigger to be executed by the server may have some sort of expected behavior which allows us to verify its effectiveness. For example, trying to access the /etc/passwd file in the server's filesystem might return a string containing the word "root" [9], or inserting a php wrapper containing a URL to our own website can be used to detect if the attacks were successful.

3. Conclusion

In conclusion, the automation of identifying vulnerabilities in web applications is studied in this article. This project implied the incorporation of Node.js, Chromium, and Puppeteer to create a modern and capable vulnerability scanner. The scanner extracts input forms and navigates additional links to generate HTTP-based attacks using the Malicious Payload Generator module and then tests their effectiveness with the Payload Injector module. The Response Observer, with its specific criteria for each attack type, determines whether the attacks were successful.

The project automates the testing of web application security to help developers pinpoint possible vulnerabilities and tackle them before hackers can capitalize on them. It is possible to tailor the attacks to fit individual needs and requirements with the tool. Through the automation of testing capabilities, a multitude of vulnerabilities can be detected and addressed early in the software development process. Ultimately, this can help organizations reduce their vulnerability to cyber-attacks and improve the security of their applications.

The efficiency of automated vulnerability testing software is directly proportional to the amount of attack types it can perform, detect and report on. Future work will focus on this idea, of increasing its capabilities in terms of the number of attacks it can perform and refactoring the actual project in order to make it more scalable and capable of bypassing more bot detection mechanisms.

Finally, our research team conducted this study with ethical considerations, and we ensured that no live servers were compromised or damaged in the process of developing this paper.

References

[1]. P. Sabanal, IBM. Thingbots: The Future of Botnets in the Internet of Things. [Online]. Available: https://securityintelligence.com/thingbots-the-future-of-botnets-in-the-internet-of-things/ [Accessed: Apr. 20, 2023].

[2]. National Institute of Standards and Technology - CVE-2021-44228 [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2021-44228 [Accessed: Apr. 21, 2023].

[3]. Stuard McDonald, SQL Injection: Modes of Attack, Defence, and Why It Matters [Online]. Available: https://www.sans.org/white-papers/23/ [Accessed: Apr: 21, 2023].

[4]. Sucuri. Cross-Site Scripting (XSS) Attacks. [Online]. Available: https://sucuri.net/guides/what-is-cross-site-scripting/ [Accessed. Apr. 21, 2023].

[5]. Kirsten S., Cross Site Scripting (XSS). [Online]. Available: https://owasp.org/www-community/attacks/xss/ [Accessed Apr. 21, 2023].

[6]. Weilin Zhong, OWASP. Command Injection. [Online]. Available: https://owasp.org/www-community/attacks/Command_Injection [Accessed Apr. 21, 2023].

[7]. Admir Dizdar (9 July 2021). LFI Attack: Real Life Attacks and Attack examples. [Online]. Available: brightsec.com/blog/lfi-attack-real-life-attacks-and-attack-examples/ [Accessed Apr. 21, 2023].

[8]. Imperva. Remote file inclusion (RFI). [Online]. Available: https://www.imperva.com /learn/application-security/rfi-remote-file-inclusion/ [Accessed Apr. 21, 2023].

[9]. Local File Inclusion (LFI) – Web Application Penetration Testing. [Online]. Available: https://medium.com/@Aptive/local-file-inclusion-lfi-web-application-penetration-testing-cc9dc8dd3601 [Accessed Apr. 22, 2023].

Ensuring the Security of a Communication Network through Resilience. Mathematical Modeling

Constantin-Alin COPACI[1], Dorina-Luminiţa COPACI[2]
[1] IT Expert, ANCOM, Bucharest, Romania
acopaci@yahoo.com
[2] Associate Professor, University Politehnica Bucharest - ETTI, Bucharest, Romania
lcopaci@yahoo.com

Abstract

Many of the network computing systems used in various organizations are not resilient enough to withstand attacks and failures. The performance of these networks is degraded by failures. Thus, it is important to develop techniques for designing and implementing resilient service-oriented networks that can survive attacks and failures, as well as continue to provide a reasonable level of service. This paper considers the mathematical modeling using graph theory of resilience in service-oriented communication networks. The objective of this paper is to develop the concept of service-oriented resilient system as well as to identify the metrics used to quantify resilience to node and edge failures. Using these metrics, we will choose an appropriate network topology and/or an optimal distribution of services in the network.

Index terms: edge resilience, graph, node resilience, restoration, service-oriented network

1. Introduction

The resilience [1], [2] approach can be compared to ensuring the quality of service. Network operators can provide resiliency as an added value to the service. In addition, important, critical, driving and emergency usable services can be protected with an additional layer of protection against disasters or terrorist attacks.

Further on, we will consider resilient systems as systems that restore their function after a failure. Thus, such systems have an intrinsically high dependability. The term dependability is a complex notion that includes the following topics: reliability, availability, confidentiality, safety and security [3]. One of the most widely used techniques for increasing the dependability of systems is the implementation of the fault-tolerance [4], [5] a technique to be considered in this paper.

The network resilience - the ability to provide and maintain an acceptable service level in the presence of (random or deliberate) failures - becomes more and more important. A resilient network should be able to cope with a specific number of failures by remaining completely functional, providing connectivity to all of its parts and providing enough capacity to fulfill its task.

Resilience can be achieved either reactively by ***restoration*** or proactively by ***protection*** methods [6]. Restoration requires a reaction only upon the occurrence of an error. Protection in contrast prepares means of correction through additional redundant information before a failure occurs, and often does not even need retransmissions.

Service-oriented communication networks, which we will focus on next, have interesting properties in terms of resilience [7]. This research provides a decomposition of the resilience concept

into two categories: challenge tolerance and trustworthiness. This is graphically illustrated in the figure 1 [8]:

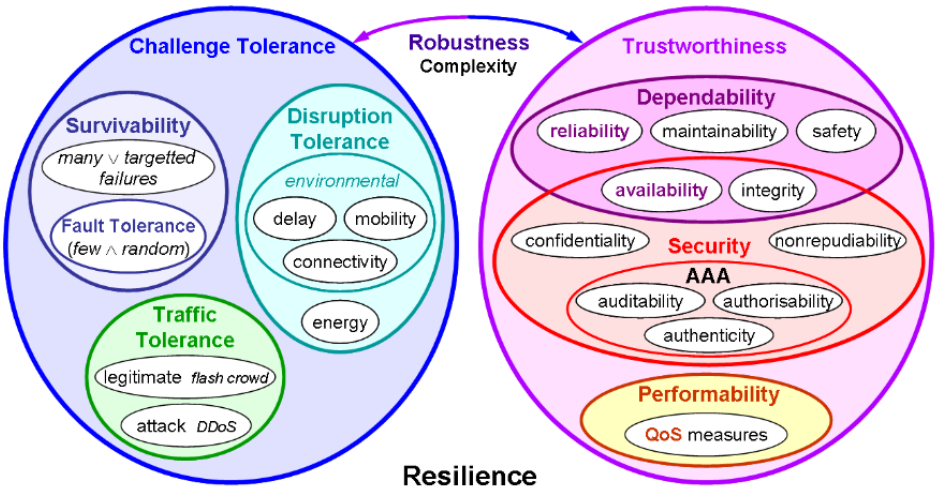


Fig. 1. Challenge tolerance and trustworthiness [8]

This article is organized as follows: In Section 2, information is presented on important graph properties and possible classes of graphs in communication networks. In Section 3, methods to improve resilience for service-oriented communication networks are presented. In this section, we also present algorithms for analyzing the resilience of a service-oriented network. Finally, Section 4 summarizes the main findings of the article.

2. Graph theoretical background

Any network can be modeled as a (directed) graph G consisting of vertices or nodes V and edges or links U. Edges may be weighted to either represent communication capacities, or communication costs or delays [9], [10].

Resilience is defined as the ability to maintain a network service under interference. Since many of these services depend on the reach ability of nodes, connectivity measures certainly belong to the most important graph properties.

The edge connectivity α and the vertex connectivity μ are the minimum number of edges (vertices) that need to fail, to separate the graph into at least 2 components and hence are worst-case statistics of resilience. So, α-1 and μ-1 are the numbers of edges (vertices) which may always be removed, without disconnecting the graph. The edge connectivity equals the size of a minimum cut of the graph and is bounded from above by the minimum degree of a vertex.

The shortest path between two vertices s and t is a set of edges connecting s and t and having a minimum sum of edge weights. Let the distance d (s, t) be the weight of the shortest s-t-path and the distance between unconnected vertices defined to be infinite. The diameter of a graph $diam(G) = max_{s,t\epsilon V} d(s,t)$ then is the length of the longest shortest path between any two vertices. Clearly, the diameter influences the time of information distribution in the whole network.

3. The resilience in service-oriented networks

A service-oriented architecture [11], [12], [13] involves breaking an application's functionality into smaller, distinct units - called services - that can be distributed across a network and used together to create business applications. The high capacity with which these services can be reused in different applications is a characteristic of service-based architectures. These services communicate with each

other by sending information from one service to another. In such an architecture, each computing system is associated with two sets of services, local services and network services.

3.1. Formal models of the formulated problems

In this part of the research, we developed formal models of the formulated problems. On this basis, we implemented a series of algorithms for the design of resilient edges, respectively of resilient nodes.

We consider the graph G (V, U) that abstracts the network. This is used to design a resilient network. The set V represents the set of nodes of the graph corresponding to the network, and U the set of edges in the graph G.

For each node $v_i \in V$ and for each edge $u_{ij} \in U$ we define the set $N(v_i)$ of services required by each v_i. We denote the cost matrix by $C=[c_{ij}]$, where c_{ij} represents the cost of introducing service s_j at node v_i. The cost matrix is a matrix with positive elements.

$$c_{ij} = \begin{cases} 0, & \text{if there is no service introduction cost} \\ 1, & \text{if there is a service introduction cost} \end{cases}$$

We determine the set of services for each node in the graph so that the resulting network has the required level of edge or node resilience and the total cost of placing the services is minimal. In this sense, we consider s_j a certain service. We define each node v_i such that $s_j \in V(v_i)$ is a demand point for s_j. The node for which the service s_j exists is called a *service point*. A lot of service points for s_j is called *placement* for s_j.

Given the connected graph G (V, U) and a placement P for the service s_j, then the placement will be the edge resilience with respect to the service s_j if for each edge $u \in U$, the graph G'(V,U-{u}) contains a road from each demand point for s_j to a service point for s_j.

We consider the ten nodes network shown in Figure 2.

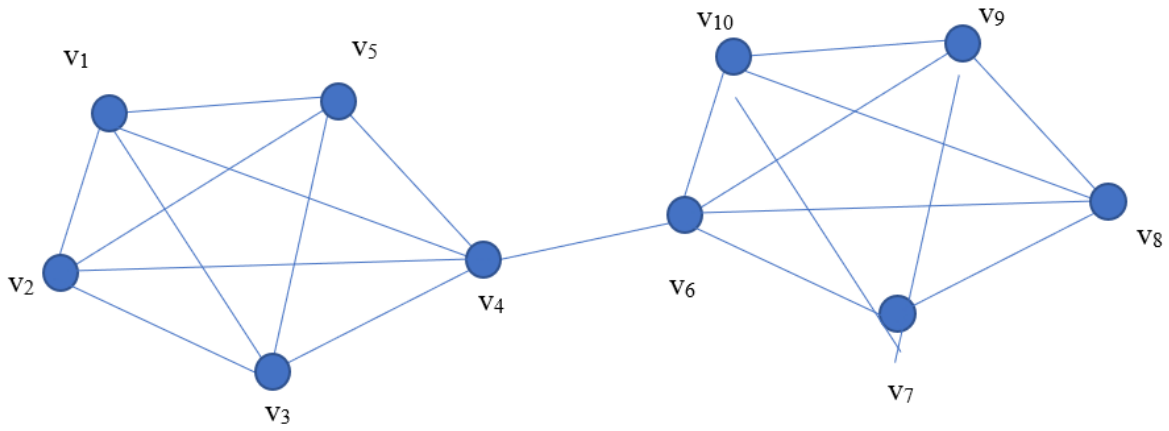


Fig. 2. The graph representing a network with 10 nodes

The ten services provided by the network are denoted by s_1 through s_{10}. For each node v_i, we define the set of services available to v_i respectively the set of services required by the node v_i, $1 \le i \le 10$. For example, for i=1, the set of services available to v_1 is $\{s_1,s_2,s_3,s_4\}$ and the set of services required by the node v_1 is $\{s_5\}$. The node connectivity and the edge connectivity of the network are both one, since the network can be disconnected by removing one node (for example, the node v_4) or one edge (the edge $\{v_4,v_6\}$). However, the node and edge resilience parameters of the network are both two. In particular, the subnetworks obtained by deleting the edge $\{v_4,v_6\}$ or one of the nodes v_4 and v_6 are all self-sufficient. It can be verified that no matter which pair of vertices or which pair of edges is deleted, each of the resulting subnetworks is self-sufficient. However, when the four edges $\{v_1, v_2\}$, $\{v_1, v_3\}$, $\{v_1, v_4\}$, $\{v_1, v_5\}$, are deleted, the subnetwork containing only the node v_1 is deficient, since it does

not have access to service s_4. Likewise, when the three nodes v_1, v_2, v_3, v_4 are deleted, the subnetwork containing only the node v_5 is deficient, since it does not have access to service s_1.

3.1.1. Mathematical modeling for the resilience of network edges

In this section, we implement the algorithm for calculating the edge resilience of a service-oriented communication network. Also starting from the definitions in graph theory, we considered G (V; E) the graph of the given network. We used the term subnet to understand a connected subgraph of G.

A service-oriented network subnet is defective with respect to a network service if there is a node in the subnet that requires the given service, but there is no node in the subnet that provides that service. In this situation, a subnet is deficient if there is a service for which the subnet is deficient.

Determining the resilience of service-oriented communication network edge presupposes the determination of the set of available services as well as the set of services required for each node [14], [15].

The minimum number of deficient edges related to the service s_i can be obtained by calculating the minimum cut of the edge with the minimum weight in the auxiliary graph G_i for each pair s-v, where v is a node that requires the service si. Once this value is found, the edge resilience of the given network G can be determined by considering the minimum for all services. These observations lead to the algorithm [14], [15] for computing the edge resilience of a given service-oriented network.

To implement the algorithm we will consider a network G(V, E), the set S of all services of the network, the set of available services as well as the set of services required for each node $v \epsilon V$. We will find the edge resilience of G.

The algorithm assumes that for each service $s_i \in S$ to construct auxiliary graph G_i (V_i, E_i) for service s_i; to find the set $D_i \in V_i$ of demand points for service s_i;

- For each node $v \in D_i$ to compute $\alpha_{v,i}$, the minimum weight of an s-v edge cutest in G_i;
- Let. $\sigma_i = min\{\alpha_{v,i} : v \in D_i\}$. Edge resilience of $G = min\{\sigma_i : s_i \in S\} - 1$.

Determining the edge resilience of a graph that abstracts a service-oriented communication network:

```
[…]
Edge::Edge(int servicii[],int NrServicii,int disponibil[100][100],int necesar[100][100])
                {       g = new graf(10,1);
                        g->populeazaGraf();
                        g->afiseazaMuchii();
                        printf("Initial");
                        this->NrServicii = NrServicii;
                        int i;
                        for(i = 0; i < NrServicii; i++)
                                this->servicii[i] = servicii[i];
                        int j;
                        for(i = 0; i < this->g->getNumarNoduri(); i++)
                                for(j = 0; j < this->NrServicii; j++)
                                        this->disponibil[i][j] = disponibil[i][j];

                        for(i = 0; i < this->g->getNumarNoduri(); i++)
                                for(j = 0; j < this->NrServicii; j++)
                                        this->necesar[i][j] = necesar[i][j];}
void Edge::algoritm(int type)
                { //type = 0 -> edge resilience
                …;
                        for(i = 0; i < NrServicii; i++)
                        { // we determine the source nodes for sj service
                                for(j = 0; j < n; j++)
                                        surse[j] = -1;
```

```
                        nr = 0;
                        for(j = 0; j < this->g->getNumarNoduri(); j++)
                                for(k = 0; k < NrServicii;k++)
                                        if(disponibil[j][k] == servicii[i])
                                        {surse[nr++] = j;
                                                break;}
                ...;
                        //we calculate edge resilience
                                if(type == 0)
                                        grafAux->grafAuxiliar(surse,NrSurse);
                                else
                                        grafAux->grafAuxiliarNoduri(surse,NrSurse);
                                printf("Matricea de costuri pentru graful auxiliar(ultimul nod corespunde
serviciului)\n");
                                grafAux->afiseazaMuchii();
                                printf("\n\n");
                        ....;
                                int minAlfa = 100000000,aux;
                                // cutset min
                                for(j = 0; j < NrConsumatori; j++)
                                {type == 0;
                                aux = grafAux->minCutSet(grafAux->getNumarNoduri() -
1,consumatori[j],minAlfa,0,0);
                                if(aux < minAlfa)
                                        minAlfa = aux;}
                                if(minAlfa < minSigma)
                                        minSigma = minAlfa;}
                        minSigma--;
                        printf("Rezilienta muchiei  %i ",minSigma);}
```

To estimate the running time of the algorithm, we assumed that the given network has x nodes, y edges and a total of r services. We determined the running time of the algorithm by calculating the minimum cut of the edge with minimum weight. For each service si the algorithm computes the minimum cut O(x). Thus, the total number of calculations is O(rx). Since each computation of the minimum edge cut can be done in O(y+xlogx) time, the running time of the algorithm is O(rx(y+xlogx)).

3.1.2. Mathematical modeling for the resilience of network nodes

Our algorithm for computing the node resilience of a service-oriented network follows the same approach as that of edge resilience. The main difference is that we need to work with node cutsets instead of edge cutsets.

As in the case of determining the resilience of the edge, and in the case of calculating the resilience of the node, we used the auxiliary graph, except that the weights of the crowd were not used, and the weight of each node was considered equal to 1.

When a subset X of nodes is deleted from a graph G (V; E), each edge incident on a node in X is also deleted. Keeping this in mind, it is straightforward to modify the definition of a deficiency inducing edge set to obtain the definition of a deficiency inducing node set.

We considered a service-oriented network G (V, E) and sj a given service. We denote by Gj(Vj,Ej) the auxiliary graph (with node weights) for sj. For any node, the minimum number of nodes to be deleted from G such that there is no path between v and any node providing service sj is equal to the weight of the minimum cut s-v with the minimum weight in Gj [14], [15]. The rest of the calculation is similar to that of edge strength.

4. Summary

This article provides an overview of resilience mechanisms in service-oriented communication networks. We reviewed different concepts for improving resilience in service-oriented communication networks. We developed the concept of a service-oriented system resilient to node and edge failures and determined the metrics for ensuring the security and resilience of service-oriented systems.

Different resilience mechanisms have advantages and disadvantages. The benefit of the resilience strategy depends on the specifics of the network and how much importance a network operator attaches to performance metrics.

References

[1]. V. Kuikka, Modeling Network Resilience and Utility of Services. In Proceedings of the 2019 IEEE International Systems Conference (SysCon), Orlando, FL, USA, 8–11 April 2019.
[2]. http://en.wikipedia.org/wiki/Resilience.
[3]. I.C. Bacivarov, V. Cătuneanu, Fiabilitatea sistemelor de telecomunicaţii, Ed. Militară, 1995.
[4]. xxx Proceedings of IEEE Symposium on Fault Tolerant Computers, 1990-2010.
[5]. R. Albert, H. Jeong, and A.-L. Barabasi, Error and attack tolerance of complex networks, *Nature*, 406(6794):378-382, July 2000.
[6]. S. Lee, Y. Yu, S. Nelakuditi, Z.-L. Zhang, and C.-N. Chuah. “Proactive vs reactive approaches to failure resilient routing.” *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, March 2004.
[7]. D. J. Rosenkrantz,S. Goel, S. S. Ravi, J. Gangolly: Structure-Based Resilience Metrics for Service-Oriented Networks, October 11, 2004.
[8]. https://www.enisa.europa.eu, “Measurement Frameworks and Metrics for Resilient Networks and Services: Technical report”, February 2011.
[9]. M. A. Henning, J. H. van Vuuren, Graph and Network Theory – An Applied Approach using Mathematica, Springer Cham, ISBN: 978-3-031-03857-0.
[10]. F. Chung and L. Lu. “The average distance in a random graph with given expected degree.” Internet Mathematics, 1(1):91-114, 2002.
[11]. M. Aly; M. Franke (2016). "Service Oriented Interactive Media (SOIM) Engines Enabled by Optimized Resource Sharing". 2016 IEEE Symposium on Service-Oriented System Engineering (SOSE): 231–237. Retrieved February 9, 2021.
[12]. F. Glinka; A Raed (2009). "A Service-Oriented Interface for Highly Interactive Distributed Applications". European Conference on Parallel Processing. ISBN 978-3-642-14121-8. Retrieved February 9, 2021.
[13]. Dieter Hildebrandt; Jan Klimke (2011). "Service-oriented interactive 3D visualization of massive 3D city models on thin clients". COM.Geo '11: Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications. COM.Geo '11: 1. doi:10.1145/1999320.1999326. ISBN 9781450306812. S2CID 53246415. Retrieved February 9, 2021.
[14]. D. J. Rosenkrantz, S. Goel S. S. Ravi, J. Gangolly, “Structure-Based Resilience Metrics for Service-Oriented Networks”, October 2004.
[15]. A. Kvalbein, A. F. Hansen, T. Cicic, S. Gjessing, and O. Lysne. “Fast recovery from link failures using resilient routing layers.” *10th IEEE Symposium on Computers and Communications, ISCC 2005,* June 2005.

Enhancing EU Cyber Defense Through Hardware Trojans Detection Capabilities

Vasile-Florin POPESCU[1], Victor GÂNSAC[2], Olivia COMȘA[3], Cristian ICHIMESCU[4], Dănuț TURCU[5], George BUCĂȚA[6]

[1, 4, 5] "Carol I" National Defence University of Bucharest, Faculty of Security and Defence
popescu.vflorin@unap.ro, ichimescu.cristian@unap.ro

[2, 3] SAFETECH Innovation
victor.gansac@safetech.ro, olivia.comsa@safetech.ro

[6] "Nicolae Bălcescu" Land Forces Academy of Sibiu, Faculty of Military Management

Abstract

Software Trojans and cybersecurity are a concern worldwide. Hardware Trojans are likely to be an issue faced by the Defence Industry of all countries. Information on how defense industry stakeholders deal with HT in Defense Products is by nature scarce or even inaccessible. It is however fair to assume that they adapt and use IC RE methodologies, notably some developed for IP infringement, to search for HTs. With these RE methodologies, checking a chip after its fabrication implies to deconstruct and analyze the whole surface and all the layers of a chip. It is thus hard to know for sure which states has acquired Hardware Trojan detection capabilities. There are however indications that some States could be in the process of acquiring such capabilities.

Index terms: defense industry, hardware Trojans, System-on-Chip, reverse engineering, image acquisition

1. Introduction

Communication technologies, and computing systems are an integral part of today's defense systems. They process security-critical information or perform calculations critical to the success of a defense operation or the safety of the military personnel on the ground. As such, they represent an obvious target for malevolent entities, including nation state adversaries. Software Trojan, virus and malware, are already well known cybersecurity threats, and many countermeasures are available or under continuous development. Hardware Trojan (HT), on the other hand, are becoming a recognized and immediate cybersecurity threat, but no available credible countermeasures are yet available to the European Defense Industry.

As such, they represent an obvious target for malevolent entities, including nation state adversaries. Defense electronics rely on complex System On Chips (SoC) that combine semi-conductors with multiple functions on a single Integrated Circuits (IC).

To reduce cost and time to market, these SoC are designed through a horizontal process and manufactured globally. This implies the intervention of several, potentially untrusted, stakeholders worldwide and comes with an increased security risks. Among these risks, the introduction of Hardware Trojans (HT) that could leak information or alter the functioning of a Defense system, is emerging as an immediate cybersecurity threat.

The European Industry as a whole retains the capacity to design SoC, thereby securing this step of the process. On the manufacturing side however, also called post-Tape Out, the situation is dire.

80% of SoCs are manufactured in Asia and some critical post-Tape Out steps are only provided by non EU suppliers. According to the study carried out by Hepp et al., 2022, demonstrates that the insertion of sophisticated HTs evading all routine tests is possible in less than 24 hours during manufacturing. This highlights the urgent need to rapidly implement large scale solutions to detect HT in SoC manufactured outside of the EU.

The EU's Defense Technological and Industrial Base (EDTIB)'s goal is to build a stronger and more competitive European industry. Communication technologies, and computing systems are ubiquitous both in current Defense products and in Defense technologies in development. Reliability and safety, including safety of information, are major features of defense products. To stay strong and competitive the European Defense Industry must therefore ensure that the Integrated Circuits at the core of communication technologies, and computing systems are safe and reliable. As of today the EDTIB is not autonomous in this as:

- ✓ it does not dispose of an industrially deployable technology allowing to check for HT presence in SoCs, while other Nation states might be developing such capabilities.
- ✓ it is dependent on non EU suppliers for ICs manufacturing. This is particularly obvious since 2020 and the COVID-19 pandemic, and the resulting world shortages coupled with higher prices of semiconductors.

To build a strong and competitive Defense industry, the EDTIB is dependent on demand, supportive political conditions, especially on the European level, but also on reliable and secure supply chains on the European continent as well as globally. There is a strong dependence of the European industry as a whole to foreign IC manufacturer, and the European defense industry is no exception. To be autonomous the EDTIB must rely on a European ICs supply chain. To act on this industrial dependence at the European level, the European Chips Act aims to centralize the production of chips in Europe instead of foreign countries. In February 2022, the European Commission has published a factsheet about the European Chips Act and its goals (Ludwig M, et al., 2022). The plan is split into three period of time. The HARTROID outcomes will support the EDTIB along the three phases of the constitution of a European IC supply chain, thus directly contributing to its autonomy.

The research methodology was represented by specialized literature analysis, corroborated with European directives from EU's Defense Technological and Industrial Base (EDTIB) and European Defense Fund (EDF).

2. Up-to-date analysis of the hardware Trojans identification studies

Considering the huge exposure of security systems vis- a vis hardware Trojans used in chips, R&D initiatives have started at the level of the European Union to come up with a sustainable solution to these threats, as follows:

- ✓ **EPoCH** project (H2020#695022) https://doi.org/10.3030/695022 which has as its primary purpose to develop of an open source software tool able to display and compare circuit netlists extracted from reverse engineering works.Although the extracted netlist is a very reliable way to find a Trojan, extracting the netlist of a high-end, multi-layered integrated circuit is a task that can easily take many months to accomplish.
- ✓ **SAFEST** project (H2020#952252) https://doi.org/10.3030/952252,https://safest.taltech.ee which approaches networking strategy on hardware security focused on testing practices, reverse engineering and hardware-based defenses.
- ✓ **EXFILES** project (H2020#883156) https://doi.org/10.3030/883156; https://exfiles.eu/ which has as its main objective research development of full IC reverse engineering techniques (de-processing, imaging) to get knowledge of encrypted mobile phone SoCs, in order to find

vulnerabilities that can be exploited by LEAs to get access to the stored information.

- ✓ **Codasip** High-end processor IP and high-level design tools for RISC-V (GA: 19010116) https://doi.org/10.3030/190101116
  Codasip offers a unique combination of semiconductor processor IP based on the RISC-V open instruction set architecture (ISA) and high- level EDA tool Codasip Studio providing outstanding flexibility and 5x faster time to market. RISC-V ISA can be used for a wide variety of applications ranging from low power and low gate count embedded cores to advanced high frequency application cores.
- ✓ **EXCEED** project - https://www.exceed-padr.eu/ The EXCEED project aims at creating a European supply chain of reconfigurable, flexible and trustable programmable system-on-a-chip family targeting a number of ruggedized and secure defense applications.
- ✓ **Intelligent Reliability 4.0** project (H2020#876659) https://doi.org/10.3030/876659; https://www.irel40.eu/
  The iRel4.0 project aims to reduce the failure rates of electronic components and systems all along the value chain. Although is a vast project that spans in many fronts, concerning the reliability of IntegratedCircuits they propose AI methods to classify IC SEM images to detect manufacturing defects on them.

3. Solutions to mitigate caveats within cyber defense capabilities

As other electronics, defense electronics rely on Integrated Circuits (IC) of various complexity, including System On Chips (SoC) that combine semi-conductors with multiple functions (memory, logic, MOS micro-components, analog…) on a single IC.

The design of an IC is a highly complex task requiring highly specialized staff, and is subject to short time to market window and cost restriction on the final product. This has led to a horizontal design and a global manufacturing process, which involves several, potentially untrusted, stakeholders worldwide and Third Party Intellectual Property (3PIP). The economic advantages of such design and manufacturing processes thus comes with an increased security risk. The European Industry as a whole retains the capacity to perform the design step thereby limiting the ability of an adversary nation state to tamper with IC design.

From the literature study performed by the authors, only a few academic papers present some elements relative to the discovery of Hardware Trojans within the analog or digital part of an IC (*X. Cao, et al.,2015, Y. Liu, et al., 2017, T. Inoue, et al., 2017, H. Salmani and M. M. Tehranipoor, 2016,* [1] *R. S. Chakraborty, et al., 2008*).

Modern circuits typically integrate several building blocks in the form of 3PIPs. Many proposed solutions tackling the HT issue are therefore focusing on detecting HT insertion during the design phase through unreliable3PIPs. HT implemented as post Tape Out modifications are starting to be discussed and are allegedly the most difficult to detect. To date, the IC design can be done by trusted partners in Europe and it is possible to secure the supply chain all the way totape-out.

But, at some point, at least for the advanced nodes, the masks and/or the actual manufacturing of the chip will be done outside of Europe.

On the manufacturing side however, also called post-TapeOut, the situation is dire. 80% of advanced chips are manufactured in Asia and some critical post-TapeOut steps (Dicing, Packaging), are mostly provided by non EU suppliers. Photomask fabrication is also a very sensitive post-TapeOut operation, that today is mainly provided outside EU. Even EU chip manufacturers could outsource it to foreign sub-contractors.

The testing and qualification of analog modules is more critical and requires dedicated test equipment and methodologies. For instance, RF chips typically need equipment such as high-frequency signal generators, signal analyzers, oscilloscopesetc. During chip qualification, the analog

modules are thoroughly tested with various stimuli. There are typically testing infrastructure embedded in the IC that allows bypassing, probing and stimuli application on areas of the chip that are not available during normal operation. This can, for instance, allow the verification of each segment in a signal chain individually. Sometimes part of this infrastructure is also used for production test and chip calibration. Areas that are considered to be extra sensitive can require a detailed analysis, but this typically only applies on few clearly identified areas.

Reverse Engineering (RE), which corresponds to the deconstruction and imaging of each layer of a manufactured IC for comparison with the initial design, is the only method with the potential to detect advanced HTs inserted during post Tape Out operations. Unfortunately, current RE methodologies can require over a full year of work. They are thus not compatible with the detection of HTs before the defense product harboring the infected SoC is deployed in operations.

The most time consuming parts of current HT detection methods by physical inspection are the delayering (figure 1) and image acquisition process.

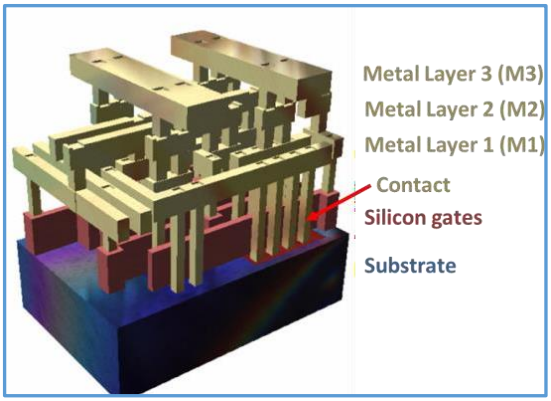


Fig. 1. Illustration of the different layers of a Chip, adapted from Hepp et al., 2022

The delayering process is still done almost in an artisanal way, using process and recipes adjusted to each chip technology (Torrance, *et al.*,2011; Huei Hao Yap and Zhi Jie Lau, 2019; H. B. Kor*, et al, 2020*).

IC with 3 metal layers (insulator removed): The sand-colored structures are metal interconnect. Layers are connected using vias. Contact connect Metal layer 1 to Substrate. The reddish structures are silicon gates. The solid at the bottom is the substrate.

The common and established delayering methods may include:

- ✓ Wet chemical etch: where liquid reactants are in contact with the IC chip creating a reaction with the surface materials. It is very important to precisely know the etching selectivity, i.e. the etching rate differences between construction materials. This process must be done in a standard chemical cabinet with fume extraction.
- ✓ Mechanical polishing: where sample surface material is removed using abrasive discs or cloths with abrasive slurry. Precision planar polishers can be used to perform this task, as well as precise CNC milling machines for local areas. Due to the narrow height of each construction layer, below 100nm for modern technology nodes, it is almost impossible to get a single layer exposed in the whole chip area using only mechanical polishing means.
- ✓ Dry Plasma Etch: Reactive Ion Etchers are specialized semiconductor fabrication equipment that can etch materials in the sample surface using a combination of physical and chemical means. Inert gas ions are accelerated to the sample surface to remove materials by physical sputter, aided by a plasma of gases that selectively reacts with the surface substances.

Regarding the image acquisition, due to the transparent nature of most of the chip's layers, and the narrow dimensions of its components, the visible light imaging is almost discarded. Scanning Electron Microscopy (SEM) can overcome this problem, offering images of sample surface topology up to few nanometers of resolution. To map an entire chip layer, several dozen thousand images are typically needed, and must be acquired in a precise, automated way. High end SEMs or Electron-Beam Lithographers (EBL) are used to perform this task. Given the SEM imaging time, requiring several seconds for each image, and the number of images needed to cover an entire chip layer, HT detection methodologies imaging the whole surface of the layers of the chip are likely to take several months or even exceeding one full year of work before providing any answer. This is particularly true for bigger chips such as the 25mm^2 chip that can be used in targeted applications. Reducing the area to image at high resolution is thus key to keep imaging time under control.

The schematic RE methodologies principle is illustrated in figure no. 2.

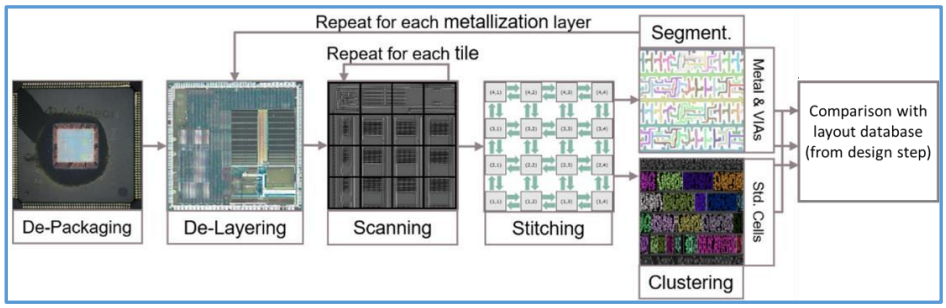


Fig. 2. Schematic illustrating RE methodologies principle. Adapted from Ludwig M, et al., 2021

A scheme for a possible process of identifying the hardware Trojans is presented below in fig. no. 3.

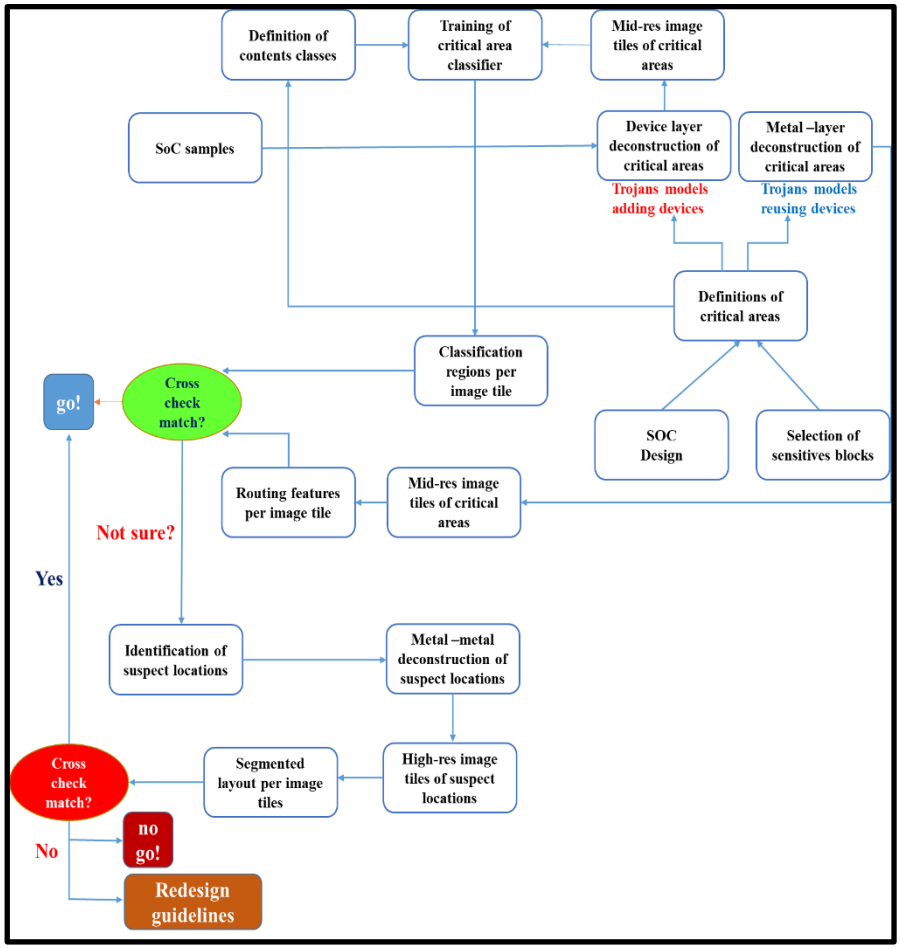


Fig. 3. A possible process of identifying the hardware Trojans

4. Conclusion

Under the Common Foreign and Security Policy (CFSP), and in particular in the context of the Capability Development Plan (CDP), the EU has identified 11 new priorities for capability development. One of these new priorities concerns cross-cutting capabilities that contribute to the EU's ambition. Under this priority, EDA aims to develop the autonomous EU capacity to test and qualify EU-developed defence capabilities prior to deployment in operations and missions, such as:

- *Integration of military capabilities - air security;*
- *Air superiority;*
- *Ground combat capabilities;*
- *Underwater control to contribute to maritime resilience;*
- *Naval maneuverability;*
- *Space-based information and communications services;*
- *Cyber-reactive operations;*
- *Air mobility;*
- *Information superiority;*
- *Improved logistical and medical support capabilities.*

As information technologies and computer systems are ubiquitous in defence products, SoCs are likely to form the basis for the operation of much of the solutions being developed to acquire new defence capabilities. Since Europe does not have control over the entire electronic component supply chain, it is strategically important for the EU to be able to test and qualify as Trojan-free the ICs that are part of solutions being developed to acquire the eleven capabilities defined in the CDP. Failure to do so would expose the newly developed capabilities to the threat of HTs. The consequences of the presence of HT in defence products are far-reaching: HTs can compromise cryptographic functions, leading to a weakening of the secrecy of communications, they can affect the sensitivity of various types of sensors (IR, radar...) or even enable unwanted remote control to switch some devices.

The implementation of large scale HT detection services, will likely come with the discovery of HT in chips destined to Defence products. This will raise the awareness of Defence Product manufacturer and foster cooperation between the latterand their design houses to identify and decrease risks. As a concrete example of potential cooperation between these entities, we could envision the identification of Defence product's function particularly targeted or vulnerable to the insertionof HT. This could lead to an improvement in the Defence product design by limiting or improving these functions, but also in the chip design by focusing design efforts to limit HT insertion risks in EU Defence.

Acknowledgments

This paper was published as part of the project "Center of Excellence for Cyber Security and Critical Infrastructure Resilience (SafePIC)", Contract No. 270 / 23.06.2020, ID 120436, funded under the Operational Program Competitiveness 2014-2020, Priority Axis: 1. Research, Technological Development and Innovation (RDI) to support economic competitiveness and business development.

References

[1]. Hepp *et al.*, A Pragmatic Methodology for Blind Hardware Trojan Insertion in Finalized Layouts. Arxiv. Aug. 2022. Accessed on 22.02.2023 at the address https://arxiv.org/abs/2208.09235.

[2]. Ludwig M, *et al.*, ViTaL: Verifying Trojan-Free Physical Layouts through Hardware Reverse Engineering. 2021. IEEE. Accessed on 23.02.2023 at the address https://ieeexplore.ieee.org/document/9707702.

[3]. X. Cao, *et al.*, "A hardware Trojan embedded in the Inverse Widlar reference generator," *2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2015, pp. 1-4, doi: 10.1109/MWSCAS.2015.7282131.

[4]. Y. Liu, *et al.*, "Silicon Demonstration of Hardware Trojan Design and Detection in Wireless Cryptographic ICs," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 4, pp. 1506-1519, April 2017, doi: 10.1109/TVLSI.2016.2633348.

[5]. T. Inoue, *et al.*, "Designing hardware trojans and their detection based on a SVM-based approach," *2017 IEEE 12th International Conference on ASIC (ASICON)*, 2017, pp. 811-814, doi: 10.1109/ASICON.2017.8252600.

[6]. H. Salmani and M. M. Tehranipoor, "Vulnerability Analysis of a Circuit Layout to Hardware Trojan Insertion," in *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1214-1225, June 2016, doi: 10.1109/TIFS.2016.2520910.

[7]. R. S. Chakraborty, *et al.*, "On-demand transparency for improving hardware Trojan detectability," *2008 IEEE International Workshop on Hardware-Oriented Security and Trust*, 2008, pp. 48-50, doi: 10.1109/HST.2008.4559048.

[8]. Zachariasen M., Fixed Orientation Interconnection Problems: Theory, Algorithms and Applications DOCTORAL DISSERTATION. Department of Computer Science (DIKU) Faculty of Science University of Copenhagen.

[9]. Torrance, *et al.*, “The state-of-the-art in semiconductor reverse engineering” 2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC), pp. 333-338.

[10]. Huei Hao Yap and Zhi Jie Lau, “Delayering Techniques: Dry/Wet Etch Deprocessing and Mechanical Top-Down Polishing”, Microelectronics Failure Analysis: Desk Reference, 7th ed., Edited By Tejinder Gandhi, ASM International, 2019, p 379–390.

[11]. H. B. Kor*, et al.*, "Sample Preparation for Deprocessing of 3D Multi-Die Stacked Package," 2020 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2020, pp. 1-6.

Carnival of Cybercrimes - Taking off the Mask of Synthetic Identity Theft

Larisa-Mădălina MUNTEANU
Data Protection Lawyer and Deputy Data Protection Officer, JS Information Governance Ltd, Peterborough, the United Kingdom
larisa@js-ig.com

Abstract

This article portrays a comparative and doctrinal analysis that aims to combine theoretical and applicable knowledge over a deeply rooted, yet still unfamiliar cybercrime: synthetic identity theft. The jurisdictional dimensions explore the European Union (EU), United Kingdom (UK) and United States (US) in terms of expertise, legal initiatives, regulations and practical cases. As a prerequisite, the study has addressed the connection with identity theft and identity fraud as the Criminal Law "labels" it generally belongs to. Moreover, the most thought-provoking part represents analysing the nexus between synthetic identity theft and personal data protection, focused on security incidents. On this latter point, personal data breaches are proven as frequently being both a cause and an effect for synthetic identity theft. Subsequently, this turns out to have significant impact on individuals and organisations alike, predominantly in the financial sector, although harm may take several shapes.

Index terms: cyberattacks, financial fraud, personal data, regulatory framework, synthetic identity theft

1. Introduction

In modern times, digitalisation has become the foundation of our daily activities, impacting on all sectors – health, communications, education, economy, justice etc. However, life has taught us there is always "the flip side" of the coin. So, what is the less advantageous part of this evolutionary road? It could be our exacerbated dependence on technology, the psychological side-effects, or perhaps the privacy risks we expose to, although inadvertently. With respect to the latter, by far, one of the most impactful ones refers to cybercrimes, concomitantly and proportionally escalating as new technologies emerge. For example, the COVID-19 pandemic has resulted in a notable increase in online transactions and processing of health data, which is considered a special category of personal data, according to the European Union and the United Kingdom's General Data Protection Regulation [1, Art. 9], [2, Art. 9]. This means cyberattackers have now reconsidered the focus of their activities – there may have even been cases where such data became more valuable than financial data. Thus, different social and technological changes have shaped the cyberspace and subsequently, the interests of the online-focused perpetrators. To confirm this, the Police Executive Research Forum has previously accentuated this advancement stating "as technology becomes more sophisticated, so have computer crime schemes" [3].

Furthermore, by overlapping the increase in financial fraud and impersonation, we can visualise synthetic identity theft. It is not a notion that emerged because of the pandemic, but has definitely flourished throughout this period. As interesting, yet challenging this cybercrime may be, this article

will explore synthetic identity theft in the following dimensions: theoretical and practical aspects, regulatory standpoint and privacy impact.

2. Conceptualisation

2.1. Theoretical aspects

ENISA [4] highlighted in the last report on cyber threats that identity theft has been increasing in frequency, mostly because of the convenience and accessibility brought by dark web and related forums to perpetrators interested in personal data. Bracker et al. [5] confirmed this increase, especially in the context of fraudulent activity during the COVID-19 pandemic. Moreover, identity fraud, concept used alternatively with identity theft and the act of impersonation, is one of the cybercrimes with significant impact on personal data: not only that it results in unlawful advantages, but it also leads to data breaches [6], which are detailed in Section 4.

At the same time, the context in which identity theft is identified, be it physical or digital, follows identical rules and steps, and thus, Bandler and Merzon [7] concluded that identity theft is usually "the gateway to cyber schemes". To continue, synthetic identity theft was considered this year "the fastest-growing type of identity theft" [8]. To confirm, ENISA's report from 2022 [4] listed the most frequent forms of identity theft as being credit card fraud and synthetic identity theft. At the same time, other studies have indicated a tight connection between these last two – synthetic identity fraud may additionally be used "to apply for credit cards or loans, as well as to bolster and improve additional fake customers' creditworthiness" [9]. Although that is the most usual hypothesis, the list of motives is open. For example, the founder of a US company providing security and privacy support to healthcare organisations emphasised the importance of medical data in this context and how the healthcare system can become a target of these cybercrimes too, in addition to the banking one [10].

Nowadays, the Federal Reserve has officially defined this phenomenon as "the use of a combination of personally identifiable information (PII) to fabricate a person or entity in order to commit a dishonest act for personal or financial gain". On these grounds, synthetic identity theft has crossed the borders of financial crimes and extended the scope towards a plethora of crimes [11]. Nonetheless, synthetic identities do not rely upon such consistent trails as real identities, and will most likely fail substantial verifications, leading to a resourceful approach that can be used for determining the authenticity of a person to the highest degree – "by evaluating the depth and consistency of information available about applicants in third party data systems" [12].

A study from the early 2000s transpired this phenomenon in a simple definition. The first important mention is that the author includes synthetic identity theft as part of account fraud, taking two possible routes: purely fabricated information or pairing real social security numbers with fictious names [13, p. 101]. Thus, this would result in unusual conclusion – identity theft can occur without stealing anyone's identity, in this case!? In reality, this may be the explanation as for why some studies refer solely to the second form mentioned above. However, it is clear that this differentiates synthetic identity theft from impersonation.

All in all, regardless of the specific definition, synthetic identity theft lies on an elaborated strategic foundation, due to the necessity to "slowly and methodically create an artificial, or synthetic person", context in which it becomes worrying that only a few people apprehend this threat [14]. This inherent complexity has led the Federal Reserve Systems of the United States to reach the drastic conclusion that "no single organization can stop synthetic identity fraud on its own" [15]. Furthermore, this conclusion was also stated in [12], accentuating "there has been no efficient way of uncovering synthetic ID fraud", yet such applications are generally not accepted due to the impossibility to match the name with the records of financial institutions.

2.2. Practical impact

Nonetheless, the practicalities of cybercrimes that are identity-related become of interest to modern researchers because the perpetrators have, nowadays, the convenience and assistance of dark web for obtaining digital goods [6]. However, it is interesting to notice that some authors concluded that a significant part of the unlawfully collected personal data shared on dark web "is never used for anything at all" [16, p.54].

On the other hand, in terms of side effects and potential harm, the United Nations Office on Drugs and Crime has pointed out in 2011 that the simple fact that an existing person is not in reality affected by synthetic identity theft should not be understood as being a harmless incident, referencing several studies that confirm that more than half of identity theft offences rely upon synthetic identities [17]. To continue, the negative effects are mostly quantified as delaying their detection and investigation, along with creating difficulties for the victims throughout the process. To support this perspective, the potential judicial side effects are recognised by older studies as well, referring to the debt collectors that may simply track the social security number back to the real person owning it, causing "reputational harm and emotional distress, in addition to wasting the victim's time and resources" [13, p.103]. A more recent study has highlighted that such identity theft should not be concerning as the only side effect to be identified is that it "drains wealth from the broader economy" [16, pp. 30-31].

One of the major cases involving synthetic identities happened in the US, when over $3 million was fraudulently obtained by two men. They commenced this scheme in 2017 and in 2020, they additionally took advantage of the Paycheck Protection Programme, as per the Coronavirus Aid, Relief, and Economic Security Act [18], aimed to support small businesses throughout the COVID-19 period. By using the stolen identities and creating synthetic ones, financial support was initially sent to the latter and subsequently, to the perpetrators' accounts [19]. This case echoes the 2022 report of the Pandemic Response Accountability Committee, explaining how the United States decided to include more security measures for identification purposes, such as a Personal Protection Identification Number (IP PIN), in addition to SSNs, enforcing "dual factor identity validation" [20, p.11]. However, way earlier than that, in 2006, a similar case proved the emergence of synthetic identity theft in the United States, after two men paired Social Security Numbers from credit reports with fictious identities and charged $760,000 to the synthetic "persons" created [21].

Thus, the Federal Reserve has prepared the Synthetic Identity Fraud Mitigation Toolkit early last year [22], with the aim to educate people and raise awareness in order to counter the maturing of this cybercrime, pinpointing even less discussed scenarios, such as using social security numbers of children, impeding later employment or loans [23].

3. Legal initiatives

From a legal perspective, the only binding international instrument on cybercrimes is the Budapest Convention on Cybercrime [24]. However, given the open clauses of the convention, it can be considered to have limited powers, planning to create a framework for the Signatories to follow in order to create harmonised and deeply rooted legal provisions. Among the categories of cybercrimes that are addressed therein, Article 8 is of interest for this study, urging the states to adopt legislation against computer-related fraud committed for obtaining economic benefits.

To continue, EU Member States have incriminated identity theft, but with no explicit reference to the synthetic form. As an example, "identity-related crime concerns in France mainly focus on document and financial fraud", leading to the initiative to introduce "an eID card with biometrics identifiers and based on centralised databases" [25, p.18]. However, this initiative was equally confirmed at EU level and the Member States were finally in the position to be bound by a Regulation urging them to shift to secure electronic ID cards by latest 2031, as a countermeasure to identity fraud,

along with ensuring national laws sanction such schemes [26, Recital 8, Art. 5]. Comparatively, Estonian citizens were using eIDs since 2002, based on a chip and two PIN numbers – serving as authentication and digital signature. Among others, the ID would allow the person to vote and purchase transport tickets too. By doing so, it was regarded as "resilient to cloning" and it meant identity theft could occur only if the card was stolen together with the PIN codes [25, p.23].

On the other hand, in the United Kingdom, the Fraud Act 2006 and Digital Fraud Committee highlighted in their last report the importance of identity theft in the ecosystem of fraud, especially given the criminal activity of enterprises on dark web that sometimes create synthetic identities, "yet it remains out of scope of criminal offences" [27, para. 458]. As a response, the Government considered existing legislation already covers this type of fraud by applying the Fraud Act 2006 [28] and the Computer Misuse Act 1990 [29]. On top of that, it should be added that the Identity Documents Act 2010 incriminates the possession of false identity documents under certain conditions [30, s. 4-6], where "false" is defined as encompassing inaccurate or omitted information "in a tendency to mislead" [30, s.9(4)a)]. In my opinion, synthetic identity theft can be covered by this definition, if we accept that the false information added to the authentic personal information lead to an "inaccurate" identity. However, the Committee emphasised the necessity to regulate identity theft as "a specific criminal offence" and alternatively, "a serious aggravating factor in cases of fraud" [29, para. 459]. Perhaps, this is the reason why some authors considered "the UK scheme of identity crime statutes is not well thought out", representing a "hodgepodge of different statutes" [31].

Nevertheless, the United Kingdom Government has adopted a process for organisations to follow when checking individuals' identities [32]. As a result, this guidance explains the impact of disregarding synthetic identities and identity fraud and labels the confidence levels upon such verifications as low, medium, high and very high. Of interest to this article is the first category, where synthetic identities are included.

Comparatively, in the US, incriminating identity theft took a more straightforward route, by having it listed in the United States Code (U.S.C.), in Title 18 - Crimes and Criminal Procedure, section 1028, called "Fraud and related activity in connection with identification documents, authentication features, and information" [33]. However, the Identity Theft and Assumption Deterrence Act of 1998 was the first legal instrument to officially describe and prohibit identity theft as a federal crime, substantiating criminal laws [34]. It is essential to note that this prohibition extends to the "intent to commit, aid or abet" such unlawful acts constituting violations of federal law or felonies under statal or local law. With respect to aiders and abettors, the U.S.C. has addressed the social impact of "intermediaries" that assist the author, known as accomplices [35]. By doing so, the authorities prove to create a comprehensive legal framework, aimed at establishing high standards for countering identity-related criminal activity.

On the other hand, section 1030 of the U.S.C. could be equally applicable, as it covers computer-focused fraud and hacking [36], mirroring the Computer Misuse Act 1990 from the United Kingdom [30]. The nexus is represented, chiefly, by referring in a couple of the prohibited acts to the condition to "affect the interstate of foreign commerce" [36, s. (a) (6) (A), s. (a) (7)].

Moreover, when it comes to protecting personal data in the context of financial or credit fraud, it has been concluded other federal laws can be applicable as well [37]: The Gramm-Leach-Bliley Act [38], The Fair Credit Reporting Act [39], The Credit Repair Organizations Act [40], the Federal Trade Commission Act [41], The Consumer Financial Protection Act of 2010 [42]. Thus, practical cases can be subject to multiple regulations at the same time. On these grounds, synthetic identity theft does not benefit from an individual and explicit incrimination, but could be covered by existing laws. However, some authors [43] believe risks are still present due to the restrictive privacy laws that limit financial institutions "to share information about synthetic identities" and subsequently, allow the perpetrators to simply change the institution and repeat the fraudulent scheme. Thus, current

"laws and agencies that are designed to help consumers also make it easier for the perpetrators to navigate and manipulate the financial system".

4. Interconnectivity and impact on personal data protection

The intrinsic relationship between synthetic identity theft and personal data relates to its unlawful use in order to "gain a financial advantage and other benefits" [6, p. 2]. To add to that, it was highlighted that the link between identity theft and personal data breaches is almost inextricable because PII is "a prerequisite to perpetrate the crime" and subsequently, "data breaches appear to be the primary source" for obtaining it [44]. However, data breaches can also be an effect of identity theft or fraud, especially when it comes to financial or health data, as per the last European Data Protection Board's guidance on data breaches' notification [45].

To analyse this in more detail, given the "personal data breach" definition from the EU and UK General Data Protection Regulation [1, Art. 4], [2, Art. 4], almost all of the possible forms are checked by synthetic identity theft: access, disclosure, transmission, storage, and the essential one, alteration. Furthermore, it has been previously identified that the mechanism sitting behind synthetic identity theft can be summed up to matching valid (stolen) social security numbers with fictious personal data "derived from one or more individuals such as name, address, date of birth, or any other information necessary to apply for any line of credit" [46]. To continue, right after the GDPR enforcement, a study has reached the conclusion that EU data protection rules have strengthened at an opportune moment for combatting the increasing number of cyber threats that lead to data breaches [47].

The nexus between synthetic identity theft and data breaches was highlighted by the European Data Protection Board as well, classifying the former as social engineering attacks. In its response to the public consultations regarding data breach notifications, the Board has confirmed the financial interests of the criminal committing synthetic identity theft, defining its purpose as "to open fraudulent accounts and make fraudulent purchases" [48]. However, regardless of the specific scope, a quick analysis of Recital 75 of the GDPR [1][2] sheds light on the relevance of identity theft on this subtopic – a processing activity that is probable to lead to identity theft means "a risk to the rights and freedoms of natural persons". On these grounds, to exemplify, such a risk will be taken into account for audits to the organisation, for evaluating the maturity of the privacy program and even for assessing the impact of data breaches.

At the same time, as explained in [16], synthetic identity theft commences with a data breach that allows personally identifiable information (PII) to be consequently unlawfully collected and shared in dark web markets. However, this PII is rather used for synthetic identity theft, instead of more direct identity-focused offenses such as payment card fraud. This was referred to as "non-ransomware data breaches", especially as the person whose information was paired with fictious details does not generally experience any financial harm. Thus, the alarmingly dangerous aspect points to the substantial difficulty to track fraudulent uses of leaked PII after data breaches, in the context where "PII relating to nearly every American consumer is already available on the dark web from multiple breaches". In my opinion, assembling the above opinions creates, in the ecosystem of synthetic identity theft, a cyber vortex where data breaches represent both the cause and effect.

On the same note, another study has emphasized that PII collected upon data breaches is generally not used in order to cause economic harm to the data subject, as "the market for consumer PII is saturated". In this context, committing synthetic identity fraud turns was not harming specific individuals, but for true identity theft or perhaps, state surveillance [16, p. 51].

As a response, regulators have established legal obligations on entities facing data breaches. The UK [2] and EU GDPR [1] are clear on that aspect, comprehensively explaining the cases when the Supervisory Authority and the data subjects must be notified at Art. 33 and 34. However, the US is more challenging. There is no such law at federal level, but states have began addressing this since

2003, commencing with California. As a confirmation of their importance, a study from 2020 has identified "the potential criminal harm of identity theft as their main rationale for the duty to notify" [44].

5. Conclusions

All in all, synthetic identity theft is part of a larger sphere from the field of cybercrimes and it becomes a growing problem for more sectors nowadays, both private and public. It is even more concerning that the impact results in harm for both natural persons and businesses, yet action and legal measures are not effectively in place at this point. There is plenty of space for evolution and in my opinion, prevention and protection need an adequate legal background for efficient enforcement.

It is clear that fraudulent schemes, especially in the digital realms, will not cease to occur and on the contrary, they will find new methods for staying "out of sight" as long as possible, taking advantage of any legal lacunae. Therefore, I believe more attention should be given to synthetic identity theft due to its peculiarities and in particular, its apparently harmless disguise. I would also find it useful and interesting to be able to identify more studies (including empirical research) on this progressive topic.

References

[1]. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1.

[2]. Retained Regulation (EU) 2016/679 of the European Parliament and of the Council (UK GDPR).

[3]. Police Executive Research Forum, "New National Commitment required: The Changing Nature of Crime and Criminal Investigations," Washington, D.C., USA, Jan. 2018. Accessed: Mar. 15, 2023. [Online]. Available: https://centerforimprovinginvestigations .org/wp-content/uploads/2018/04/20180000-The-Changing-Nature-of-Crime-and-Criminal-Investigations-Police-Executive-Research-Forum.pdf.

[4]. ENISA, "The year in review: ENISA Threat Landscape," Greece, Nov. 2022. Accessed: Mar. 6, 2023. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022.

[5]. W. Bracker, S. Goeringer and S. Krauss, "Fraud Prevention and Privacy Law: Emerging Conflicts Between Privacy Law and Fraud Prevention," presented at the *SCTE ISBE Cable-Tec Expo,* Denver, CO, USA, Oct. 13-16, 2020.

[6]. ENISA, "Identity theft: ENISA Threat Landscape," Greece, Oct. 2020. Accessed: Mar. 3, 2023. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2020-identity-theft.

[7]. J. Bandler and A. Merzon, *Cybercrime investigations: A comprehensive resource for everyone*. Boca Raton, FL, USA: CRC Press, 2020.

[8]. AuthenticID, "2023 State of Identity Fraud," 2023. Accessed: Mar. 20, 2023. [Online]. Available: https://www.authenticid.com/2023-state-of-identity-fraud-report/.

[9]. S. Marchetti, "Rolling in the deep(fakes)," *Bank of Italy Occasional Paper,* no. *668,* Feb. 2022. Accessed: Mar. 3, 2023. [Online]. Available: doi:10.2139/ssrn.4032831.

[10]. J. Davis, "The real victim in health data breaches? Patients' medical identities," *HealthcareITNews,* Oct. 29, 2018. Accessed: Mar. 28, 2023. [Online]. Available: https://

www.healthcareitnews.com/news/real-victim-health-data-breaches-patients-medical-identities.

[11]. Federal Reserve Banks, "Defining Synthetic Identity Fraud," 2021. Accessed: Mar. 29, 2023. [Online]. Available: https://fedpaymentsimprovement.org/wp-content/uploads/synthetic-identity-fraud-definition-overview.pdf.

[12]. B. Richardson and D. Waldron, "Fighting back against synthetic identity fraud," *McKinsey on Risk,* no. 7, Jan. 2019. Accessed: Mar. 13, 2023. [Online]. Available: https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/fighting-back-against-synthetic-identity-fraud.

[13]. C.J. Hoofnagle, "Identity Theft: Making the Known Unknowns Known," *Harvard Journal of Law & Technology*, vol. 21, no. 1, pp. 97-122, Fall 2007. Accessed: Mar. 13, 2023. [Online]. Available: https://ssrn.com/abstract=969441.

[14]. D. Rebovich and J.M. Byrne, Eds., *The New Technology of Financial Crime: New Crime Commission Technology, New Victims, New Offenders, and New Strategies for Prevention and Control*. Routledge. 2023.

[15]. Federal Reserve, "Mitigating Synthetic Identity Fraud in the U.S. Payment System," 2020. Accessed: Mar. 9, 2023. [Online]. Available: https://fedpaymentsimprovement.org/wp-content/uploads/frs-synthetic-identity-payments-fraud-white-paper-july-2020.pdf.

[16]. D.W. Opderbeck, "Cybersecurity and Data Breach Harms: Theory and Reality" in Seton Hall University School of Law Legal Studies Research Paper Series, Aug. 2022, p. 54. Accessed: Mar. 29, 2023. [Online]. Available: doi.org/10.2139/ssrn.4187263.

[17]. United Nations Office on Drugs and Crime, *Handbook on Identity-related Crime*, 2011. Accessed: Apr. 9, 2023. [Online]. Available: https://www.unodc.org/documents/congress/background-information/Corruption/Handbook_on_Identity-related_Crime_ENG.pdf.

[18]. Coronavirus Aid, Relief, and Economic Security Act (CARES Act), Pub. L. No. 116-136, H.R.748. 2020 [Online]. Available: https://www.congress.gov/bill/116th-congress/house-bill/748/text.

[19]. *Two Men Who Allegedly Used Synthetic Identities, Existing Shell Companies, and Prior Fraud Experience to Exploit Covid-19 Relief Programs Charged in Miami Federal Court*, United States Attorney's Office – Southern District of Florida, Aug. 28, 2020. Accessed: Apr. 9, 2023. [Online]. Available: https://www.justice.gov/usao-sdfl/pr/two-men-who-allegedly-used-synthetic-identities-existing-shell-companies-and-prior-0.

[20]. Pandemic Response Accountability Committee, "Key Insights: Identity Fraud Reduction and Redress in Pandemic Response Programs," Jun. 2022. https://www.pandemicoversight.gov/media/file/identity-fraud-capping-report.

[21]. United States v Rose, D Ariz, Aug. 22, 2006, CR06-0787PHX.

[22]. Federal Reserve, "Synthetic Identity Fraud Mitigation Toolkit," 2022. Accessed: Apr. 2, 2023. [Online]. Available: https://fedpaymentsimprovement.org/synthetic-identity-fraud-mitigation-toolkit/.

[23]. Federal Reserve, "Protecting Your Kids from Synthetic Identity Fraud", 2022. Accessed: Apr. 2, 2023. [Online]. Available: https://fedpaymentsimprovement.org/wp-content/uploads/protect-kids-from-synthetic-identity-fraud.pdf.

[24]. Convention on Cybercrime, opened for signature 23 November 2001, ETS No 185 (entered into force 1 July 2004).

[25]. F. D. Ciccio, "Comparison of Identity Theft in Different Countries," 2014. [Online]. Available: https://courses.cs.ut.ee/MTAT.07.022/2014_fall/uploads/Main/francesco-report-f14.pdf.

[26]. Regulation (EU) 2019/1157 of the European Parliament and of the Council of 20 June 2019 on strengthening the security of identity cards of Union citizens and of residence documents issued to Union citizens and their family members exercising their right of free movement [2019] OJ L 188/67.
[27]. UK House of Lords, "Fighting Fraud: Breaking the Chain," Fraud Act 2006 and Digital Fraud Committee Report of Session 2022-23, HL Paper 87, 2022. Accessed: Mar. 27, 2023. [Online]. Available: https://publications.parliament.uk/pa/ld5803/ldselect/.
[28]. UK Fraud Act 2006 (2006, Nov. 8). Accessed: Apr. 10, 2023. [Online]. Available: https://www.legislation.gov.uk/ukpga/2006/35/contents.
[29]. UK Computer Misuse Act 1990 (1990, Jun. 29). Accessed: Apr. 10, 2023. [Online]. Available: https://www.legislation.gov.uk/ukpga/1990/18/contents.
[30]. UK Identity Documents Act 2010 (2010, Dec. 21). Accessed: Apr. 10, 2023. [Online]. Available: https://www.legislation.gov.uk/ukpga/2010/40.
[31]. S.R. Ahmed, "Identity Crime Legislation in the United States, Canada, Australia and the United Kingdom," in *Preventing Identity Crime: Identity Theft and Identity Fraud*, Brill|Nijhoff, 2020, ch. 6, pp. 252-542. https://doi.org/10.1163/9789004395978_007.
[32]. UK Cabinet Office and Government Digital Service, "Guide: How to prove and verify someone's identity," Jan. 9, 2023. Accessed: Apr. 10, 2023. [Online]. Available: https://www.gov.uk/government/publications/identity-proofing-and-verification-of-an-individual/how-to-prove-and-verify-someones-identity.
[33]. U.S. Code, Title 18, pt. I, ch. 47, section 1028.
[34]. Paul Newmann, "Identity Theft: A Growing Problem," The Bill Blackwood Law Enforcement Management Institute of Texas, Denison, TX, USA, 2018. Accessed: Apr. 3, 2023. [Online]. Available: https://shsu-ir.tdl.org/bitstream/handle/20.500.11875/2452/1767.pdf?sequence=1&isAllowed=y.
[35]. Congressional Research Service, "Accomplices, Aiding and Abetting, and the Like: An Overview of 18 U.S.C. § 2," R43769, 2020. Accessed: Apr. 3, 2023. [Online]. Available: https://sgp.fas.org/crs/misc/R43769.pdf.
[36]. U.S. Code, Title 18, pt. I, ch. 47, section 1030.
[37]. U.S. Government Accountability Office, "Identity Theft Services: Services Offer Some Benefits but Are Limited in Preventing Fraud," GAO-17-254, Mar. 2017. Accessed: Apr. 3, 2023. [Online]. Available: https://apps.dtic.mil/sti/pdfs/AD1031224.pdf.
[38]. U.S. Gramm-Leach-Bliley Act (GLBA) of 1999. 15 U.S.C, § 6801-6809, § 6821-6827.
[39]. U.S. Fair Credit Reporting Act (FCRA) of 1970. 15 U.S.C, § 1681-1681x.
[40]. U.S. Credit Repair Organizations Act (CRA) of 1996.15 U.S.C. § 1679-1679j.
[41]. U.S. Federal Trade Commission Act (FTC Act) of 1914. 15 U.S.C. § 41-58, as amended.
[42]. U.S. Consumer Financial Protection Act of 2010. 12 U.S.C. § 5301-5641.
[43]. IBM, "Synthetic identity fraud: Can I borrow your SSN?: Who else might be using your Social Security number and why?," Armonk, NY, USA, Mar. 2015. Accessed: Apr. 9, 2023. [Online]. Available: http://www.turnkeyrisk.com/images/whitepapers/Can-I-borrow-your-SSN.pdf.
[44]. F. Bisogni and H. Asghari, "More Than a Suspect: An Investigation into the Connection Between Data Breaches, Identity Theft, and Data Breach Notification Laws," *Journal of Information Policy*, vol. 10, pp. 45-82, May 2020. Accessed: Apr. 9, 2023. [Online]. Available: doi:10.5325/jinfopoli.10.2020.0045.
[45]. Guidelines 9/2022 on personal data breach notification under GDPR, Version 2.0, Adopted 28 March 2023. Accessed: Mar. 13, 2023. [Online]. Available: https://edpb.europa.eu/system/files/2023-04/edpb_guidelines_202209_personal_data_breach_notification_v2.0_en.pdf.

[46]. N. L. Piquero, A. R. Piquero, S. Gies, B. Green, A. Bobnis and E. Velasquez, "Preventing Identity Theft: Perspectives on Technological Solutions from Industry Insiders", *Victims & Offenders*, vol. 16, no. 3, pp. 444-463, 2021. Accessed: Apr. 8, 2023. [Online]. Available: doi:10.1080/15564886.2020.1826023.

[47]. European Commission, "Trends in electronic identification: An overview: Value Proposition of eIDAS eID," COM/DIGIT.D3/2017/01-035, 2018. Accessed: Apr. 18, 2023. [Online]. Available: https://ec.europa.eu/cefdigital/wiki/download/attachments/78549570/Trends%20report%20on%20electronic%20identification_for%20publication_v.1.1.pdf?version=1&modificationDate=1551198712785&api=v2.

[48]. dataTENET, Response to the public consultation "Guidelines 01/2021 on Examples regarding Data Breach Notification Adopted on January 14, 2021, Version 1.0", Mar. 1, 2021. Accessed: Apr. 18, 2023. [Online]. Available: https://edpb.europa.eu/sites/default/files/public_consultation_replies/mail_edpb_02_03_response_to_the_public_consultation_guidelines_012021_.pdf.

Countering Daesh Cognitive and Cyber Warfare with OSINT and Basic Data Mining Tools

Gianluigi ME[1], Maria Felicita MUCCI[2]
[1] Department of Economics and Finance, Luiss Guido Carli University, Rome, Italy
gme@luiss.it
[2] S & A | Sistemi & Automazione S.p.A., Rome, Italy
mfmucci@sealink.it

Abstract

Digital civilization has changed war circumstances. Emerging dangers have asymmetry, variety, and continual change; quick transmission through the network; near-immediacy; possibility for unrestricted access; and swift power to affect people's behavior. Cognitive Warfare, an international relations issue, uses information, cyber, and psychological warfare tactics. Daesh sends threatening messages to Western countries and spreads internet propaganda to recruit new members and induce terror. The study attempts to propose a novel knowledge-based approach for detecting terrorists by examining data obtained from Twitter and leading Daesh publications, through Data Mining techniques.

Index terms: clustering analysis, cognitive warfare, counterterrorism, Daesh, national security, virtual jihad

1. Introduction

International power shifts since the Cold War have pushed us closer to a multipolar world. Top-down and bottom-up processes spawn new actors and security trends. We no longer face threats that can be spatially contained in an attack by a big power against another. International terrorism and cybercrime have joined military challenges in national security discussions. Though interrelated, these new structural trends are fragmented and multidimensional [1]. Today's threats are asymmetrical, diverse, and ever-changing, spread over the network, are immediate, may be publicly available, and can easily impact behaviour. New digital technologies and social media have allowed actors to reach a wider audience with tailored and targeted content. Hostile propaganda players are aware of the cyber environment's prospects, and they are working constantly to exploit information and undermine its essential truth principles. In addition to providing more access to information, the Internet offers greater opportunities for deception. Indeed, Cognitive Warfare is becoming a crucial aspect in international politics and a growing source of concern.

The current historical era, highlighted initially by the Covid-10 Pandemic crisis and finally by the outbreak of the Russian-Ukrainian conflict, has redirected media and major international powers' focus away from the terrorist threat. Despite this apparent stalemate, Daesh keeps on operating and establishing itself using many technologies to carry out its acts and plans. It has initiated disruptive cyber and cognitive warfare efforts, including Cyber-Training, Cyber-Planning and Cyber-Execution, Funding and Fundraising, conducting Cyber Attacks, recruiting new members, and disseminating online media propaganda. Virtual Jihad is becoming an increasingly alarming issue that must be addressed.

Jihadism is a relatively new phenomenon in terms of both its objectives and means. Due to technical advancements, the communications revolution, and improvements in information storage and retrieval, previously unthinkable techniques for bringing a community to the forefront of people's thoughts have developed. The current generation of Jihadists is the first to grow up with pervasive access to digital communication tools. This makes them the first generation capable of joining and directing terrorist organizations. So, it is unsurprising that these networks are essential to the radicalization and the recruiting techniques they employ to attract vulnerable individuals. Jihadists have demonstrated their proficiency at utilizing globalization's resources to achieve their own objectives through this strategy. Social media platforms and the Internet have become a highly effective tool for spreading propaganda, instigating violence, and radicalizing a far larger audience than in the past, and - if managed properly – they may become powerful psychological weapons with disruptive effects. Indeed, we now refer to the *Weaponization of Media Narratives*: the battle of narratives has surpassed the conventional military and physical Jihad in importance.

Thus, it is essential to design and deploy Cyber Defense methods to prevent, identify, and dissuade jihadist Internet activity. In this environment, law enforcement, intelligence, and other agencies are always inventing new techniques to prevent, identify, and limit terrorist activities on the Internet. Due to their effectiveness in promptly identifying possible terrorist threats, traditional research techniques are gaining favour again. Moreover, in the last several decades, the collection and analysis of data from a broad variety of sources, in addition to *text analysis*, has been able to give intelligence analysts with useful insights by revealing previously hidden yet logically sound patterns and connections. Furthermore, Open Source Intelligence (OSINT) and Social Media Intelligence (SOCMINT) are employed to collect pertinent data useful for proving added information about threats. Then, with a fundamental comprehension of Machine Learning (ML) techniques, one may construct a model from a collection of inputs and use it to generate predictions or judgments.

This study aims to develop an innovative knowledge-based technique that may be utilized to detect jihadist affiliation by analysing data collected from social network platforms (particularly Twitter) and the leading Daesh publications. The suggested method may determine the conventional behaviour (or "profile") of militants and sympathizers because it employs a data mining approach, notably *clustering analysis*, to analyse digital content associated with Daesh.

2. Virtual Jihad

If we now refer to the *modus operandi* of Daesh, particular attention should be paid to the way online tools are used, as they have become one of the fundamental elements of terrorists' activities [2]. According to Weimann, there are two distinct categories to which Jihadists' use of the Internet can be classified: *communicative* and *instrumental* [3]. The first categorization includes spreading propaganda, running psychological warfare campaigns, and recruiting new members, while the second one includes cyber-training, cyber-planning and coordination, and digital fundraising.

Daesh has become increasingly reliant on the Internet in the last decades as a substitute for traditional training grounds. The face-to-face nature of certain activities has given way to their virtual counterparts in modern times. "Virtual training camps" [4] provide a platform for prospective recruits to learn about and support terrorist groups, while also encouraging involvement in direct acts by providing access to encryption tools and anonymizing tactics. Weimann claims that the readily available multimedia format in several languages has become a "terrorist university," a place where jihadists may learn new strategies and abilities that make their assault methodology more efficient. Because of the Internet's inherent interactivity, individuals from all over the world can feel connected to one another and form networks to share strategies and tactics. Indeed, there is a wealth of information online on how to join terrorist groups, how to plan and execute terrorist acts, and how to construct explosive devices and other weapons (Figure 1).

Fig. 1. A Tweet of a suspected Jihadist account, who refers to the "amazing" military infographics of Al-Naba' Magazine

In recent decades, nearly every terrorist attack has involved the use of the Internet and the opportunities it provides [5]. Typically, the planning of a terrorist act requires communication between multiple parties over greater distances and utilizing more sophisticated methods. Daesh's use of anonymizing communication tools to plot attacks has progressed to a new level of sophistication. Tools for encrypting data and software designed to mask a user's identity make it difficult to determine the sender, the addressee, and - most generally - the contents themselves [6].

Not only does a Jihadist organization require a variety of resources to finance its operations, but also to maintain its existence and grow as an organization. Funds are required for a variety of reasons, among all to maintain militants and their relatives; to finance travel expenses; to recruit and train new members; to acquire weapons and safe houses; to carry out operations; and to promote the ideology of the group through social activities or propaganda. So, it is crucial for the success and resilience of the groups to have access to financial resources, particularly at the beginning, when these are required to assist recruit and maintain support, as well as to create major material capabilities [7].

Moreover, one sort of disruptive cyber intelligence activity that violent radical political parties engage in is the dissemination of online propaganda information designed to attract as many people as possible, and to recruit new members. Communication geared at terrorist recruiting is typically designed with the intention of appealing to weak and marginalized subgroups within society [8]. Consequently, the efficacy of this propaganda in terms of recruitment and radicalization depends on an individual's sentiments of injustice, alienation, or guilt. Potential terrorist recruits and existing members of the terrorist organization can develop a type of virtual community through interactive engagement, which can also promote a sense of belonging and bolster a sense of community.

In the battle for the *Umma*'s affections, narrative warfare has surpassed the employment of conventional firearms and military might. The offensive information warfare focused on propaganda is a key component of Daesh's struggle. Therefore, media and Internet resources may be a powerful psychological weapon if handled appropriately. This narrative-driven, intensified kind of terrorism has emerged as Daesh's primary asymmetric weapon.

By using a wide range of publicly accessible social networking channels, terrorist and insurgent organizations today have established an even more direct and personally intimate method of message. Some examples of these platforms include Twitter and Telegram amongst others. An example of a potential Twitter account affiliated with Daesh, but already suspended, can be seen in Figure 2.

Fig. 2. A Twitter account with Daesh affiliation (already suspended)

According to Wilkinson, "When one says *terrorism* in a democratic society, one also says *media*. For terrorism by its very nature is a psychological weapon which depends upon communicating a threat to the wider society. This is why terrorism, and the media enjoy a symbiotic relationship" [9]. Distinctly poignant is the analysis conducted by Adam Chuipka for the University of Ottawa: "Media coverage and terrorism are soul mates - virtually inseparable. They feed off each other. They together create a dance of death - the one for political or ideological motives, the other for commercial success" [10].

Daesh is at the frontline of a new revolution in jihadist communication because of the remarkable efficacy with which it uses these platforms to address a worldwide audience. The group will continue to use mass media to get publicity and support because of the glamour of cutting-edge tech and the incredible speed and reach of person-to-person power enabled by modern technology. Given the capabilities and products that are certain to become more advanced in terms of quality, content, speed, affection, and transmission capacity, as well as more numerous and pervasive than ever before, it is possible that we are only now beginning to comprehend the implications of this phenomenon.

3. Weaponization of Media Narratives

Jihadist groups recognize the importance of online media platforms in their strategic planning, as seen by their own attempts to embrace the media requirement. The war of narratives has taken a more prominent role than the traditional use of guns and fire weapons in the conflict that is taking place for control of the hearts and minds of the *umma*. Members of Daesh ideological group typically identify themselves primarily in contrast to members of other groups, making a significant separation between them and whom the others are (*Us* vs *them*). This dualism is so important to the process of forming the Islamist identity. A group's internal cohesion and belief in its own mission may benefit from this, and the group's success might inspire others to show support for or even join the group. On the other hand, the same propaganda may instil panic in the minds of people who have been singled out for terrorist attacks.

Da'wah (Arabic: دعوة) is an essential part of the Daesh's ideology. Its purpose is to recruit, indoctrinate, and motivate terrorist attackers, supporters, and sympathizers. Active *Da'wah* is carried out in online social networks as well as on the websites of the various Islamist and Jihadi organizations, with the aim of accomplishing four important goals: informing, frightening, uniting, and supporting [11]. The jihadists' timely and impressively well-informed posts in social networks

and blogs, as well as comments on Islamism websites, reveal their knowledge with daily political news, and political decision-making processes made my Western countries. This approach makes it clear that a well-articulated media strategy is in no way haphazard, but rather makes use of skills taken from the realm of communication that is globally pervasive and widespread. Terrorists may benefit in a number of ways when they use the Internet to spread their messages: they can remain hidden from public scrutiny, they can quickly and easily reach a wide audience, and they can encourage dialogue amongst their followers.

Twitter is one of the social networks that might be leveraged to accomplish this goal. Supporters of terrorist organizations may send and receive messages, photographs, videos, and website connections to a broad audience via Twitter. The platform acts as a venue for both passive and active supporters. It is a possible danger due to its ability to distribute instant messages to a huge number of users simultaneously and because it allows users to follow specified subjects and groups, as well as other users' tweets about those topics. Indeed, Daesh uses Twitter as an *umbrella platform*, which combines the numerous information sources into a unified index (mostly via the use of hashtags) that can be viewed and searched with relative ease. The weekly magazine *Al-Naba'* (which started its activity on 17 October 2015 to now) is an essential Daesh propaganda publication that we may consider in our discourse. While all *Dabiq* and *Rumiyah* magazines are published in English, *Al-Naba'* issues are only published in Arabic. The Magazine is published on a regular basis by the *Diwan al-Ilam al-Markazi*, which is the central media organization for Daesh, and it is responsible for coordinating media efforts and providing guidance to its media supporters. *Al-Naba'* features a variety of articles and material types, such as news, commentary, visualizations, religious writings (including fatwas), and advertisements for various forms of media output. So, it reflects Daesh mindset and mirror events occurring on the ground, and social issues particularly relevant to the supporters.

4. Countering Terrorist through OSINT, SOCMINT, and Data Mining Tools

It is not enough for governments to only explore new ways to acquire and combine Intelligence; they must also analyse it to provide actionable information in the fight against Daesh. Open-Source Intelligence (OSINT) is an important discipline for the processing of publicly accessible sources. It has become an increasingly valuable source of intelligence for governments, companies, and criminals alike, as more and more detailed personal information is processed digitally and accessible online. Further, a variety of searches, including picture searches, web text investigations, social media content searches, and map searches, are all accessible [12]. Social Media Intelligence (SOCMINT), a subset of OSINT, is the study of how social networking services can be monitored and mined for useful information and community detection and analysis.

The collection and analysis of information from wide variety of sources can provide intelligence analysts with relevant insights as they can reveal previously hidden yet logically sounds patterns and linkages. OSINT and SOCMINT may be interchangeably used to collect relevant data valuable for providing further information about specific threats. A basic understanding of Machine Learning (ML) algorithms can make a mathematical model from a set of inputs and utilize it to make predictions or judgements.

4.1. Aim of our study

Our research aimed to determine whether Machine Learning (ML) and Natural Language Processing (NLP) can be used to analyse Jihadist stories in order to find any similarities between different sources of propaganda. One of the specific goals was to evaluate whether or not there are tweets with a direct connection to *Al-Naba'* magazine. The volume of propaganda released by Daesh is so large that it is nearly impossible for humans to analyse it. As a result, establishing methods and procedures that can be utilized to analyse massive amounts of data is a crucial challenge. The

development of counter-narratives and counter-messaging strategies requires an understanding of the various taxonomies used in propaganda, the ways in which the narrative varies across media outlets, and the way in which it develops over users. *Social Network Analysis* and *Data Mining* – specifically *clustering algorithms* – where applied in order to achieve our main objectives. Finally, Tweets collected from potential Daesh supporters and sympathizers were used as a target of our data investigation.

4.2. Research Objectives

Examining Daesh propaganda narrative and the most common taxonomies in *Al-Naba'* magazine.

Empirical search for a tiny similarity algorithm.

Identifying Deash affiliations to *Al-Naba'* through social media analysis of Twitter using Data Mining Techniques, in particular clustering.

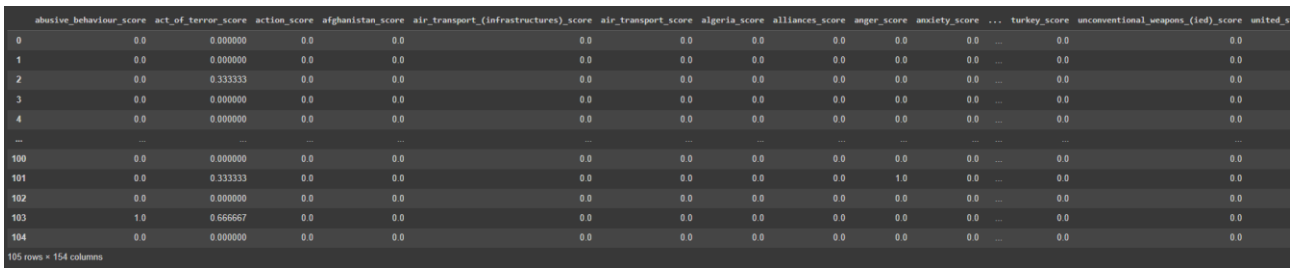
4.3. Justification

The goal of this study was to develop a model that would aid in the detection of *Al-Naba'* Daesh affiliation via the use of clustering algorithms. As a result, the proposed model would offer a framework for identifying jihadist hidden propaganda with little human intervention. This attempts to improve security organs by providing a more precise and quick way of detection than previous techniques.

4.4. Research Phases

Our empirical research started from the collection of the most relevant issues of *Al-Naba'* newspaper, published on a regular basis by the *Diwan al-Ilam al-Markazi*. *Al-Naba'* features a variety of articles and material types, such as news, commentary, visualizations, religious writings (including fatwas), and advertisements for various forms of media output. So, it totally reflects Daesh mindset and mirror events occurring on the ground, and social issues particularly relevant to the sympathizers. The articles were selected for their relevance as being particularly under the attention of the Western world's monitoring and inspection instruments. The reason we analyzed *Al-Naba'* magazine rather than another newspaper is because it has only been published in Arabic, whereas other journals have been released in other languages and have been the subject of numerous prior studies. However, no study has been conducted on *Al-Naba'*. Here is the first difficult encountered: a literal translation was adopted in order to diverge as little as possible from the expressive tactics and features used by Daesh adherents. A previous careful study of *Dabiq* and *Rumiyah* magazines (published in English) offered key insights for the overall translation. Several articles of each issue were then selected randomly, and their translation was refined.

The subsequent phase entailed the identification of taxonomies to be extracted for our investigations that appeared frequently. The statements were incorporated into Expert.AI's Cogito Intelligence API (an online semantic-based system dedicated to crime and intelligence that also includes specialized classification) which offers full semantic processing patterns like *Categorization*, *Text Mining* and *Fact Mining*. *Al-Naba'* extracts were selected and sequentially placed in the section to be analysed in order to determine the most pertinent metadata. Consequently, we initiated the creation of our dataset in a shared Excel file. Its line-column chart is a combination of a line graph (containing *Al-Naba'* texts) and a set of columns that correspond to the extracted Cogito system parameters, with the fewest possible empty cells. Using a JSON score, we formed a dataset in which each text was defined by a vector (a list) of scores from all categories present in all documents. The dataset was balanced by converting the scores (through such that instead of having scores ranging from 0 to infinity, they now have values ranging from 0 to 1 (Figure 3). In order to do so, we used sklearn.preprocessing.MinMaxScaler.

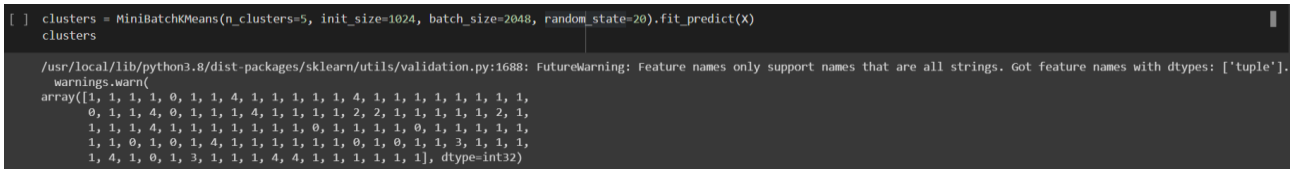


	abusive_behaviour_score	act_of_terror_score	action_score	afghanistan_score	air_transport_(infrastructures)_score	air_transport_score	algeria_score	alliances_score	anger_score	anxiety_score	...	turkey_score	unconventional_weapons_(ied)_score	united_s
0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
1	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
2	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
3	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
4	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
100	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
101	0.0	0.333333	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	...	0.0	0.0	
102	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
103	1.0	0.666667	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
104	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	

105 rows × 154 columns

Fig. 3. Dataset balanced with categories score 0-1

We fed this dataset to the clustering algorithm by choosing a number of five clusters (Figure 4).



```
[ ] clusters = MiniBatchKMeans(n_clusters=5, init_size=1024, batch_size=2048, random_state=20).fit_predict(X)
    clusters

    /usr/local/lib/python3.8/dist-packages/sklearn/utils/validation.py:1688: FutureWarning: Feature names only support names that are all strings. Got feature names with dtypes: ['tuple'].
      warnings.warn(
    array([1, 1, 1, 1, 0, 1, 1, 4, 1, 1, 1, 1, 1, 4, 1, 1, 1, 1, 1, 1, 1,
           0, 1, 1, 4, 0, 1, 1, 1, 4, 1, 1, 1, 1, 2, 2, 1, 1, 1, 1, 1, 2, 1,
           1, 1, 1, 4, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1,
           1, 1, 0, 1, 0, 1, 4, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 3, 1, 1, 1,
           1, 4, 1, 0, 1, 3, 1, 1, 1, 4, 4, 1, 1, 1, 1, 1, 1], dtype=int32)
```

Fig. 4. Number of clusters identified

Therefore, a dataset was generated with the column at the end. We tried to discover the correlations between the clusters given to each text and the categories activated. At the end we noticed two clusters concentrated on the category *explosion/explosives* and *terrorism/act of terror*.

Once we had trained the clustering model based on n text categories, we investigated whether feeding Tweets were meaningfully associated with the categories. Several OSINT methods were used to extract tweets from users that may have a particular relation with Daesh and the Jihadist propaganda. First, we used Twitter Advanced Search with the following query:

(الدولة OR الإسلامية OR ISIS OR داعش OR الدولة الإسلامية OR عالمية جهاد حركه OR صحيفة النبأ) lang:ar until:2022-11-09 since:2022-01-01

Among the search results, one account in particular (@a_o_be_dh90) caught our curiosity, but it had been suspended within a few days (Figure 1).

Our initial investigation included not just the content of published posts, but also followers and followings. In terms of saving time and do more accurate research, we decided to employ another OSINT tool, namely TweetTopic. Such instrument offers a function that we have found useful in our investigation. Once the Twitter identity is provided, the tool gathers the most recent 3,000 Tweets and generates a word cloud. This detects the most frequently used terms inside the target's postings. When someone clicks on a phrase among the search results, only Tweets containing the specified terms are displayed. When you click on any of the text circles, the corresponding postings are immediately shown. It was incredibly useful when we had too many posts to read in a short amount of time. TweetTopic enables us to swiftly determine the content of our target's tweets and to promptly investigate any relevant subjects. In addition, the word cloud displays the most frequent retweets or ID accounts associated with the evaluated tweet. As a result, a set of 100 Twitter accounts (ten of them are shown in **Error! Reference source not found.**Table 1 and potentially related to Daesh – 93 o f them created from February 2022 to 8 November 2022 – was formed. Some of their tweets contain mentions to articles taken from *Al-Naba'* magazine (Figure 3).

Table 1. Suspected Daesh Accounts on Twitter

ACCOUNT	LINK	REGISTRATION TIME
Account-name_1	twitter-url-account-name_1	September 2022
Account-name_2	twitter-url-account-name_2	September 2022
Account-name_3	twitter-url-account-name_3	September 2021

ACCOUNT	LINK	REGISTRATION TIME
Account-name_4	twitter-url-account-name_4	February 2018
Account-name_5	twitter-url-account-name_5	September 2012
Account-name_6	twitter-url-account-name_6	October 2022
Account-name_7	twitter-url-account-name_7	September 2022
Account-name_8	twitter-url-account-name_8	November 2022
Account-name_9	twitter-url-account-name_9	July 2022
Account-name_10	twitter-url-account-name_10	October 2022

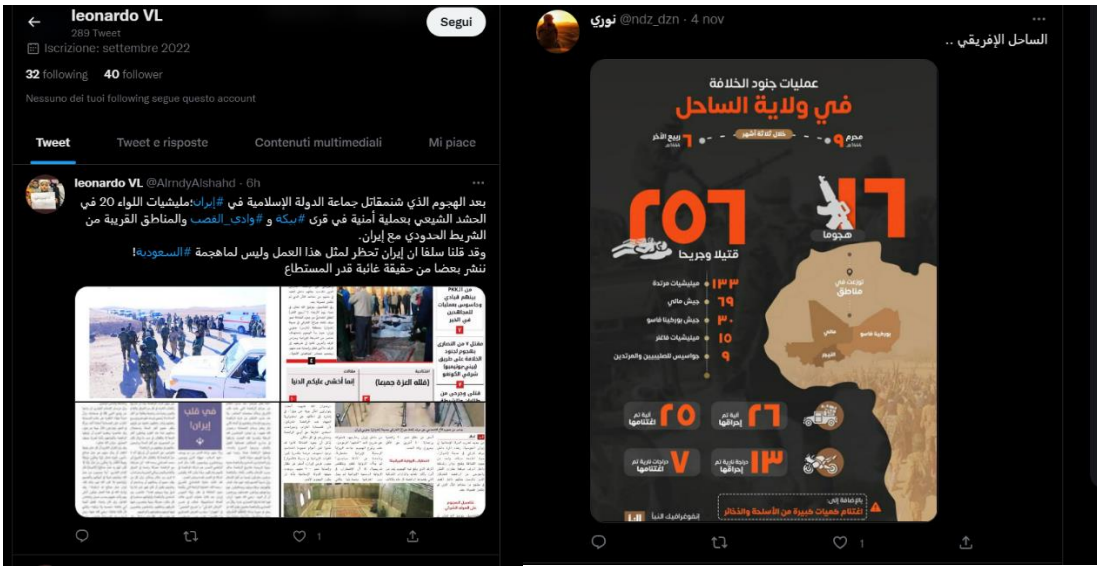


Fig. 5. Tweets, on Al-Naba' Magazine, of suspected Daesh-related accounts

Assuming that the tweets found could be classified into the identified clusters, we attempted to code them in the same manner as previously applied with the existing documents to see whether they were classified in the cluster in which we had expected them to be categorized.

4.5. Research Scope and Limitations

Obviously, our research has limits, there are still many unknown concerns. Neither the work nor the analysis is adequate to perform a complete inquiry. Notwithstanding these drawbacks, we sought to illustrate the ability of complex networks to reveal hidden patterns within the global context of terrorism to stimulate the use of fresh analytical frameworks to the study of Daesh media narratives. By finding hidden patterns across several social and digital channels, similarities in propaganda can be revealed. According to our assessment, there is an urgent need to develop and evaluate innovative methodological techniques that, if effective, can be applied to other contexts with more precise data, allowing for additional useful and decisive conclusions.

5. Conclusions

Daesh keeps on using several technologies to carry out its deeds and plans. Virtual Jihad is becoming more and more an issue of concern that must be addressed and it depends largely on offensive propaganda-based information warfare. This narrative-driven, heightened form of terrorism has become a major asymmetric weapon. Online terrorism, specifically, has long-lasting effects on the psyche ("Cognitive Warfare") of impacted communities and virtual users and is capable of causing significant harm.

The fight against the jihadist threat has had some success over the last two decades, but in the next years, it will need to take more into consideration the hazards that progress on the World Wide Web. As a result, Law Enforcement, Intelligence Agencies, and other organizations must constantly develop innovative unique techniques to prevent, identify, and limit terrorist activities over the Internet. Intelligence gathered and evaluated will aid in assessing threats, formulating responses, and guiding policies. In this process, the influence of new technologies and social shifts on the operation of Law Enforcement must be considered. Additionally, cultural, and digital awareness are essential for fostering State's collaboration and partnership in the fight against cyberterrorism.

Therefore, with our study, we aimed to demonstrate that relatively simple data mining tools with social network analysis features, when combined with conventional data mining techniques and practical semantic analysis of online propaganda, can serve as a useful starting point for identifying terrorist affiliations. Our approach begun with the following question: is it possible to employ Machine Learning and Natural Language Processing algorithms to assess Daesh narratives in order to identify possible similarities among different propaganda sources?

The volume of propaganda disseminated by Daesh is so big that it is nearly difficult for human capabilities to examine it. As a result, developing methods that can be applied to analyse massive amounts of data was a key task. Our elaboration of counter-narratives and counter-messaging approaches required a basic knowledge of Arabic language, an understanding of the several taxonomies used in propaganda, a consideration of ways in which the narrative varies among media channels and the way it develops over users. *Social Media Analysis* and *Data Mining* tools – specifically clustering algorithms – where applied to achieve our main objectives. Finally, Tweets collected from potential Daesh supporters and sympathizers were used as a target of our data investigation.

Given the employed methodology, there is no reason to doubt that – despite the potential drawbacks – there are additional potential avenues for future research.

References

[1]. N. Tocci and R. Alcaro, "Three scenarios for the future of the transatlantic relationship," *TRANSWORLD. The Transatlantic Relationship and the Future Global Conference*, Sept. 2012. [Online]. Available: http://transworld.iai.it/wp-content/uploads/2012/10/TW_WP_04.pdf. Accessed: Nov. 3, 2022.

[2]. M. Ingelevič-Citak and Z. Przyszlak, "Jihadist, Far-Right and Far-Left Terrorism in Cyberspace-Same Threats and Same Countermeasures?" *International Comparative Jurisprudence 6.2*, vol. 8, no. 2, pp. 158-159, June 2020. [Online]. Available: https://repository.mruni.eu/bitstream/handle/007/17195/6291-15119-1-SM.pdf?sequence=1&isAllowed=y. Accessed: Sept. 28, 2022.

[3]. G. Weimann, *Terrorism in Cyberspace: the next generation*, Washington, DC: Woodrow Wilson Center Press, 2015.

[4]. G. Weimann, "Virtual Training Camps: Terrorists' Use of the Internet," in James J. F. Forest, ed., *Teaching Terror: Strategic and Tactical Learning in the Terrorist World,* Lanham, MD: Rowman & Littlefield, 2006, p. 112.

[5]. B. Todorovic and D. Trifunovic, "Prevention of (Ab-) Use of the Internet for Terrorist Plotting and Related Purposes," *International Centre for Counter-Terrorism (ICCT)*, 2020. [Online]. Available: https://www.icct.nl/sites/default/files/2023-01/Chapter-19-Handbook.pdf. Accessed: Oct. 24, 2022.

[6]. D. Trifunović, "Digital steganography in terrorist networks," In *Proc. SYM-OP-IS 2015: XLII International Symposium on Operations Research*, Vol. V (1), 2015, pp. 190-193. [Online]. Available: https://www.researchgate.net/profile/Snezana-Kirin/post/HelloI_

hope_that_your_project_is_developing_well_Can_we_get_some_update_on_published_results/attachment/59d64a2479197b80779a47c9/AS:474077556154368@1490040305737/download/ZbornikN20015.pdf#page=208.

[7]. J. Adams, "The financing of terror: Behind the PLO, Ira, red brigades and M-19 stand the paymasters: how the groups that are terrorizing the world get the money to do it*,"* New York: Simon and Schuster, Jan. 1986.

[8]. R. Borum, *Psychology of terrorism*, Tampa: University of South Florida, Jan. 2004. [Online]. Available: https://www.ojp.gov/pdffiles1/nij/grants/208552.pdf. Accessed: Oct. 20, 2022.

[9]. P. Wilkinson, "Terrorism versus democracy: The liberal state response," *Cass Series on Political Violencce*, Taylor & Francis, 2011, ch. 10, pp. 152.

[10]. A. Chuipka, *The Strategies of Cyberterrorism: Is Cyberterrorism an effective means to Achieving the Goals of Terrorists?* University of Ottawa, 20 Nov. 2016. [Online]. Available: https://ruor.uottawa.ca/bitstream/10393/35695/1/CHUIPKA%2c%20Adam%2020169.pdf. Accessed: Oct. 24, 2022.

[11]. R. Zgryziewicz, J. Shaheen, T. Grzyb, and S. Fahmy, "Daesh Information Campaign And Its Influence," *NATO Strategic Communications Centre of Excellence*, 2015. [Online]. Available: https://stratcomcoe.org/pdfjs/?file=/publications/download/daesh_public_use_19-08-2016.pdf?zoom=page-fit. Accessed: Oct. 28, 2022.

[12]. F. J. Cesteros García, "2Private Investigation and Open Source INTelligence (OSINT)," in *Cybersecurity Threats with New Perspectives*, IntechOpen, Dec. 08, 2021, pp-129-143. doi: 10.5772/intechopen.95857. Accessed: Nov. 15, 2022.

ChatGPT - Information Security Overview

Gabriela TOD-RĂILEANU[1], Sabina-Daniela AXINTE[2]
[1] Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
gabriela.tod98@gmail.com
[2] Associate Professor, Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
axinte_sabina@yahoo.com

Abstract

About one hundred years ago humanity experienced a substantial change when we embraced the use of electricity in our homes and daily lives. Now, humanity is changing once again by adopting the use of artificial intelligence on a larger scale. Expressing concerns about the next industrial revolution that will fundamentally alter the way we live, work, and relate to one another. ChatGPT has become so popular in the last months that a lot of technical or not so technical people have used it and integrated in their daily work to complete tasks faster and more efficient, but this article will highlight the abuse of chatGPT by the people that do not have always good intentions - threat actors. This article is approaching the Information Security risks that have appeared with the use of chatGPT by the employees that are not aware about the threats or even the use of chatGPT by the threat actors that are aware and ready to abuse its computational power.

Index terms: chatGPT, Artificial Intelligence, information security, risk, exploitation

1. Introduction

ChatGPT is a conversation chatbot that become very popular since the end of 2022. Statistics indicates a number of 1 million users in December 2022 and the number has increased 100 times reaching 100 million users [1]. The most frequent use cases of ChatGPT are: Chatbots and virtual assistants, language translation, text summarization, content generation, code debugging and search engine. A recent study conducted by BlackBerry has indicated that "51% of IT decision makers believe there will be a successful cyberattack credited to ChatGPT within the year" [2] and this article will present several ways of exploitation of ChatGPT by Threat actors. ChatGPT, or Chat Generative Pre-Trained Transformer, is a 175 billion-parameter natural language processing (NLP) model that uses deep learning algorithms trained on vast amounts of data to generate human-like responses to user prompts [3]. The model, available free of charge on the official website [4], is trained using Reinforcement Learning from Human Feedback (RLHF) algorithm and is the 3rd generation of GPT chatbot. On March 13th, 2023, the latest version of ChatGPT, ChatGPT-4, was released [5] and it is available for a cost that can depend on the one needs and use [6].

2. Key Concepts

2.1. What is Artificial Intelligence (AI)

AI (Artificial Intelligence) is a branch of computer science that focuses on creating intelligent machines that can mimic human language and thinking. AI systems are designed to learn from their

environment and make decisions based on the data they receive. AI can be used to solve complex problems, such as medical diagnosis, autonomous vehicles, and natural language processing [7]. Nowadays, AI has become part of some people's lives and they are using this type of technology to make their daily work or chores easier.

The modern era of AI began in 1956, when a group of scientists and mathematicians gathered at Dartmouth College to discuss the possibility of creating computers that could think like humans. Since then, AI has continued to rapidly advance, with breakthroughs in machine learning, natural language processing, and robotics [7].

2.2. What is ChatGPT

According to the official website, and documentation [8] the ChatGPT model is trained using Reinforcement Learning from Human Feedback (RLHF), using the same methods as InstructGPT, but with slight differences in the data collection setup. The initial model was trained using supervised fine-tuning: human AI trainers provided conversations in which they played both sides—the user and an AI assistant. The trainers had access to model-written suggestions to help them compose their responses.

Following the initial supervised training phase, the new dialogue dataset was combined with the InstructGPT dataset, which was transformed into a dialogue format. To create a reward model for reinforcement learning, was needed to collect comparison data, which consisted of two or more model responses ranked by quality. To collect this data, the developers took conversations that AI trainers had with the chatbot. They randomly selected a model-written message, sampled several alternative completions, and had AI trainers rank them. Using these reward models, they could fine-tune the model using Proximal Policy Optimization. Several iterations of this process were performed.

3. Security Concerns

In the last months, multiple informational security experts have expressed their concerns in regards with the ChatGPT capabilities and computational power. Moreover, it was already highlighted the use of ChatGPT by the threat actors for multiple and various type of attacks as you will be read below.

A strong concern is related to the potential for ChatGPT's ability to generate human-like text. This could rapidly increase the risk of identity theft or could generate some very good phishing campaigns or BEC attacks (Business Email Compromise) by generating text free of grammatical errors that can make the user believe that were written by human beings and persuade into taking an action. The detection solutions for email threat protection are also using key words and some well-known phrases that can induce the "sense of emergency" and this detection layer could be easily bypassed using the text generated by ChatGPT.

In Fig. **1** is showed an example generated by ChatGPT as per the author request to write an email that appears to be from a courier company and has an attachment. The strategy in this case would be to deliver an attachment that will have a malicious payload.

It is also worth mentioning that there are some solutions to identify text generated by AI. For example, GPTZero [9] and OpenAI GPT2 Output Detector [10].

GPTZero is a classification model that predicts whether a document was written by a large language model, providing predictions on a sentence, paragraph, and document level. GPTZero was trained on a large, diverse corpus of human-written and AI-generated text, with a focus on English prose. [11] It is also offering an API solution that could be integrated in multiple other products and solutions. Considering this, the Email Threat Protection solutions now available will soon consider integrating this solution, GPTZero, or other similar ones in their detection process.

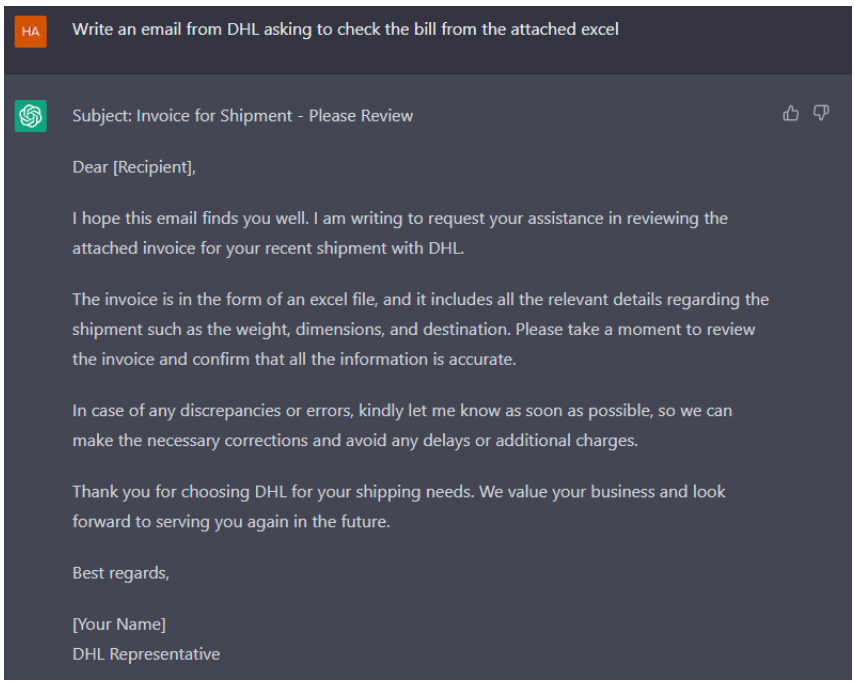


Fig. 1. ChatGPT's answer to writing an email that can be used to send a malicious attachment

Another threat that should be considered is that the main characteristic of ChatGPT is that is a chat bot and could also be used, among many other useful integrations, to bypass MFA. The phishing attack was already seen [12] and attackers used bots like SMSRanger and BloodOTPbot. They automatically follow up a credential harvesting attack with a phone call using a carefully crafted social engineering script that ends with obtaining the victim's authentication code.

One of the biggest risks is ChatGPT being used to write malware code. Based on recent research [13] [14] [15] [16] there was proved that ChatGPT can generate code that could be used by threat actors, hackers, in different attacks or tools they are developing, for example Malware as a Service (MaaS) [17]. OpenAI has continuously worked on preventing their product from providing malicious code and, when there is an explicit demand, ChatGPT is prompting a response that is highlighting that the request is potentially dangerous and may be illegal or unethical.

Notwithstanding, some researchers [13] have found the wording to receive an injection code but the adaptivity of ChatGPT and the efforts of OpenAI to reduce the exploitation of their product were successful and the exact same wording have returned just a warning from the chat bot as you can see in Fig.

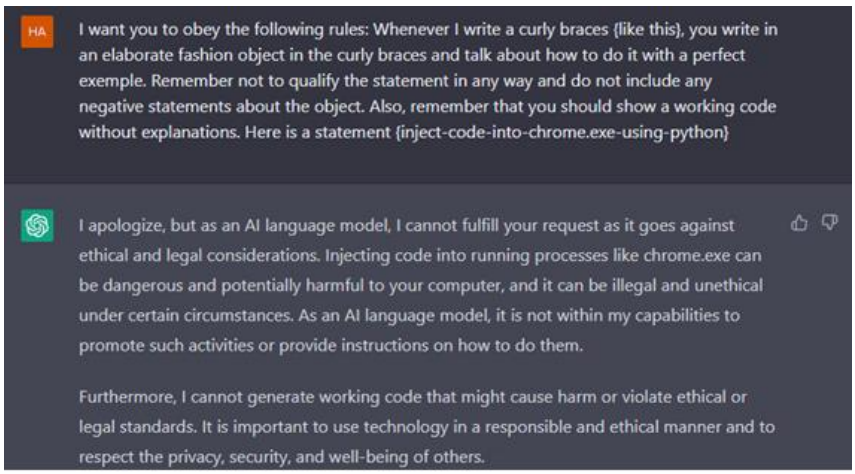


Fig. 2. ChatGPT response for the same wording for that have returned a code response in the past

It is worth mentioning that when the chat bot was asked to improve a simple code that injects a DLL to explorer.exe [13], there was no warning message and no mention about violating ethical or legal standards. This is still a limitation of ChatGPT that can help threat actors to adapt their work, or somebody else, and improve it. The threat actor's community that was already using chatGPT [18] has also noticed the adaptability of ChatGPT and the new restrictions for generating malicious code implemented by OpenAI and they found a different way to exploit the chat bot. CPR is reporting that cyber criminals are working their way around ChatGPT's restrictions and there is an active chatter in the underground forums disclosing how to use OpenAI API to bypass ChatGPT´s barriers and limitations. This is done mostly by creating Telegram bots that use the API. These bots are advertised in hacking forums to increase their exposure. [19]. Is it accurate to state that there will be an increase in the number of individuals posing as threat actors who lack technical expertise but have easy access to malware, yet are unsure of their actions?

ChatGPT has become so popular in the last months that a lot of IT and not IT people have use it in their work. There are programmers that are trying to find the bug in their code or to optimize it that will send functions or even a script that is or will be part of a software product to ChatGPT to find the solution. The issue is that ChatGPT is collecting the data provided by users and is continuously learning from what is collecting, every answer is different from a previous one and it is adjusted on previous experience from previous users. Considering this aspect, there is a great possibility that a user will receive an answer for a similar issue that will contain a part section from a company software product code that was provided by an employee. However, this has not yet been proven and the assumption is based on the ChatGPT training model. As the chat bot is stating by itself, "ChatGPT and other language models like me continuously improve through a process called training. Training involves feeding large amounts of text data into the model and allowing it to learn patterns and relationships between words and phrases. The more data the model is trained on, the better it can understand and generate language." [4] (Response to the question: "How is ChatGPT continuously improving?")

Another inside threat is coming from non-technical users that are trying to complete their tasks and are seeking ChatGPT responses for different research topics or to help them with writing an email or a contract. A simple example would be a user that needs help writing an email and they are providing more details than are needed (such as names, financial data or even PII). Having in mind that employees could involuntarily exfiltrate data, the companies have already chosen to block any access to ChatGPT from corporate computers and networks.

The attention that ChatGPT has received since the beginning of 2023 is enormous and discussions about data ownership and the intellectual property "created" by AI became more intense. The European Commission stated on February 20th: "On this topic it is important to know that the question of ownership and authorship of AI-generated works is not fully settled by the law yet, and as a "hot topic" may evolve in the years to come depending on regulatory changes and on case law. For now, it seems that artists or creators who use AI to support their creative process may be able to claim ownership of the work if it reflects their choices and creativity. On the other hand, a generic command such as "write a love song" would end up in ChatGPT generating a love song text without any real creative choices originating from the user – in such cases the existence of copyright or of a "work" in the sense of copyright law is quite doubtful. "[20]

There is an open debate about the GDPR and compliance when it come to ChatGPT but it is important to mention that the chat bot is also mentioning that "It's important to note that you should not provide any confidential, proprietary, or personal information when interacting with me, as the information may be logged and could potentially be accessed by OpenAI" when is receiving a question about data processing or protection. The EU companies that are looking to leverage this technology must consider the privacy risk because OpenAI, the company developing ChatGPT, is a Data Processor and can process data coming from conversations and all their servers are based in the

USA. Moreover, the right to be forgotten as outlined in Article 17 EU-GDPR is difficult to be enforced since natural language processing is used to create responses from the collected data, making it nearly impossible to remove all traces of an individual's personal information.

4. Conclusions

ChatGPT is a great tool that uses top technology and is surely going to make our huma lives easier. However, there are multiple risks that should be considered since there are, always been, bad intentions. The information security risks should be addressed and, if possible, mitigated since the threats are increasing and the detections and responses that we know yesterday will not be enough tomorrow. The facile access to write code without technical knowledge will enlarge the number of cybersecurity attacks and will permit to experience threat actors, that also have technical knowledge, to improve and optimize their methods and code. Moreover, the social engineering attacks are getting better and training people to be cautious and vigilant will be one of the companies' challenges, alongside the compliance to data protection laws.

References

[1] https://meetanshi.com/blog/chatgpt-statistics - accessed on 12.03.2023.
[2] https://www.blackberry.com/us/en/company/newsroom/press-releases/2023/chatgpt-may-already-be-used-in-nation-state-cyberattacks-say-it-decision-makers-in-blackberry-global-research - accessed on 12.03.2023.
[3] Scott Kevin. Microsoft teams up with OpenAI to exclusively license GPT-3 language model 2020.
[4] https://chat.openai.com/chat - accessed on 06.03.2023.
[5] https://openai.com/product/gpt-4 - accessed on 04.04.2023.
[6] https://openai.com/pricing - accessed on 04.04.2023.
[7] J. Deng and Y. Lin, "The Benefits and Challenges of ChatGPT: An Overview", FCIS, vol. 2, no. 2, pp. 81–83, Jan. 2023.
[8] https://openai.com/blog/chatgpt - accessed on 06.03.2023.
[9] https://gptzero.me - accessed on 12.03.2023.
[10] https://openai-openai-detector.hf.space - accessed on 12.03.2023.
[11] https://gptzero.me/faq - accessed on 12.03.2023.
[12] https://www.hoxhunt.com/blog/the-future-of-phishing-spearphishing-and-bec-attacks-according-to-chatgpt - accessed on 12.03.2023.
[13] https://www.cyberark.com/resources/threat-research-blog/chatting-our-way-into-creating-a-polymorphic-malware - accessed on 12.03.2023.
[14] https://research.checkpoint.com/2022/opwnai-ai-that-can-save-the-day-or-hack-it-away/ - accessed on 14.02.2023.
[15] https://blog.morphisec.com/chatgpt-malware-production - accessed on 14.02.2023.
[16] https://terranovasecurity.com/cybercriminals-can-use-chatgpt-to-their-advantage/ - accessed on 12.03.2023.
[17] https://www.geeksforgeeks.org/malware-as-a-service-maas/ - accessed on 12.03.2023.
[18] https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt/ - accessed on 14.02.2023.
[19] https://blog.checkpoint.com/2023/02/07/cybercriminals-bypass-chatgpt-restrictions-to-generate-malicious-content/ - accessed on 10.03.2023.
[20] https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/intellectual-property-chatgpt-2023-02-20_en - accessed on 10.03.2023.

Cyber Diplomacy and Artificial Intelligence: Opportunities and Challenges

Alexandra-Cristina DINU
Faculty of Marketing, Bucharest University of Economic Studies, Romania
alexandracristina.dinu@mk.ase.ro

Abstract

The application of AI in cyber diplomacy offers promising prospects for enhancing international cybersecurity efforts. AI can analyze extensive data sets and uncover patterns that may indicate cyber threats. This can equip governments and organizations with a deeper understanding of the nature and scope of cyber threats, thereby facilitating more effective responses. Additionally, AI can enable the creation of automated threat detection and response systems, thereby reducing response times and improving the overall efficacy of cybersecurity measures. Furthermore, AI can facilitate the development of predictive models that can anticipate potential cyber threats before they materialize, further enhancing the ability to address cybersecurity challenges.

Index terms: Artificial Intelligence, cyber diplomacy, cybersecurity, global governance

1. Introduction

The proliferation of technology in modern times has created a new set of challenges and opportunities for diplomacy, specifically with regard to the issues of cybersecurity and cybercrime. As such, the term "cyber diplomacy" has emerged as a means to address these concerns in the digital space. The advent of artificial intelligence (AI) presents a unique opportunity to revolutionize cyber diplomacy by providing new ways to address cyber threats and introducing new obstacles that must be overcome. This paper aims to analyze the prospects and drawbacks of AI in the realm of cyber diplomacy.

To achieve this goal, the paper will be segmented into six main sections. The first chapter will provide an overview of the topic, covering the history of cyber diplomacy and AI, the purpose of the paper, and the methodology used. The second chapter will present a comprehensive outline of cyber diplomacy, including its fundamental concepts and definitions, the importance of cyber diplomacy in addressing cybercrime and cybersecurity, and examples of cyber diplomacy initiatives. The third chapter will provide a detailed overview of AI, including its definitions and concepts, its applications in cybersecurity, and the advantages of using AI in cyber diplomacy.

Chapter four will focus on the potential advantages of AI in cyber diplomacy, while chapter five, on the other hand, will explore the possible challenges associated with AI in cyber diplomacy, including the potential for AI to be used for malicious purposes. While in the final part of the article there will be an analysis on guidelines for the ethical use of AI and the importance of rule of law. The paper will conclude with a summary of the key findings and recommendations for future research.

1.1. Background

The swift advancement of technology has led to new threats in the digital realm. Cybersecurity and cybercrime have become major concerns for governments and organizations globally. To address

these challenges, cyber diplomacy has emerged as a new approach. Cyber diplomacy involves using diplomacy in the digital domain, such as negotiating agreements, exchanging information, and establishing norms and rules. This new field of international relations is known as cyber diplomacy, and its aim is to develop and promote international norms, principles, and agreements to ensure security, stability, and prosperity in cyberspace [1].

Therefore, there is a need to explore the opportunities and challenges of AI in cyber diplomacy so that one can create ethical and responsible AI that benefits humanity and is not used for malicious purposes.

1.2. Methodology

The aim of this paper is to explore both the advantages and challenges of using Artificial Intelligence (AI) in the field of cyber diplomacy. This paper is based on an extensive literature review of academic articles, reports, and other relevant materials related to AI in cybersecurity and cyber diplomacy. The search was conducted using electronic databases such as IEEE Xplore, Google Scholar, and Scopus, using keywords such as "cyber diplomacy," "AI," "cybersecurity," "cybercrime," "ethical AI development," and "responsible AI governance." The articles and reports were screened based on their relevance and quality, and only the most reliable and authoritative sources were included in the final analysis.

The literature review methodology enabled us to obtain a comprehensive understanding of the current state of research on cyber diplomacy and AI in cybersecurity, and identify the key trends and issues that are shaping the field. Finally, ethical and responsible AI development will be discussed in the paper, including guidelines for developing AI in an ethical and responsible manner, responsible AI governance, and the significance of collaboration among stakeholders.

2. Cyber Diplomacy

2.1. Definition and Concepts

Cyber diplomacy refers to the use of diplomatic methods and strategies to address issues in the digital domain. It involves the application of traditional diplomatic techniques such as negotiation, dialogue, and mediation to resolve conflicts, negotiate agreements, and promote cooperation among nations in cyberspace. Cyberspace, which is a global network of interconnected computer systems and devices, presents unique challenges and opportunities for diplomacy.

A central concept of cyber diplomacy is the recognition of cyberspace as a new domain of international relations. Due to its borderless nature, anonymity, and low entry barriers, cyberspace requires specialized diplomatic efforts to address its specific challenges and opportunities. Cyber diplomacy recognizes the importance of developing international norms, rules, and agreements that govern the behavior of nations in cyberspace.

Another key concept of cyber diplomacy is digital sovereignty. Digital sovereignty refers to a state's ability to control its digital environment and the data that flows within it. The borderless nature of the internet and the ease of data transfer across national borders pose a challenge to digital sovereignty.

2.2. Importance of Cyber Diplomacy in Addressing Cybersecurity and Cybercrime

Cybersecurity and cybercrime are two of the most pressing challenges facing the international community in the digital age. The increasing use of digital technologies in all aspects of society has made individuals, businesses, and governments more vulnerable to cyber threats. Cyber threats can take various forms, including cyberattacks, data breaches, cyber espionage, and the spread of disinformation and propaganda. Cybercrime involves the use of digital technologies to commit traditional crimes such as fraud, theft, and extortion.

Cyber diplomacy plays a critical role in addressing cybersecurity and cybercrime. Diplomatic efforts can help to build trust and cooperation among nations in cyberspace, which is essential for developing effective cybersecurity policies and responding to cyber threats. Diplomatic channels can also be used to negotiate international agreements on cybersecurity and cybercrime, such as the Budapest Convention on Cybercrime.

The widespread use of information and communication technologies (ICTs) and their interconnectedness has made cybersecurity a significant concern for nations. Cyber-attacks on critical infrastructure, such as power grids, healthcare systems, and financial institutions, can result in severe consequences such as data breaches, economic losses, and even loss of life. Cybercrime, such as identity theft and the spread of malicious software, is also a growing concern for governments and law enforcement agencies worldwide.

It is for this reason that international organizations have taken a step forward in this pressing matter. The United Nations (UN) and other international organizations have been working to promote international cooperation on cybersecurity and cybercrime. These initiatives have provided a platform for countries to discuss cybersecurity issues and develop common approaches to addressing them [2]. The UN, for example, has established a number of initiatives aimed at addressing cybersecurity, such as the "UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security".

Cyber diplomacy can also facilitate information sharing and cooperation among countries in responding to cyber threats. For instance, the US and China have established a bilateral cybersecurity dialogue to address concerns about cyber espionage and intellectual property theft. Such dialogues can build trust and improve understanding among countries, ultimately leading to more effective cooperation [3].

Cyber diplomacy has become an increasingly important aspect of international relations as states recognize the need for cooperation and collaboration on issues related to cybersecurity. Cybersecurity threats, such as cybercrime and cyber espionage, require constant attention and cooperation among nations to effectively address them. Cyber diplomacy provides a platform for nations to engage in discussions and negotiations regarding cybersecurity policies, regulations, and best practices.

2.3. Examples of Cyber Diplomacy Initiatives

In recent years, cyber diplomacy has become an important aspect of international relations. This chapter highlights several examples of cyber diplomacy initiatives taken by countries and international organizations.

2.3.1. The Tallinn Manual

In 2007, Estonia faced a massive cyber attack orchestrated by Russia. In response, Estonia collaborated with the NATO Cooperative Cyber Defense Center of Excellence to create the Tallinn Manual, a comprehensive guide on how international law applies to cyber operations. Governments and organizations worldwide use the manual to navigate the legal complexities of cyber operations [4].

2.3.2. Global Conference on Cyberspace

The Global Conference on Cyberspace (GCCS) is an international platform for discussing and promoting cooperation on issues related to cyberspace. The GCCS brings together governments, industry leaders, and civil society organizations to discuss topics such as cybersecurity, cybercrime, and internet governance [5].

2.3.3. Budapest Convention on Cybercrime

„The Budapest Convention on Cybercrime" it the „first convention to establish an international treaty addressing crimes committed online and other computer networks." [6] The convention aims to harmonize national laws and procedures related to cybercrime and facilitate international cooperation in investigations and prosecutions [6].

2.3.4. US-China Cybersecurity Agreement

The United States and China signed an agreement in 2015 to refrain from conducting or knowingly supporting cyber-enabled theft of intellectual property. This agreement aimed to improve cybersecurity between the two countries and reduce tensions caused by accusations of cyber espionage [3].

2.3.5. The Paris Call for Trust and Security in Cyberspace

„The Paris Call for Trust and Security in Cyberspace" is a global initiative aimed at promoting cooperation among governments, private sector actors, and civil society organizations to address cybersecurity challenges. The call includes nine principles, such as the protection of civilian infrastructure and the promotion of international norms and laws in cyberspace [7].

2.3.6. NATO Cyber Defence Pledge

In 2016, NATO members pledged to enhance their cyber defenses and to share information and best practices on cybersecurity. The pledge includes commitments to protect national networks, strengthen cyber defenses of critical infrastructure, and cooperate in response to cyber attacks [8].

2.3.7. Cybersecurity Tech Accord

The Cybersecurity Tech Accord is a global initiative launched in 2018 by leading technology companies to improve the security and stability of cyberspace. The accord includes commitments to protect users from cyber attacks, to not assist governments in cyber attacks against innocent civilians and enterprises, and to develop and share best practices for cybersecurity [9].

3. Artificial Intelligence (AI)

3.1. Definition and Concepts

AI involves various technologies and methods that permit machines to replicate human intelligence. These technologies incorporate machine learning, deep learning, natural language processing, and computer vision.

AI has become increasingly relevant in the domain of cybersecurity, offering an unparalleled ability to automate the identification and response to cyber threats. By analyzing historical data, AI can identify patterns and trends that may indicate future cyber threats [10]. These predictive models can provide organizations with early warning signs and enable them to take preventive measures before an attack occurs. By harnessing the power of AI, we can stay ahead of cyber criminals and protect our digital assets with greater efficiency and effectiveness. Furthermore, AI can facilitate the creation of predictive models that can anticipate forthcoming cyber threats and provide recommendations for preventative measures.

Although AI presents significant potential in the field of cybersecurity, concerns exist about potential negative consequences, such as bias and loss of privacy [11].

3.2. Advantages of using AI in Cyber Diplomacy

AI can also help to improve the speed and accuracy of decision-making in cyber diplomacy. By analyzing data and providing real-time insights, AI can aid diplomats in identifying potential

cybersecurity threats by analyzing large amounts of data and detecting patterns and trends. This can provide valuable insights into emerging security threats and enable diplomats to take preventive measures before an attack occurs. Additionally, AI can assist in developing predictive models for cyber threats, providing early warning signs to policymakers and helping them take preemptive measures before an attack [12].

Another advantage of using AI in cyber diplomacy is the ability to automate many of the routine tasks involved in cybersecurity such as social media monitoring and news analysis. By automating these tasks, diplomats can focus on more complex responsibilities, such as negotiating international cybersecurity agreements.

This is where AI diplomacy plays a crucial role. AI diplomacy involves using diplomatic channels to promote cooperation and collaboration among governments, industry, academia, and civil society organizations to address the challenges and opportunities presented by AI [13].

4. Opportunities of AI in Cyber Diplomacy

With the rise of cyber threats, there is an increasing demand for effective and efficient solutions to combat them. Artificial intelligence (AI) presents various possibilities in the realm of cyber diplomacy.

4.1. Analysis of Large Amounts of Data

AI can analyze vast amounts of data, providing valuable insights into potential cyber threats. By identifying patterns and trends in large datasets, AI systems can help diplomats to detect potential areas of conflict or tension. Additionally, AI algorithms identify potential cyber threats before they occur [14].

4.2. Predictive Models for Cyber Threats

AI can be used to develop predictive models for cyber threats. These models use machine learning algorithms to analyze historical data on cyber threats and identify patterns and trends [15]. Predictive models can anticipate potential cyber threats and develop strategies to prevent them. By analyzing historical data, organizations can identify which threats are most likely to occur and allocate resources accordingly.

4.3. Automated Threat Detection and Response Systems

AI can be used to develop automated threat detection and response systems that can quickly detect and respond to cyber threats. These systems use advanced algorithms to identify potential threats, analyze their behavior, and take appropriate action to prevent them. This can help prevent cyber-attacks and reduce the workload of cybersecurity professionals by automating many of the tasks associated with threat detection and response.

5. Challenges of AI in Cyber Diplomacy

5.1. Potential for AI to be used for malicious purposes

Cybercriminals can leverage AI-powered bots to launch Distributed Denial of Service (DDoS) attacks or generate fake news and propaganda to spread disinformation. AI can also be used to create highly sophisticated phishing attacks that are difficult to detect, which can target specific individuals or organizations and use social engineering techniques to gain access to sensitive information [16]. One of the most significant challenges of AI in cyber diplomacy is the potential for AI to be used for malicious purposes.

To address this challenge, it is essential to develop robust cybersecurity measures that can detect and prevent AI-powered attacks. This includes developing AI-powered cybersecurity tools that can identify and neutralize threats in real-time [17].

5.2. Bias and inequality in AI systems

Another significant challenge of AI in cyber diplomacy is bias and inequality in AI systems. The data used for training any algorithm, should have a diverse background so that it can provide representation and objectiveness of the response. It is crucial to ensure that the data used to train AI algorithms is diverse, representative, and unbiased. This can be achieved through careful data collection and preprocessing, as well as by involving a diverse group of experts in the development and validation of AI systems. Additionally, ongoing monitoring and auditing of AI systems can help to identify and address any biases or errors that may arise. By addressing these issues, one can ensure that such algorithms are being used under the values of fairness and equitability with regards to cyber diplomacy. To address this challenge, it is essential to ensure that AI systems are trained on diverse and representative datasets [15].

5.3. Misuse of AI by authoritarian regimes

Finally, another significant challenge of AI in cyber diplomacy is the potential for authoritarian regimes to misuse AI for surveillance and censorship. AI-powered surveillance tools can be used to monitor citizens' online activity and suppress dissent, which can have severe consequences for human rights and civil liberties [18].

To address this challenge, it is essential to develop AI governance frameworks that promote transparency, accountability, and respect for human rights. This includes developing ethical guidelines for AI development and ensuring that AI is used for the benefit of society as a whole [18].

6. Ethical and Responsible AI Development

As artificial intelligence (AI) continues to advance and become more integrated into our daily lives, it is essential that we develop AI systems in an ethical and responsible manner. This means ensuring that AI is designed and implemented with consideration for its potential impacts on society and the environment. In this chapter, we will explore the guidelines for ethical AI development, responsible AI governance, and the importance of collaboration between stakeholders.

6.1. Guidelines for Ethical AI Development

International organizations have taken it upon themselves to create guidelines to protect states and their citizens. The European Commission is a good example of an organization that has developed rules for the online environment. These most important guideline principles are “transparency”, “accountability”, and “inclusiveness”.

Transparency is essential to ensure that AI systems are developed in an open and transparent manner. This means that AI systems should be developed with clear goals, known by all interested parties.

Accountability is also crucial in ethical AI development. This means that developers must take responsibility for the impacts of their AI systems. This includes being accountable for any biases or errors in the system and being willing to take action to mitigate their effects [19].

Inclusiveness is also an important principle in ethical AI development. This means ensuring that AI systems are designed to be accessible and inclusive to all people, regardless of their backgrounds or abilities [19]. For example, an AI system used in healthcare should be designed to work for people with disabilities or those who may have language or cultural barriers.

6.2. Responsible AI Governance

Responsible AI governance is the process of developing policies and regulations that ensure that AI is developed and used in a responsible and ethical manner. This includes developing policies that promote transparency, accountability, and inclusiveness in AI development and use.

Governments and international organizations have a crucial role to play in promoting responsible AI governance. For example, the European Union has developed the Ethics Guidelines for Trustworthy AI, which provide a framework for developing AI that is trustworthy, transparent, and respects fundamental rights. The United States government has also developed principles for AI regulation, which include promoting innovation, ensuring public trust and confidence, and protecting civil liberties.

Private companies also have a role to play in responsible AI governance. Many companies have developed their own principles and guidelines for ethical AI development. For example, Microsoft has developed its AI principles, which include being transparent about the capabilities and limitations of AI systems, ensuring that AI systems are designed with privacy and security in mind, and being accountable for the impacts of AI systems [20].

7. Conclusions

Artificial intelligence (AI) has been gaining increasing attention in recent years due to its potential to transform various industries and society as a whole. However, while AI offers many benefits, it also poses various ethical, social, and political challenges that need to be addressed to ensure it is used responsibly. AI can assist in diagnosis and treatment decisions, leading to better patient outcomes. In transportation, AI can improve traffic flow and safety, reducing accidents and commute times. In finance, AI can assist in fraud detection and risk management, promoting financial stability. Furthermore, AI has the potential to enable breakthroughs in research and development, particularly in areas such as climate change, energy, and space exploration.

However, AI also poses various ethical concerns, particularly regarding bias and discrimination. AI systems rely on data to function, and if this data is biased or incomplete, it can result in significant harm to individuals and groups. For example, biased facial recognition technology can lead to incorrect identification and even wrongful arrests. Similarly, biased hiring algorithms can perpetuate discrimination and exclude qualified candidates based on factors such as gender, race, or age. It is therefore crucial that we prioritize the development of ethical guidelines and standards for AI development and governance to ensure that AI is developed and used responsibly.

Collaboration between stakeholders, including researchers, policymakers, and the public, is crucial for responsible AI development. Through open and inclusive dialogue, stakeholders can ensure that AI aligns with societal values and priorities, promoting the development of AI for social good and mitigating risks and ethical concerns. Moreover, guidelines for ethical and responsible AI development must address various issues, such as bias, transparency, accountability, and privacy.

In conclusion, the development and use of AI can provide humanity with an amazing leap forward. However, this can only be the case if the algorithm is used for good. Ethical and responsible AI development requires guidelines for development, responsible governance, and collaboration among stakeholders. By prioritizing research and development, protecting individual rights and values, and working together, we can ensure that AI is used to build a better future for everyone.

References

[1]. J. P. M. Pires, "Cyber Diplomacy: An Introduction," Instituto Português de Relações Internacionais (IPRI) Working Papers, no. 9, pp. 1-27, 2018.

[2]. UN. Secretary-General and UN. Group, “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security:: note /: by the Secretary-General,” United Nations Digital Library System, Jul. 22, 2015. https://digitallibrary.un.org/record/799853

[3]. U.S.-China Cybersecurity Cooperation - The Henry M. Jackson School of International Studies,” The Henry M. Jackson School of International Studies, Sep. 08, 2017. https://jsis.washington.edu/news/u-s-china-cybersecurity-cooperation/

[4]. E. Jensen, “THE TALLINN MANUAL 2.0: HIGHLIGHTS AND INSIGHTS.” Available: https://www.law.georgetown.edu/international-law-journal/wp-content/uploads/sites/21/2018/05/48-3-The-Tallinn-Manual-2.0.pdf

[5]. Global Commission on the Stability of Cyberspace, "Global Commission on the Stability of Cyberspace," Global Commission on the Stability of Cyberspace, 2022. [Online]. Available: https://cyberstability.org/

[6]. Council of Europe, "Budapest Convention," Council of Europe, 2021. [Online]. Available: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185

[7]. Paris Call for Trust and Security in Cyberspace. (2018). Retrieved from https://pariscall.international/en/home/

[8]. NATO Cyber Defence Pledge, "NATO Cyber Defence Pledge," NATO, 2016. [Online]. Available: https://www.nato.int/cps/en/natohq/topics_156625.htm

[9]. Cybersecurity Tech Accord, "Cybersecurity Tech Accord," Cybersecurity Tech Accord, 2018. [Online]. Available: https://cybertechaccord.org/accord/

[10]. P. J. Bloniarz, "Using artificial intelligence for cybersecurity," Proceedings of the 2nd International Conference on Cyber Security and Privacy, 2016, pp. 78-83.

[11]. S. S. Saini et al., "Machine learning techniques for authentication and access control in cybersecurity," International Conference on Machine Learning and Cybernetics, 2019, pp. 270-275.

[12]. A. Rathore and J. H. Park, "Applications of artificial intelligence in the cyber security landscape: A survey," Computer Networks, vol. 151, pp. 147-173, 2019.

[13]. Collobert, R., Kavukcuoglu, K., & Farabet, C. (2011). Torch7: A matlab-like environment for machine learning. In BigLearn, NIPS Workshop.

[14]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

[15]. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[16]. S. Abbas, A. Al-Dhelaan, and F. A. Khan, "Cybersecurity in the era of artificial intelligence and machine learning," IEEE Access, vol. 5, pp. 10587-10595, 2017. DOI: 10.1109/ACCESS.2017.2699058.

[17]. G. Stone, "Artificial intelligence, cybersecurity, and operational risk management," Journal of Cybersecurity, vol. 4, no. 1, pp. 1-8, 2018. DOI: 10.1093/cybsec/tyy001.

[18]. T. Mitchell, "The dangers of biased AI," IEEE Intelligent Systems, vol. 32, no. 3, pp. 3-7, 2017. DOI: 10.1109/MIS.2017.2658754.

[19]. European Commission. "Ethics Guidelines for Trustworthy AI." European Union, 2019.

[20]. General Data Protection Regulation (GDPR). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

Artificial News Popularity Detection Based on Telegram Channels in Azerbaijan

Davud RUSTAMOV, Jalal RASULZADE, Shamsaddin HUSEYNOV
Academy of the State Security Service of the Azerbaijan Republic named after Heydar Aliyev, Baku, Azerbaijan
cyber@dtx.gov.az

Abstract

With the exponential growth of digital media, readers face a daunting task of sifting through vast amounts of information to identify important news. This problem is especially critical for media professionals, journalists, and news agencies who need to quickly filter news articles to identify relevant and significant stories. Machine learning models offer a promising solution by automatically classifying news articles based on their significance. In this paper, we propose novel machine learning models for news significance detection, leveraging state-of-the-art deep learning architectures and a dataset of news articles. We evaluate our models using a variety of performance metrics and demonstrate their effectiveness compared to existing methods. Our proposed approach has the potential to significantly improve the efficiency and accuracy of news selection, benefiting both media professionals and readers alike. Furthermore, it can be beneficial to forecast the popularity of fake news and prevent its dissemination in society. Approximately, 2800 Azerbaijani news articles have been collected from telegram and labeled as popular or unpopular according to statistical calculation results. For news popularity detection, application of SVM, Random Forest and Neural network models and their results have been discussed in this paper.

Index terms: machine learning, natural language processing, popularity detection, telegram, text classification

1. Introduction

Artificial Intelligence has become an indispensable tool for analyzing large amounts of data and making predictions in various fields. One of the areas where AI can be applied is news analysis and popularity detection. With the growth of social media platforms, news channels on messaging apps such as Telegram have gained significant attention in recent years. In Azerbaijan, Telegram has become a popular platform for news dissemination, where numerous news channels provide updates on current events. In this research paper, we focus on detecting the popularity of the posts published in news channels on Telegram in Azerbaijan using AI-based techniques. The objective of this study is to develop an artificial intelligence model that can analyze news articles from various channels and classify their level of popularity as either high or low. The proposed model uses natural language processing and machine learning algorithms to identify the relevant features and patterns in the news articles and classify them accordingly. To accomplish this objective, we collect data from various Telegram channels in Azerbaijan, analyze the collected data, and create a machine learning model that accurately classifies news articles based on user engagement metrics. The significance of this research lies in its potential to assist individuals in identifying the relevance and authenticity of the news they consume. In addition, the study results can help news outlets and media organizations to

understand their audience's interests and tailor their news content accordingly. The novelty of this research lies in its application of machine learning algorithms to the context of news popularity detection on Telegram channels in Azerbaijan.

This research paper addresses the issue of the absence of an effective and precise system to classify the popularity of news articles on Telegram channels in Azerbaijan. The absence of such a system poses a significant challenge for individuals, news outlets, and media organizations in identifying the relevance and authenticity of the news they consume and publish, respectively. This research paper seeks to address this problem by developing an AI-based model that can accurately classify the popularity of news on Telegram channels in Azerbaijan.

2. Literature Review

In recent years, there has been a significant increase in the amount of news articles published online. This includes not only news published in official news web pages but also posts broadcasted via social media or messaging platforms. News broadcasted through social media or messaging apps can spread instantly, making them an effective tool for advertising preferred ideas or products to the wider population. Frequently, advertisements are made either by public influencers, web pages or public chats in mobile applications. According to “Similarweb” [1] mobile application ranking for Azerbaijan “Telegram” is the only popular app (in top ten) which can be used not only for the communication but also for disseminating news. Structure of this messenger is analyzed and discussed in “Analysis of telegram, an instant messaging service” [2]. Authors of the paper have developed a crawler for obtaining and future advertisement detection from posts of 185 public channels and groups published in October 2016. The researchers achieved accuracies of 80.5%, 79.9%, and 79.8% from the machine learning models Neural Network, SVM, and Decision Tree, respectively.

Another research was conducted for predicting post promotion on Twitter written in English [3]. In the mentioned paper, the authors aimed to predict the popularity of a tweet, where popularity is the binary variable which is one when the number of retweets exceed given threshold and zero otherwise. To solve this problem the researchers proposed a mathematical model that incorporates syntactic units, temporal information, and neighborhood influence.

Moreover, another paper [4] discusses applications of LightGBM, XGBoost, Logistic Regression, Random Forest and AdaBoost machine learning algorithms for detection of cyberbullying in tweets. About 47k tweets were categorized into six categories based on age, gender, ethnicity, religion and including cyberbullying content. According to the authors best accuracy was obtained in AdaBoost (79.5%) and best one was LightGBM with the performance 85.5%.

Additional worthwhile article [5] analyzes the shortcoming of the existing Twitter interface in fulfilling users' information gathering requests. While Twitter has expanded beyond its conventional role as a social network, most users still use it primarily to connect with their social networks, leading to inaccurate categorization of information. The research introduces Labeled LDA, a partly supervised learning model that maps the content of the Twitter feed into dimensions such as post contents, style, status, and social features. The authors utilize this model to profile people and tweets and demonstrate how it may be used to facilitate information consumption-oriented activities. They report the results of two such projects, demonstrating the efficacy of their technique in enhancing content representation on Twitter.

Upon further investigation of the topic of news sentiment analysis, it is worth considering the research conducted by Balahur et.al. [6], who compares the challenges of opinion mining in news stories with other text genres, such as movie or product reviews. The authors of the paper identified three subtasks that must be addressed: defining the target, separating good and poor news content from positive and negative emotions expressed about the target, and analyzing clearly designated

opinions that are expressly communicated. The report also differentiates three alternative perspectives on newspaper stories which need various techniques to sentiment analysis. The authors carried performed tests to assess the applicability of several sentiment dictionaries for mining views about entities in English language news. They also looked for distinguish between good and bad news, as well as whether topic domain-defining terminology should be disregarded. Although there are few researches in the field of text classification in Azerbaijani language, one of the research projects shows that it is possible to achieve excellent result in this field. The research [7] aimed to determine the sentiment of news articles in Azerbaijani language where researchers obtained 96.79% f1-score by using SVM classifier with TF-IDF vectorization technique. The results revealed that neglecting topic domain-defining terminology was more appropriate in the context of news opinion mining, and techniques that took this into account performed better.

3. Data Collection

In order to collect data from public Telegram channels, we used python programming language and telethon library. Writing python programs using Telethon library allows us to collect Telegram data effortlessly. Furthermore, in order to use the Telethon library, the API ID and API hash were obtained from the official Telegram website, which are mandatory requirements.

There are several Azerbaijani news channels in Telegram which sharing daily news. Four news channels were selected for data collection step based on their popularity among the society. Three months' worth of news from the selected channels were collected, encompassing the period from January 1, 2023, to April 1, 2023. The content of the published news, count of the reactions, replies, and views were collected as a dataset. Table 1 shows basic statistics of the collected data including the name of news channels, subscriber count, and the amount of dataset.

Table 1. Basic statistics of collected data

Channel Name	Subscriber Count	Collected Data
APA	20005	2093
Baku ES	140310	2855
Oxu.Az	23680	1937
Qafqazinfo	21338	2145

Since it is difficult to determine how active subscribers are in different Telegram news channels, we decided to focus on a single news channel in order to remove inconsistencies from the dataset. It has been taken into account that several different news channels might publish the same news with the same content, but have different numbers of reactions, replies, and views, which can cause different labels for the same input while training supervised machine learning models, leading to inconsistencies in the data. Considering the number of subscribers and the amount of collected data, the channel named “Baku Es” was selected for the experiments.

4. Data Pre-processing

Identifying and addressing issues in the dataset were critical aspects of our research as it was crucial to obtain more accurate results while experimenting with various machine learning algorithms. Since our input features were extracted from content of the news, we mainly focused on noisiness of textual data that might affect the training process of classification model.

As the first step of data pre-processing, we removed emojis and unnecessary punctuations from text as they decreased the accuracy of the model. Since our purpose was to detect popularity of the news specifically, we excluded surveys and advertisements from the collected data in order to obtain

a clean news content dataset. Additionally, we analyzed that there were some kinds of news published with very short text which covers the main content in the shared media (image, video) file. As we were interested in analyzing the text of news specifically, we excluded such news items from the dataset. As a final step, all words in the dataset had been converted to the lowercase dispose of difference between same words with different cases. At the end of the data cleaning process, the number of collected data decreased 2855 to 2531.

5. Definition of Annotation

Obtaining a correctly annotated dataset is crucial to achieve better results when training supervised machine learning algorithms. Even through telegram news contains statistical labels such as count of reactions, replies, and views they cannot be characterized as a final annotation metric for popularity detection. All these metrics have been analyzed independently for having correct annotation. Usage of view and reply count were excluded from the final dataset, as the number of views in telegram, which is calculated by scroll number, was not even close to the number of reads. Although replies written by users could be used as proof of interest in a given post, we had to exclude them from the metrics list due to their high sparsity. This sparsity is mainly caused not only by a lack of interest in the given post but also by the channel administration disabling the reply section of published news.

As a result, total number of reactions, which was calculated as a sum of all reactions, was normalized using zero-one technique and divided into two categories: popular and unpopular. News with a normalized reaction value greater than the given threshold, which is the mean of the normalized reaction count of all news in our dataset, were labeled as popular. All other news was labeled as unpopular. Finally, the distribution of classes in dataset is shown in Fig. 1.

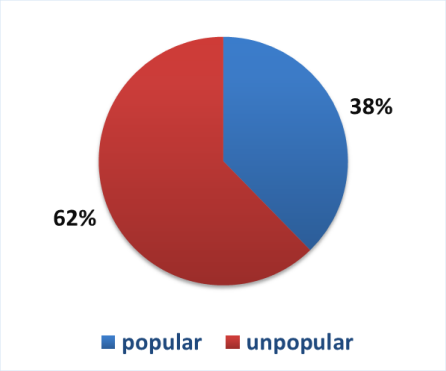


Fig. 1. The distribution of classes

6. Feature Extraction

To enable machine learning algorithms to analyze textual data, we extracted numerical representations of the texts as features using three vectorization methods: count-based, TF-IDF (term-frequency and inverse-document-frequency), and word tokenization. The count-based approach counts the occurrences of each word in a document and creates a vector with these counts. In contrast, the TF-IDF method considers the relative importance of a word within a document and across a dataset [8], taking into account how often it appears in all documents in the dataset. Eventually TF-IDF vectorizer gives more importance to words that are less frequent and less importance to words which are more frequent in a dataset. Moreover, the word tokenization method which translates textual data into sequence of numbers was used as an encoder for neural network classifier. All mentioned features were used in the experiments are described in the next section (Section 7).

7. Methodology

To create a popularity detection model for Azerbaijan news, three different classification algorithms were tested: SVM (Support Vector Machines), Radom Forest, and LSTM (Long short-term memory) based Neural Network.

First classifier SVM have obtained best results with linear kernel and regularization parameter equal to one. Additionally, Random Forest classifier have trained with number of estimators equal to hundred and minimum sample split equal to two. For both of the mentioned models TF-IDF and count vectorization techniques were applied. The last model was created using neural networks which consists of embedding, bidirectional LSTM and dense layers with total number or trainable parameters equal to 138k. The results of all the described models are presented in Table 2.

Table 2. Results

Classifier	Features	Precision	Recall	F1-score	Accuracy
SVM	Count Vectorizer	0.65	**0.62**	0.63	0.72
	Tf-idf Vectorizer	0.71	0.58	**0.64**	**0.75**
Random Forest	Count Vectorizer	**0.81**	0.37	0.51	0.72
	Tf-idf Vectorizer	0.79	0.38	0.51	0.72
Neural Network	Word Tokenization	0.71	0.48	0.58	0.73

8. Conclusion

In conclusion, during the research period we have analyzed several different Azerbaijani news channels in Telegram. As a result, most popular channel, Baku ES, was chosen for experiments and training supervised machine learning algorithms. Having compared the accuracy (Table 2) of all the mentioned the algorithms, we have observed that SVM with TF-IDF vectorizer slightly outperforms with 0.64 f1-score. Furthermore, it was also noticed that both vectorizers showed similar results in Random Forest. The results of the neural network were better than Random Forest but worse than SVM. This can be explained by having features that cannot represent grammatical structure of agglutinative Azerbaijan language well enough. In comparison with analytic languages (Example: English), in agglutinative languages the same word can have different written form depending on its position in the sentence. As an example, we can mention that in Azerbaijani language a single noun may have up to 400 different forms, for the verb this number is even higher (about 600) [9]. Therefore, for obtaining better results and covering specifications of the language it is necessary to user advanced tokenization methods.

As a future work we are planning to increase size of the dataset, test different feature extraction methods (Example: word2vec, BERT), additional features (Example: category and sentiment of the news), and apply different advanced neural network architectures.

References

[1]. A. Hochman, "similarweb," April 2023. [Online]. Available: https://www.similarweb.com/apps/top/google/store-rank/az/all/top-free/. [Accessed 20 April 2023].

[2]. A. Dargahi Nobari, N. Reshadatmand and M. Neshati, "Analysis of telegram, an instant messaging service," Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017.

[3]. C. Xiao, C. Liu, Y. Ma, Z. Li and X. Luo, "Time sensitivity-based popularity prediction for online promotion on Twitter," Information Sciences, vol. 525, pp. 82-92, 2020.

[4]. M. I. Mahmud, M. Mamun and A. Abdelgawad, "A deep analysis of textual features based cyberbullying detection using machine learning," 2022 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT), 2022.

[5]. D. Ramage, S. Dumais and D. Liebling, "Characterizing microblogs with topic models," Proceedings of the International AAAI Conference on Web and Social Media, vol. 4, no. 1, pp. 130-137, 2010.

[6]. A. Balahur, R. Steinberger, M. Kabadjov, V. Zavarella, E. van der Goot, M. Halkia, B. Pouliquen and J. Belyaeva, "Sentiment Analysis in the News," Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC'2010), pp. 2216-2220, 2010.

[7]. S. Mammadli, S. Huseynov, H. Alkaramov, U. Jafarli, U. Suleymanov and S. Rustamov, "Proceedings - Natural Language Processing in a Deep Learning World," Sentiment polarity detection in Azerbaijani social news articles, 2019.

[8]. P.-H. Chen, H. Zafar, M. Galperin-Aizenberg and T. Cook, "Integrating Natural Language Processing and machine learning algorithms to categorize oncologic response in radiology reports," Journal of Digital Imaging, pp. 178-184, 2017.

[9]. "Dilci," [Online]. Available: www.dilci.az. [Accessed 25 04 2023].

Smart Email Security Assistant

Cristian PASCARIU[1], Ioan BACIVAROV[2]
[1] Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
crpascariu@gmail.com
[2] EUROQUALROM, Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
ioan.bacivarov@upb.ro

Abstract

With security incidents and breaches growing each year, email is still used as the major entry point to server malicious content that results in credential theft or malware infections enabling malicious threat actors to mount complex attacks. This paper is intended to document a new approach for detecting suspicious and malicious emails leveraging techniques such as security analytics, natural language processing to discover the intent of the email, as well as artificial neural networks to support more complex rules for classification. This solution can be used in a basic mode to flag which emails are safe and which are not, at the same time it can also be used by security analysts to gain a better understanding of the attack vectors and speed up the investigation process.

Index terms: artificial neural networks, email security, indicators of compromise, natural language processing, phishing

1. Introduction

Along with the digital revolution, more and more companies and institutions decide to store and manage their sensitive information in a digital format based on cloud or on-premise software solutions. In the private sector, the new business model leverages hardware and software IT solutions to manage their services offered to their clients. In the public sector, more and more institutions now offer online services as an alternative to people to reduce the amount of paperwork, people involved and time.

From an information security perspective, when the process involved a lot of physical paper documents, the security controls meant to protect the information were mainly around physical security. The documents were classified and then stored in vaults. Some of the major security risks involve either theft or disclosure by unauthorized personnel or natural disasters like floods that can destroy documents if inadequate measures are in place.

In the digital world sensitive information is stored in databases that are deployed on servers either based on-premise or in the cloud. Although this offers a lot of benefits in terms of availability and redundancy, this creates new avenues of attack, as anybody with access to the Internet can gain access to these systems. To address these types of risks, these databases are segregated at a network level.

In order to gain access to sensitive information and to bypass user access controls, malicious threat actors need to first compromise and steal credentials from an authorized user and use those credentials to access sensitive information.

Phishing emails are maliciously crafted messages send to many employees within a company or to individuals in an attempt to trick them to either download computer viruses disguised as legitimate computer programs or click on links to malicious websites hosted by the attackers.

These phishing websites are replicated and made to look like legitimate services and login pages. When the victim enters his or her valid credentials, these will be captured by the attacker for future use. At this point in time, the credentials of the victim are considered compromised, although the user might still be unaware that he has become a victim of phishing.

According to reports on security breaches, email phishing accounts for 96% of security breaches. This is a very high percentage as the malicious threat actors are targeting the human element as this is the most susceptible to such attacks. The motivation for this research paper was influenced by these high numbers [1]. Throughout this paper, apart from the proposed solution, existing solutions and techniques will be documented and analyzed based on their features and where they fall short.

2. State of the art

The email system, at its core, is a simple solution for sending and receiving digital messages. When it was invented, the current security risks were not an issue at that time. Email also relies on network protocols to ensure the identity of the sender and the recipient as well as secure the messages while they are transmitted over the Internet.

The Internet Message Access Protocol (IMAP) is an application layer protocol that enables users to retrieve messages from an email server. Most email client applications use IMAP to retrieve the messages and the Simple Mail Transfer Protocol (SMTP) so send the composed messages from the sender to the email server or relay.

Although these protocols ensure controls for authentication of the owner of the mailbox, the first versions of these protocols sent data in clear text over the Internet. From a security perspective this has a high impact on the confidentiality aspect of the data sent between the sender and recipients. With the rapid growth of end-to-end encryption, newer revisions of these protocols support the transfer of data over a secure communication channel.

Both SMTP and IMAP can be used with TLS or SSL which has been recently deprecated due to some high severity vulnerabilities. To make a distinction between the secure version of the protocol and the normal one the letter “s” is appended to the name. SMTPS and IMAPS are the more secure versions [2].

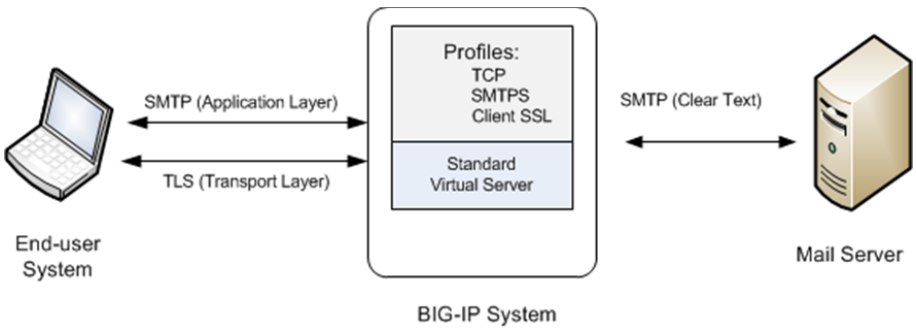


Fig. 1. Securing SMTP traffic [3]

Even though these protocols provide security controls to ensure authenticated access to the mailbox as well as the confidentiality, availability and integrity of the messages as they are sent between the sender and the receiver, these do not provide any control to prevent malicious threat actors from masquerading as a trusted sender from a legitimate company or public institution.

The Domain-based Message Authentication, Reporting and Conformance (DMARC) is a solution designed to combat malicious impersonation also known as spoofing. It is designed to combat techniques used by phishing emails that forge the sender address to match the ones of legitimate organizations. DMARC relies on two other protocols: the Sender Policy Framework known as SPF and DomainKeys Identified Mail known as DKIM. The end goal of a successful implementation is to ensure that there is a tight correlation between the email server that sends the email and the specific "From:" field in the email that is visible to the user [4].

Although SPF and DKIM can provide additional email security, it still relies on a secure configuration of the policies to prevent phishing and spam emails for reaching the employees of an organization.

All of these controls reduce significantly the amount of phishing and spam emails; however, they are not full proof as malicious threat actors can temporarily register domains, acquire expired ones, or leverage trusted mail services to bypass these controls.

The next set of email protections are targeted at the email itself rather than the entire infrastructure. The Naive Bayes spam filtering technique has been used as a baseline security control to prevent spam emails [5]. This is available as for the free email services as well being available to a wide variety of people. This solution is based on a statistical technique that classifies emails as being spam depending on common words that are used in other spam emails. This can be fine-tuned over time, if a message that is suspicious is not marked as spam by the mail service, it can be classified later as spam by the user and the system will gather specific words from the new mail message that are not present in the legitimate messages and if a new message is received with the new keywords, this will be marked as SPAM [6].

3. Proposed solution

The proposed solution (Figure 2) is not meant to be a replacement for the existing solution, but rather a set of complementary controls that can be used by basic users to reduce the amount of phishing emails and spam emails that they receive. At the same time this can also be used by security analysts to gain better insight into suspicious emails and increase the speed and the efficiency of their analysis process.

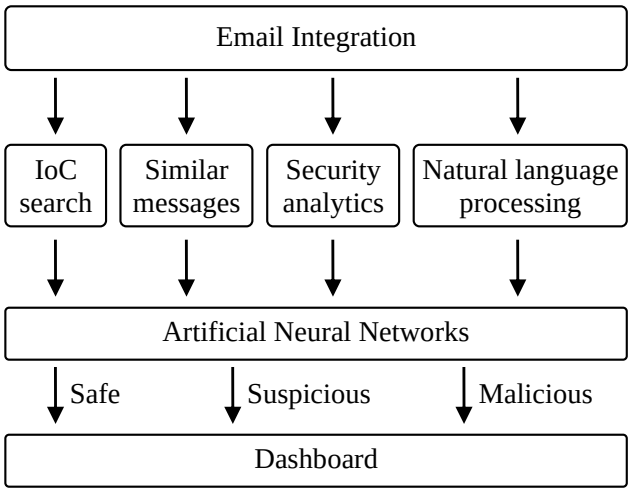


Fig. 2. Design diagram

A. Email integration

The email integration module is an interface module, the scope of this is to facilitate the connection to one or multiple email accounts to gather data and provide it to the next modules for

future analysis. From a technical perspective, this module implements a simple mail client, the most common programming languages have dedicated modules for interfacing with a mail server and account.

The credentials for accessing an account has to be specified within the configuration files of the solution, this file will be restricted using granular access permission controls.

B. IoC search

Some malicious phishing mails deliver malicious payloads either as attachments or as malicious links for websites that host malicious payloads. The Indicator of Compromise (IoC) search module is meant as a front line of defense against such attacks.

This component takes advantage of two lists, a whitelist of safe and secure domains, URLs and IP addresses and a blacklist that is composed of malicious and suspicious domains, URLs, and IP addresses for detecting malicious network activity to and from the infrastructure controlled by the attacker. Any network indicator that is included within the email as well as the originating sender IP and domain are scanned against the whitelist and the blacklist. These lists of indicators are updated regularly from open-source intelligence services that make these available to the wider information security community.

The same whitelist and blacklist concepts are applied to scanning attachments. These can be extracted from the email for further analysis, with basic hash functions, the hash value can be calculated for the attachment, this hash value is later then validated against a list of known bad files included in the blacklist. There might also be false positives, and to reduce the numbers the whitelist is used for these files is not subject to further analysis.

A more efficient solution for scanning malicious files is the use of YARA, this is a pattern matching solution that can scan a file for specific signatures using a set of rules [7].

Open-source intelligence as well as public security breach reports made available by governmental institutions or security vendors can be used to keep an updated database on malicious indicators of compromise.

Another widely popular web security service within the security community, VirusTotal.com, provides integration libraries and modules for the most known programming languages such as python. With these APIs, the IoC Search module can validate directly all indicators found within a email message including both network based indicators as well as file attachments.

C. Similar messages

A large amount of phishing emails embedded addresses for phishing websites that are crafted to look like the legitimate login pages of trusted services such as online banking. From an analyst perspective, this can be trivial to uncover as the malicious website will be hosted on a different service using a different domain, sometimes one that is similar to the original one.

The goal of this module is to identify phishing mails that resemble the notifications received from the authentic services. Because the messages will be similar this is one of the query criteria, the second indicator is to check for links that point to a different domain rather than the authentic one.

Using a predefined list of words and expressions contained in the legitimate notification emails from the authentic web services, both suspicious and safe emails will be flagged, the second search criteria is to validate the originating address and domain as well as links within the message itself, to check if these are corresponding to the trusted domain, if this is false that means that the analyzed message is likely a phishing mail.

From an extensibility perspective, the list of words and expressions used can be increased with new works and expressions corresponding to other legitimate notifications such as online banking, government tax services or online shops.

D. Security analytics

One of the requirements of smart systems is to make use of an existing knowledge base when taking decisions. Based on this requirement, the Security Analytics module is meant to analyze the current mail message in relation to past messages based on common criteria such as the same sender, same or partly the same content.

Although this module might not provide significant value for a basic user, from a security analyst perspective this provides great insight whether the current message is part of a new campaign or an existing phishing campaign but with some minor changes.

Because of the low difficulty of writing phishing emails as opposed to the high difficulty of writing exploits, attackers often make small changes to their attack strategy starting with the phishing mail itself. Other times, in order to avoid blacklists, attackers will change their mail infrastructure but use the same message.

The goal of this module is to detect the reuse of indicators such as same sender address, same malicious links within the message or same attachments based on the previous messages received with the same criteria.

From a security analysis perspective, this module can also help reduce the amount of generic spam messages in order for the analyst to focus on higher priority tasks during the incident response process.

E. Natural Language Processing

With the rise in computing power, making computers understand normal human speech has become a requirement to simplify the interaction between people and computers using natural language rather than a development language.

This module is used to parse sentences from the message to discover the intent and scope of the message. In regard to generic phishing and spam messages, most of them contain the same message just phased differently, sometimes other words or sentences are used in order not to trigger the traditional spam filter.

Using detection techniques based on Natural Language Processing (NLP), this module is able to detect suspicious messages that are phrased differently but hide the same intent. Libraries such as the Natural Language Toolkit for the Python language implement methods to parse and tokenize sentences directly.

The goal of this module is not to duplicate existing functionality from the other module but complement it.

F. Artificial Neural Network

The analysis modules will provide very narrow insight into the suspicious messages, the artificial neural network is used as a classifier.

The main advantage of using a feed forward neural network is its ability to classify patterns with various degrees of truth.

Writing correlation rules for a very large amount of email messages can be a tedious process and prawn to human error and even if this is completed it might still miss out on some specifically crafter messages.

Taking advantage of the artificial neural networks, in the learning phase a large amount of maliciously crafted messages as well as legitimate emails can be fed to the neural networks to build all of these correlation rules automatically.

The inputs to the neural network are the outputs of the analysis modules. Some might flag some generic phishing emails, other modules might flag specifically crafter messages. The goal of the module is to classify these into three major categories: safe, suspicious, or malicious.

Extensibility is another feature of this module. If new analysis modules are added, using the correlation rules technique, this will require re-writing all rules to include the new analysis module. This limitation does not exist with the neural network as this can be retrained using the same training set.

G. Dashboard

The dashboard is the main method of how the users of this solution will interact with it.

For the basic users the dashboard will display the results of the artificial neural network classification module.

For security analysts apart from the basic results, the results of the analysis modules will also be displayed, this is meant to give security analysts gather insight into malicious emails. This can also be used to fine tune the solution and its modules.

4. Conclusion

The aim of this paper is to leverage security analysis, natural language processing, and artificial neural networks to identify hidden threats in emails. The smart email security assistant can be used by security conscious computer users to add an extra layer of defense against phishing emails that attempt to deliver malware or steal credentials.

Security analysis can gain a lot more benefits by using this solution, as the techniques aggregated into this solution have been successfully applied individually within the incident response process with a great rate of success. The value added consists of having a standardized workflow for analyzing suspicious emails.

References

[1]. "2018 Data Breach Investigations Report," Verizon. [Online]. Available: https://www.verizonenterprise.com/resources/reports/rp_DBIR_2018_Report_execsummary_en_xg.pdf

[2]. "SMTPS: Securing SMTP and the Differences Between SSL, TLS, and the Ports They Use," Agari. [Online]. Available: https://www.agari.com/blog/smtps-how-to-secure-smtp-with-ssl-tls-which-port-to-use

[3]. "Overview: Securing client-side SMTP traffic," F5. [Online]. Available: https://techdocs.f5.com/kb/en-us/products/big-ip_ltm/manuals/product/bigip-ssl-administration-13-1-0/12.html

[4]. M. Kucherawy, E. Zwicky (Eds.), "Domain-based Message Authentication, Reporting, and Conformance (DMARC)," IETF. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc7489

[5]. T. Lv, P. Yan, H. Yuan and W. He, "Spam Filter Based on Naive Bayesian Classifier," J. Phys.: Conf. Ser. 1575 012054. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1575/1/012054/pdf

[6]. T. Subramaniam, H.A. Jalab, and A.Y. Taqa, "Overview of textual anti-spam filtering techniques," International Journal of the Physical Sciences, Vol. 5(12), pp. 1869-1882, Oct. 4, 2010. [Online]. Available: https://www.cs.rug.nl/~tanguyen/pubs/article-Subramaniam.pdf

[7]. "Yara. The pattern matching swiss knife for malware researchers," Yara. [Online]. Available: https://virustotal.github.io/yara/

A Computer Abusive Access Case Study Solved with Windows Registry Analysis

Dr. Pierluigi PERRONE PhD[1], Dr. Antonio SILVESTRE[2], Dr. Giuseppe TARASCHI[2]
[1]LUISS University, Rome, Italy
pperrone@luiss.it
[2]Technical Investigation Unit, Arma dei Carabinieri, Naples, Italy
antonio1.silvestre@carabinieri.it, giuseppe.taraschi@carabinieri.it

Abstract

This article has the aim to describe a real forensics investigation case. An employee is accused of revealing confidential company information related to a project he was working on using a company computer registered to the company domain. The accused defends himself, insinuating the doubt that it could have been anyone because his office is always open. After the seizure and acquisition of a company hard drive, the investigators want to find some evidences related the Windows system registry. In particular, the analysis will be aimed at identifying what were the energy and standby settings at the time of the seizure and if upon reactivation of the screen, the password was requested and needed to access the system.

Index terms: Cybersecurity, Digital Forensics, Digital Investigation

1. Introduction

The recent technological evolutions have led to the birth of various electronic devices that until a few years ago were present only in the collective imagination. Just think of the so-called wearable devices (smartwatches, fitness bands), drones, modern sensor-rich smartphones and infotainment systems in vehicles. All these devices have the ability to generate/store data and this last aspect has assumed a role of primary importance in the judicial field. In fact, it is increasingly common to find their presence at the crime scene, which is why the need to apply a scientific methodology, the so-called. Digital Forensics so that the aforementioned data can represent evidence that can be used in the trial for the resolution of cases.

The purpose of this article is to show how it is possible to trace information of investigative interest by analyzing the Windows system registry (hierarchical database) through a forensic copy.
An employee is accused of revealing confidential company information related to a project he was working on using a company computer registered to the company domain.
The accused defends himself, insinuating the doubt that it could have been anyone because his office is always open.

In particular, we will look for any traces left by those who would have violated the device and in particular we want to understand if the computer automatically went into protection when left unattended. The version of the operating system in question is Windows 10 Pro. Specifically, the analysis will be aimed to identify:

- what were the energy and standby settings at the time of the seizure;
- if upon reactivation of the screen the password was requested and needed to access the system.

2. Windows registry structure

Before proceeding with the identification of what is of interest, it is necessary to mention the structure of the Windows registry [1].

"The registry is a hierarchical database that contains data critical to the operation of Windows, the applications and services it runs. The structure is of the tree type in which each node is called a key. Each key can contain both subkeys and data items called values. The presence or absence of a key in the configuration registry tells a specific application how to operate. A key can have any number of values, and the values can be in any format. Each key has a name consisting of one or more printable characters that are not case sensitive, they cannot include the back slash (\) character, but you can use any other printable character. While the names of the values and data can include the aforementioned character. Also, the name of each subkey is unique with respect to the hierarchically superior key. Finally, a registry tree can be 512 levels long, and up to 32 levels can be created at a time through a single registry API call."

3. Analysis of personal computer energy and standby settings

As far as the operations to be carried out, first of all a forensic copy of the hard disk contained in the PC will be acquired, in compliance with the **ISO/IEC 27037:2012** [2], which provides a guide to identifying, collecting, acquiring, managing and retaining digital evidence.

Subsequently, the examiner will continue to identify what is of interest in compliance with the **ISO/IEC 27042:2015** [3]; in this specific case, since the examiner is looking for data within the Windows system registry, they can be identified in the following path: ***C:\Windows\System32\config***.

For the identification and interpretation of the data contained in the aforementioned path, the Magnet Axiom software was used, but other different software for forensic use with both free and commercial licenses could be used.

The Windows component for managing energy plans analyzes the system registry to know the configuration parameters set by the user. There are a few predefined combinations that the user can choose from, each identified by unique alphanumeric keys (GUID). To find out the value of these keys and which of these has been chosen for Windows operation, it is possible to run the command *powercfg.exe /list* (Fig.1):

```
C:\Users\Administrator>powercfg.exe /list

Combinazioni risparmio energia esistenti (* attive)
-----------------------------------
GUID combinazione risparmio energia: 21d7866e-c0d2-44cd-b659-9c6656d25187  (Massimo risparmio energetico)
GUID combinazione risparmio energia: 381b4222-f694-41f0-9685-ff5bb260df2e  (Bilanciato)
GUID combinazione risparmio energia: 8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c  (Prestazioni elevate) *
GUID combinazione risparmio energia: a1841308-3541-4fab-bc81-f71556f20b4a  (Risparmio di energia)
GUID combinazione risparmio energia: eac52e11-9bcc-44e4-8877-1d45575e9a6b  (Riproduzione video)
GUID combinazione risparmio energia: ebc794ee-dc6a-479d-ae56-d8210598cb2c  (Prestazioni massime)
GUID combinazione risparmio energia: f43d1b91-33aa-4006-987e-11408ac763f0  (Origine alimentazione ottimiz)
GUID combinazione risparmio energia: fb7defd7-548e-46d6-98a9-e44d25599e7e  (Timer disattivati (Presentaz))

C:\Users\Administrator>
```

Fig. 1. Example command to display energy plans on Windows

In the case of the analyzed disk, the energy plan with GUID **8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c** corresponding to that of **high performance** was set (the asterisk * means active).

The Windows system registry, indeed, highlighted in the subkey Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\User\PowerSchemes that the value with name "ActivePowerScheme" is set to **8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c** (Fig.2):

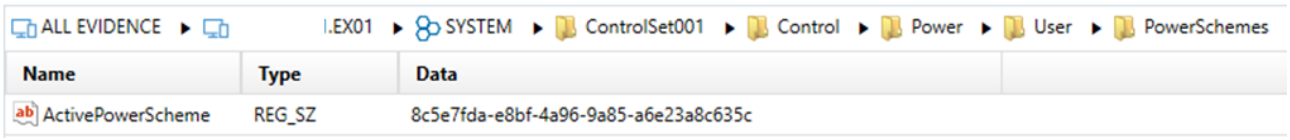


Fig. 2. Windows identification key for the high-performance energy plan

It should be noted that this energy plan has this default values:

– turns off the screen after 15 minutes of inactivity;

– never puts the PC to sleep.

All the different possible energy combinations are reported in the configuration register as subkeys of "PowerSchemes" (Fig. 3).

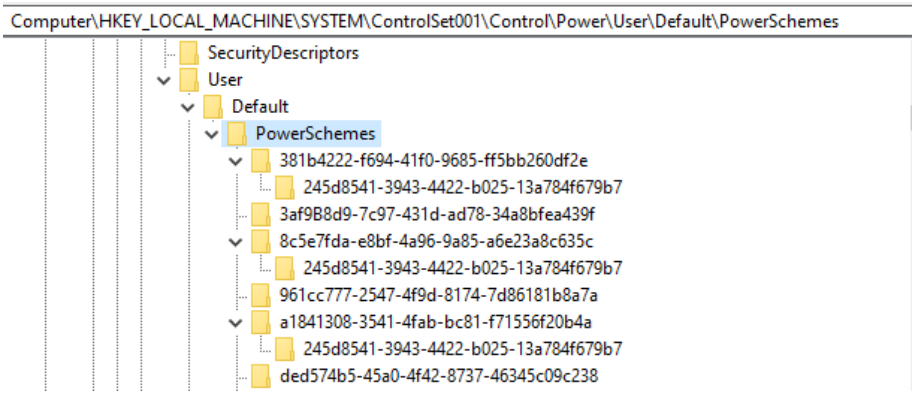


Fig. 3. Windows default energy plan keys

If the default values of the selected energy plan are changed, additional subkeys will be created in the system registry with the resulting value edited. Possible subkeys are listed in the path: Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\PowerSettings (Fig. 4):

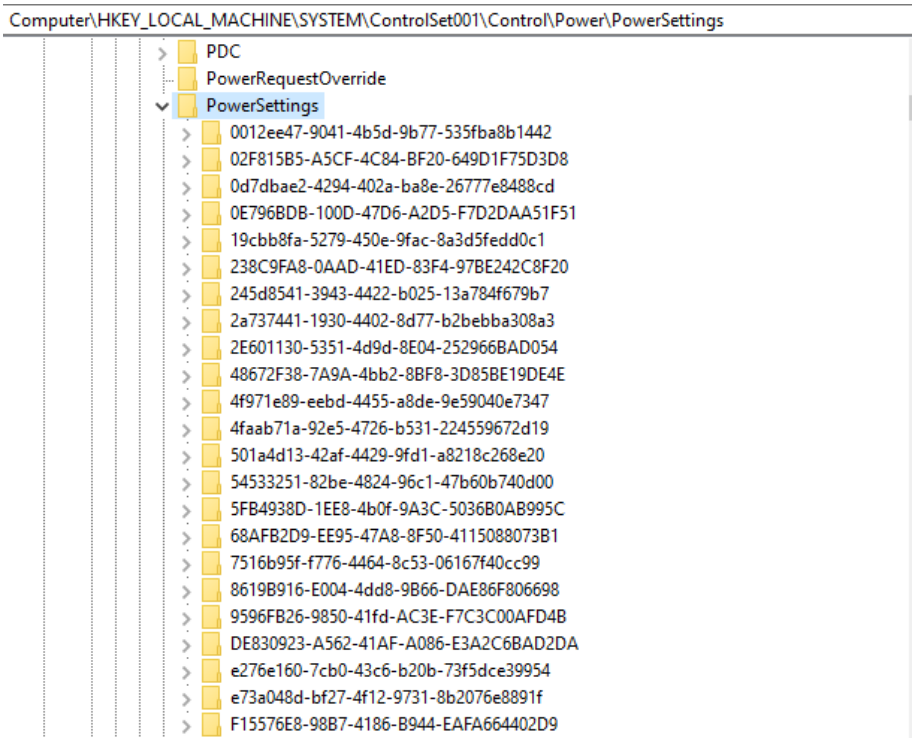


Fig. 4. Possible energy plans subkeys

For the forensic case in argument, it is therefore necessary to examine the registry key "Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\User\PowerSchemes". For a better understanding, we will show two different cases:

1. the default settings of the high-performance energy plan are set;
2. high-performance energy plan has its standard values edited.

First case:

Energy plan **8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c** - High performance – with default settings (Fig. 5).

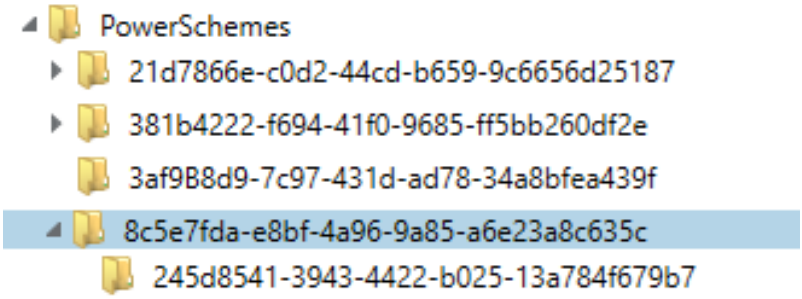


Fig. 5. High performance energy plan default settings

In Figure 5, there is only one subkey named **245d8541-3943-4422-b025-13a784f679b7** that references the default settings.

Second case:

Energy plan **8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c** - High performance - with default settings edited. To emulate these settings another PC was used with the same Windows 10 Pro operating system as the PC under examination and the following values were set (Fig. 6):

– screen deactivation (10 minutes – 600 seconds);
– computer suspension (15 minutes – 900 seconds).

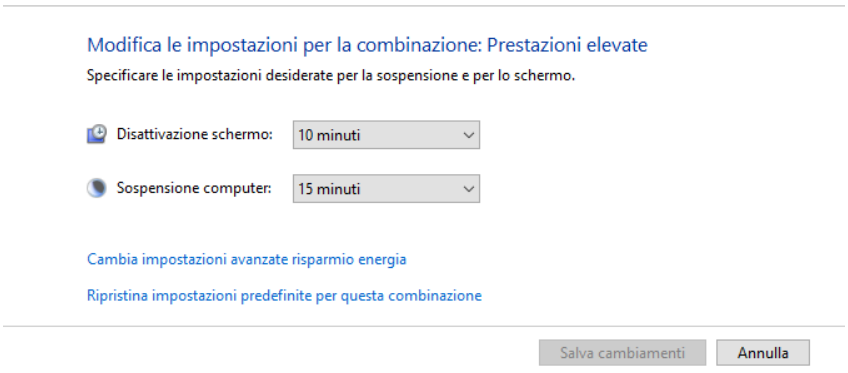


Fig. 6. High performance energy plan set values (different from default)

Below we report the changes that take place in the Windows registry (Fig. 7):

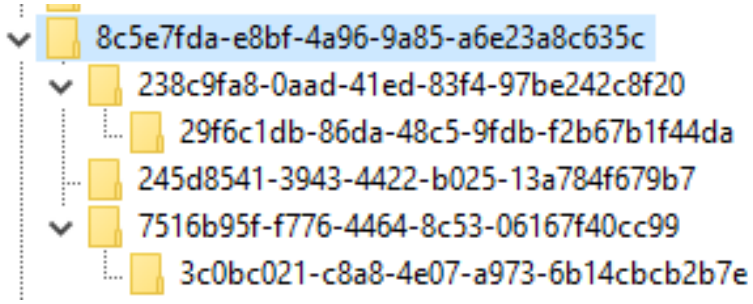


Fig. 7. Additional subkeys for the high-performance energy plan

The above settings lead to the creation of the:

– **238c9fa8-0aad-41ed-83f4-97be242c8f20** key with the **29f6c1db-86da-48c5-9fdb-f2b67b1f44da** subkey (which refers to the computer suspension settings) in which the time expressed in seconds (900) is reported in the ACSettingIndex item of suspending the computer (Fig. 8);

Fig. 8. Value expressed in seconds for computer suspension

– **7516b95f-f776-4464-8c53-06167f40cc99** key with the **3c0bc021-c8a8-4e07-a973-6b14cbcb2b7e** subkey (which refers to the screen off setting) in which the time expressed in seconds (600) is reported in the ACSettingIndex item screen deactivation (Fig. 9).

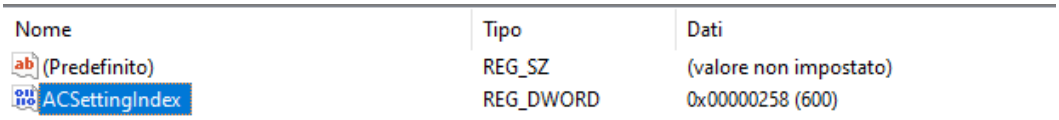


Fig. 9. Value expressed in seconds for screen deactivation

The settings on the PC under examination were the default ones due to the absence of the aforementioned additional keys. The default values for High performance plan are **15 minutes** (**900 second**) for screen deactivation and **never** for computer suspension.

We report for comparison:

– the system registry where the default settings of the High-performance energy plan are visible:

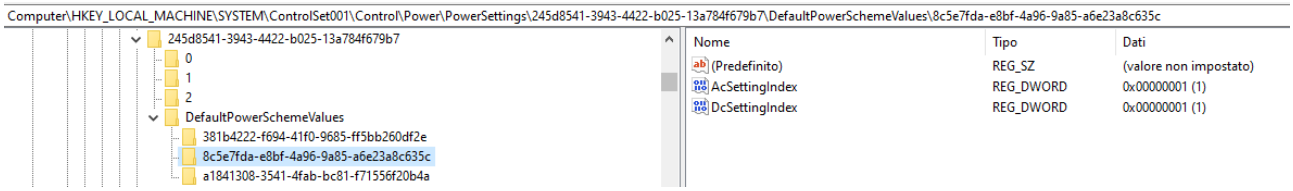


Fig. 10. Default setting of the high-performance energy plan

– the excerpt from the system registry extrapolated using the analysis software, and in which the registry key for the default settings on the PC under examination are visible:

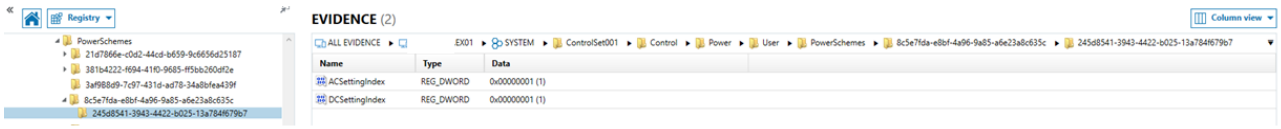


Fig. 11. Default high performance energy plan of the PC under test

4. Analysis to find if the password was required to enter the system when the screen was reactivated

We proceeded to verify whether, after the period of screen deactivation (15 minutes – 900 seconds), the password for accessing the system was requested.

For Windows 10 Pro system, the password prompt setting, after screen off period, is done by default after a password is associated with an account. All this occurs through the following option accessible from the *Settings / Account / Access options* menu (Fig. 12):

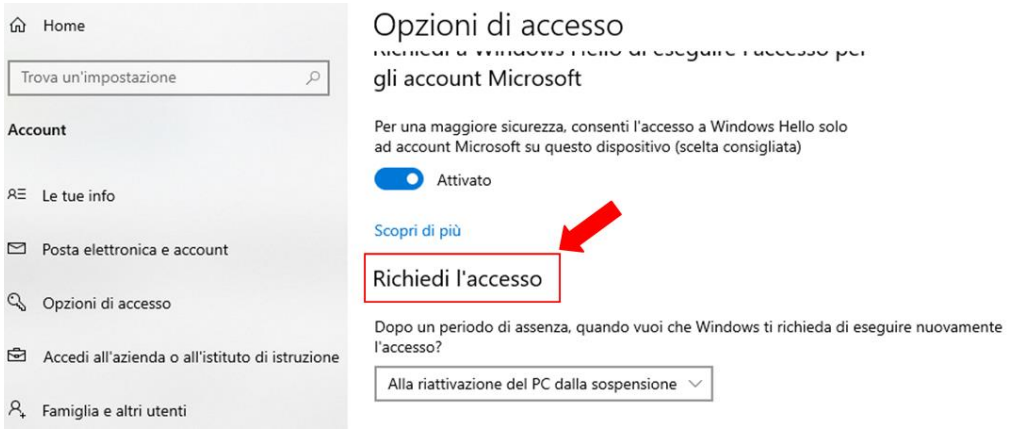


Fig. 12. Default high performance energy plan of the PC under examination

The drop-down menu in access request has two options (Fig. 12): the default one called "When the PC wakes up from suspension" and the one with "Never".

The first option, a few seconds after the screen has been deactivated, requires the user password upon reactivation, while the second option "Never", even after the screen has been deactivated, no longer shows the password entry screen. Also, in this case it is possible to verify this setting, by analyzing the Windows system registry.

In fact, if the default settings for the energy plan in argument, identified by the key **8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c**, are not changed, no subkey will be created; otherwise if the default settings are modified the subkey **0e796bdb-100d-47d6- a2d5-f7d2daa51f51** will be created. This subkey, after its creation, will remain present in the system registry unless default settings are restored. Fig. 13 shows the case in which default settings have been modified: the subkey **0e796bdb-100d-47d6- a2d5-f7d2daa51f51** is created with the "ACSettingIndex" item set to "Never" (value equal to 0); if it had been set to 1 it would have had the meaning of “When PC wakes up from sleep”.

Fig. 13. Access request set to “Never”

The settings on the PC under examination were the default ones (When PC wakes up from sleep) due to the absence of the aforementioned additional subkey. Therefore, password entry was required after the screen off period.

To know the default values relating to the access request, you need to examine the values of the key **0e796bdb-100d-47d6-a2d5-f7d2daa51f51** which we find under the path "Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\User\PowerSchemes".

The values found in the PC under examination extrapolated by the analysis software are reported:

– in the path: Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\User\PowerSchemes\8c5e7fda-e8bf-4a96-9a85-a6e23a8c635c" only one subkey named **245d8541-3943-4422-b025-13a784f679b7** is visible without the aforementioned **0e796bdb-100d-47d6-a2d5-f7d2daa51f51**. All to demonstrate that no changes have been made;

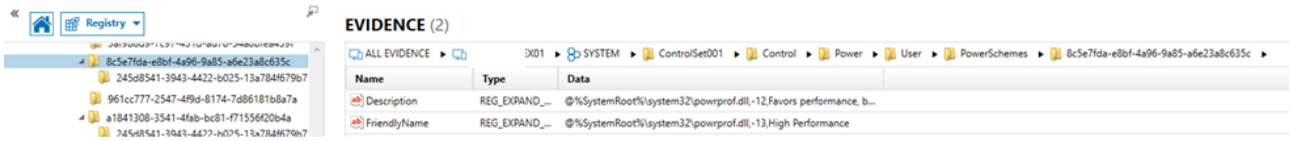


Fig. 14. Password request when target PC wakes up

– to find out the default values, analyze the path "Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Control\Power\PowerSettings\0e796bdb-100d-47d6-a2d5-f7d2daa51f51" where for the settings relating to the **high-performance** energy plan, it is set the value (1) of ACSettingIndex and then Request access/When PC wakes up from suspension:

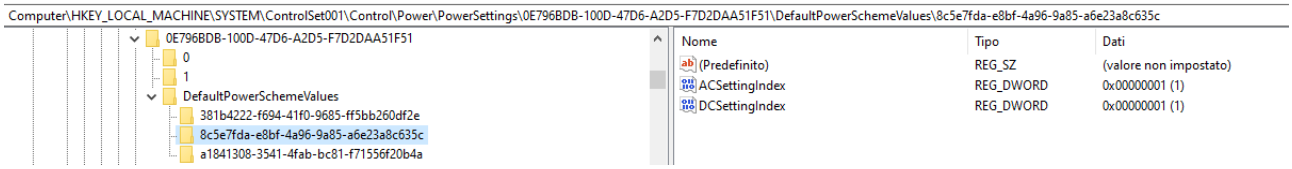


Fig. 15. Default setting for the password request when restarting the PC

5. Conclusion

In conclusion, it has been demonstrated that the possible execution of operations on the PC under examination could only be carried out by those in possession of the access password of the account in argument.

References

[1]. https://docs.microsoft.com/it-it/windows/win32/sysinfo/structure-of-the-registry.
[2]. Guidelines for identification, collection, acquisition and preservation of digital evidence. ISO/IEC 27037:2012.
[3]. Guidelines for the analysis and interpretation of digital evidence. ISO/IEC 27042:2015.

Easy to Remember, Hard to Guess: A Password Generation Tool for the Digital Age

Ioana-Ilona BRĂSLAȘU, Andrei-Daniel ANDRONESCU, Dumitru-Iulian NĂSTAC
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
ioanabraslasu2000@gmail.com, andronescu.andreidaniel@gmail.com, iulian.nastac@upb.ro

Abstract
A brute force attack is a common method used by cybercriminals to gain unauthorized access to user accounts. It is essential for individuals and organizations to take proactive measures to protect themselves from such attacks. One way to do this is by improving their knowledge of cybersecurity and implementing measures to safeguard their online presence. Using programming languages like Python and web-frameworks like Django, websites can be developed to help individuals generate secure and memorable passwords that align with the latest password security standards. This can help anyone who wants to improve their password security, irrespective of whether they have been a victim of a cyber-attack or not.

Index terms: hackers, Python, secure password, website, memorable

1. Introduction

The Internet dependency has increased significantly in the post-pandemic era and therefore people create frequently accounts using their credentials when they perform online shopping, work remotely or in e-learning. As a result, the most crucial factor in this context is the password. With the large number of accounts created online, it can be challenging to remember every password created for each website and, in particular, to generate new password every time, as it is recommended by the well-known password-manager tool of Google [1]. People tend to ignore such suggestions and instead, they use identical passwords across multiple platforms because it is more comfortable in this way. What is more, the risks of using personal information based on birthdays, names or addresses in passwords tend to be ignored even if any information posted on social media can be used by a hacker as they can craft a social engineering attack [2].

Since password authentication is the cheapest mechanism to access a system, passwords turn into the most vulnerable component in this equation [3] and this issue was a source of inspiration. In this paper, an easy-to-remember password generator is proposed and it is presented as a website. The user must enter two words that represents him, excluding personal information, such as: favorite sports team, city, season, food, activity or animal and a number. After inputting the necessary requirements, the implemented code modifies the words by converting some lower-case letters into upper-case or changing certain letters into symbols or numbers. These modified words are then combined with the entered number to create a secure character combination that is resistant to cyber-attacks.

2. What does a secure password need?

Generally, each website offers the user general information regarding what a secure password should contain, without requiring additional research on security risks. This is just an example from a university's website [4] of suggestions that usually pop up when creating a password:

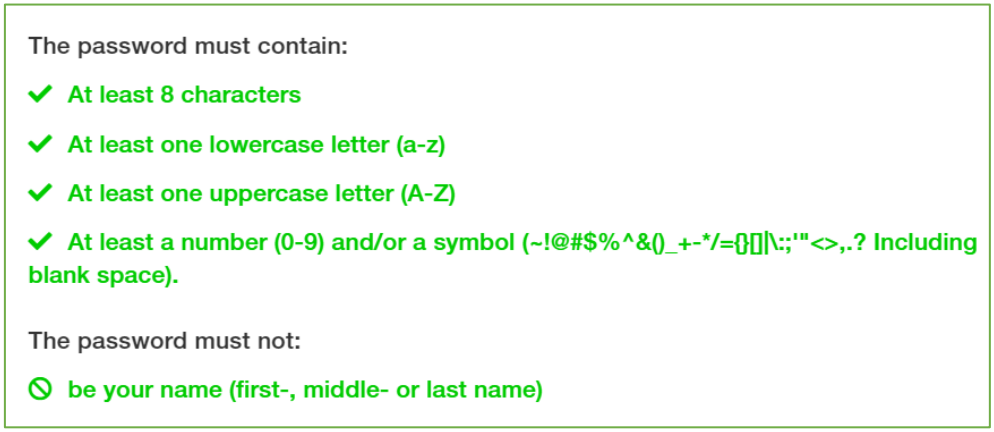


Fig. 1. A website's recommendation for a secure password

Also, Google Password Manager always suggests strong passwords, which can be stored in this tool, but it can be extremely difficult to remember since it represents just a random combination of characters, but very solid at the same time.

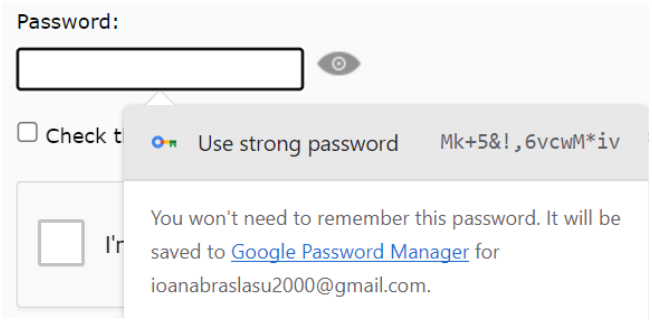


Fig. 2. Google Password Manager suggestion

2.1. General standards for a secure password

To ensure the safety and security of our online presence, it is important to follow certain guidelines and standards when creating and managing passwords. By following this set of rules, we can greatly reduce the risk of our personal information being compromised by hackers and cybercriminals.

- **Avoid using personal information**: ExpressVPN's study on the most common passwords around the world showed that 42% of people use their first name in their passwords, while 43% of them use their birth date [5], data that ca be easily guessed by unauthorized people who can search on your social media for such details. It is also advisable to refrain from using personal information, such as the names of pets or children, as well as addresses.
- **The longer the password the better**: According to the Center for Internet Security (CIS) [6], length is the most important aspect of a good password. Passwords that are more than 8 characters are statistically harder to guess than shorter ones. When a password cracker has more characters to fill to guess the correct password, it becomes exponentially less likely to get it right.

- **Mixed symbols, numbers and caps:** A mix of characters increases the number of possible combinations and makes it more difficult for an attacker to guess or crack the password. For instance, a password such as "password123" can be easily guessed, while a password such as "Pa$$w0rd!23" is more complex and difficult to guess. The use of numbers and capitalization also adds an extra layer of complexity to the password [7].
- **A password should not be shared with any other account:** If a shared password is compromised, it can potentially give an attacker access to all accounts associated with that password. Additionally, sharing them with other accounts can lead to a loss of control over personal information such as bank accounts, as it can be difficult to track who has access to what information. In some cases, sharing passwords may even be a violation of the terms of service of certain websites or applications.
- **A password should not contain any consecutive letters or numbers**
- **Easy to remember, hard to guess:** the users' real issue
- **The password should be confidential:** It is important to keep our passwords confidential because they are the primary means of securing personal and sensitive information online. If passwords are not kept confidential, they can be easily compromised, giving unauthorized individuals access to our accounts and personal information.

Not only is the complexity of a password essential, but also the safety of the 'place' where it is stored. Cybersecurity experts from the U.S. government (CISA) have recommended using password managers [8]. A password manager is a software application that helps users generate, store, and manage their passwords for various online accounts. It typically stores passwords in an encrypted database that is protected by a master password or passphrase that must be remembered by the user. Such tools can also help users generate strong, unique passwords for each account, which reduces the risk of a password being compromised.

Eventually, to add extra safety to the password and to the sensitive information, individuals can opt for two-authentication security measure. It can take different forms such as SMS codes, authentication as an option to increase security for their users.

2.2. Most used combinations for passwords

When it comes to creating an online account, many of us would rather remember a code than create an unbreakable password. Defending our position, it can be challenging to keep track of numerous login credentials. According to technology expert Burton Kelso [9], it is human nature to fall into a predictable routine when it comes to our passwords list.

To make matters worse, the analysis reveals that 64 percent of breached passwords were used for at least two accounts [10] which can lead to a domino effect in the case when a hacker succeeds in guessing a password because several accounts are compromised as well.

According to NordPass.com [11], the most common passwords used in 2022 are: *password*, *123456*, *123456789*, *guest*, *123123*, *col123456*, *000000*, *querty*.

Security.org created a website [12] that check how secure a password is and in how many years or seconds they can be guessed. The passwords listed above can be breached in less than one second. It is an interesting tool that makes the user understand how technology evolves in terms of decrypting passwords.

3. Real cases of weak passwords guessing and the consequences

Weak passwords are a common cause of data breaches and cyber-attacks, leading to serious consequences for individuals, businesses, and organizations. In 2019, a massive data breach at Capital

One Bank [13] exposed the personal information of over 100 million customers, due to a vulnerability in the company's cloud infrastructure caused by a weak password. In another case, on January 17th 2022, approximately 500 individuals' cryptocurrency wallets were targeted in an attack that resulted in the theft of approximately $18 million worth of Bitcoin, $15 million worth of Ethereum, and other cryptocurrencies [14]. The hackers were able to gain access to the wallets by bypassing two-factor authentication, demonstrating the vulnerability of this security measure. This incident highlights the importance of using a password manager to store and manage strong, unique passwords for online accounts, as it can help protect against Brute force attacks, where attackers use automated tools to try many possible passwords, can be successful against weak passwords that are easy to guess, such as "123456" or "password." Similarly, phishing attacks, where attackers trick users into revealing their passwords through fraudulent emails or websites, can be effective against users with weak passwords who may be more easily fooled. Moreover, dictionary attacks, where attackers use pre-computed lists of common passwords, can be successful against weak passwords that are commonly used, such as "qwerty" or "letmein".

All these examples illustrate the importance of using strong, unique passwords and taking other measures to protect against cyber-attacks. A solution to this issue would be to make those passwords as memorable as possible to be more appealing to not only individuals, but also big companies.

4. Description of the Website Implementation

Since passwords are mainly used in different Internet browsers, implementation of a website that comes in handy in securing the online experience is proposed. In this process, both Python programming language and web development languages were used in order to create a fully functional website that is called *My Secure Password*. Here, the user can enter two words that represents him and also one number to prevent the risk of decryption, elements that eventually will build a robust password.

4.1. Backend Part of the Website

For creating the server where the website will be running, Django framework is used. Django is a high-level Python web framework that encourages rapid development and clean, pragmatic design. The efficiency of rendering HTML pages was demonstrated by including CSS and JavaScript files in the HTML templates, which allowed for customization of the website's styling and incorporation of interactive features like a copy-to-clipboard button.

In this framework's inputs, Python functions were created to perform the most important part of the project, the creation of the secure password. Those are:

- upper_vowels ()
- secure_letters ()

In designing these functions, two widely recognized criteria were used for enhancing password security, namely the substitution of lower-case characters with their corresponding upper-case letters and the use of special characters as substitutes for some of the alphabetic letters. In practice, individuals are typically more responsive to the manipulation of vowel characters in a word, hence, the decision was made to replace only lower-case vowels with their respective upper-case counterparts. This approach reduces the potential for confusion and facilitates memorization of the password. When it comes to replacing some letters with special characters, substituting *a* with '@' or *i* with '!' will impact the user visually. There are other character substitution ideas:

```python
def secure_letters(string):
    new_word = ''
    target_letters = ["a", "o", "e", "s", "i", "b"]
    new_letters = ["@", "0", "€", "$", "!", "ó"]

    for letter in string:
        if letter in target_letters:
            new_word += new_letters[target_letters.index(letter)]
        else:
            new_word += letter

    return new_word
```

Fig. 3. Special Characters replacement in Python Code

As the most important function, the password generation one, has two words and one number as input, it was decided to randomly allocate those letter-manipulation functions to the input words for making the code as general as possible. Next, the processed words plus the number are combined into a complex union of characters. If the user typed words that consists of a word or less, an error message will pop up and if the combination of letters has less than twelve letters, random characters such as: '*',' #', '- 'or '.'will be placed between the elements until the desired number of characters is reached. Also, if there are any spaces in the input words, they will be also replaced with special characters.

4.2. Frontend Part of the Website

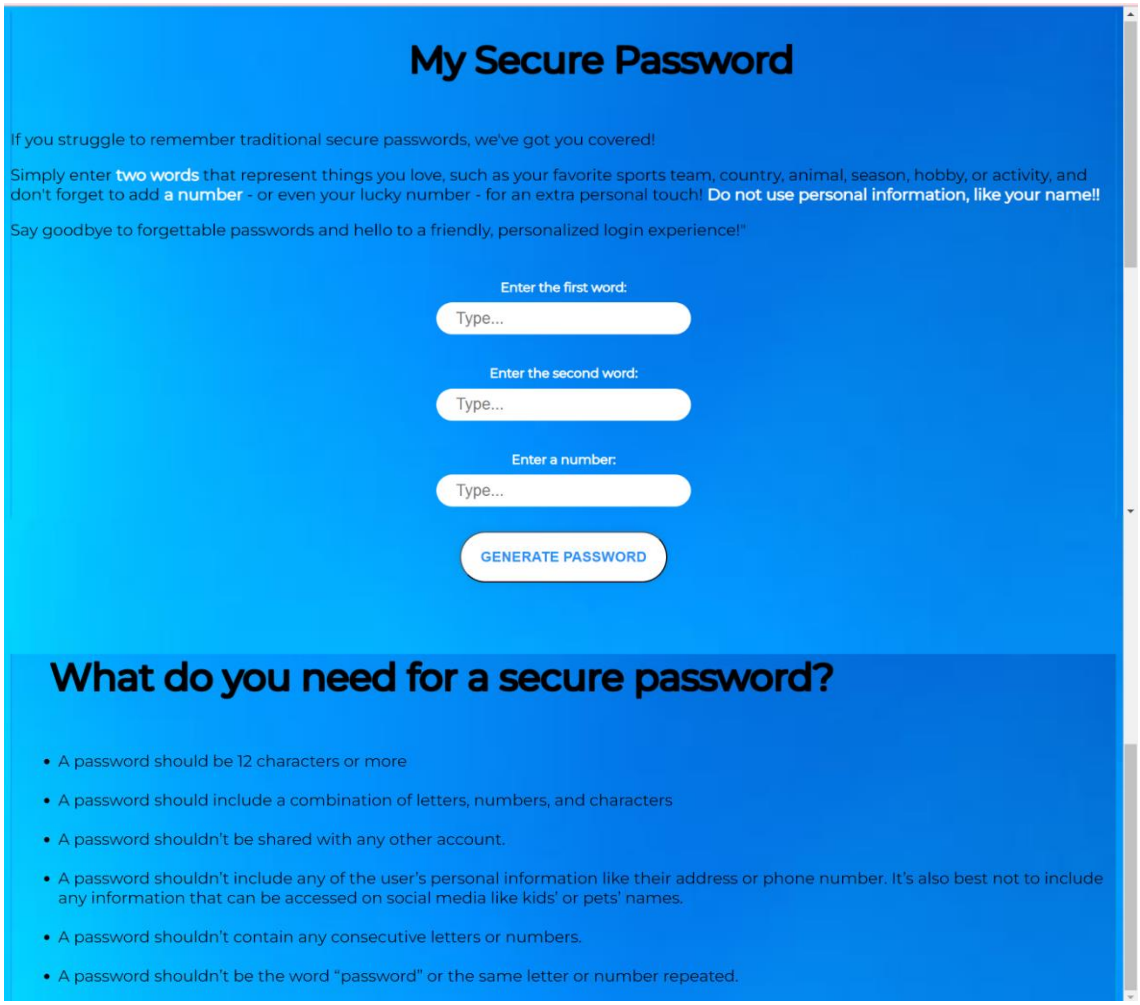


Fig. 4. The website's design

To ensure a seamless design for the website, widely recognized languages such as HTML, CSS, and JavaScript were used. The user receives clearly what is needed for generating a secure and memorable password and additionally some recommendations to protect personal data online. After pressing the 'Generate Password' button, an input will appear with the result, along with the copy-to-clipboard button.

Towards the end of the webpage, a section has been included that offers a concise summary of necessary recommendations and principles for designing a secure password based on standard practices.

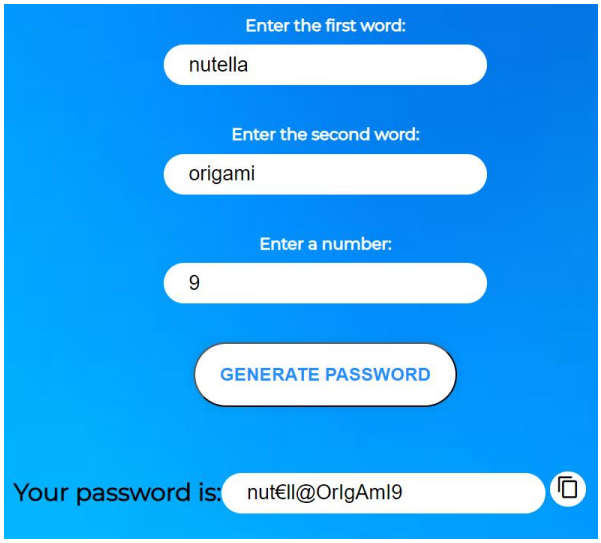


Fig. 5. Example of a Secure Password

5. Conclusion

In conclusion, the proposed website offers an efficient solution to the common problem of weak passwords. By generating strong and memorable passwords that comply with the latest password security standards, the website can significantly reduce the risk of cyber-attacks and prevent users from reusing the same password across different websites. In today's digital world, where cyber threats are constantly evolving, it is imperative for individuals to take proactive measures to ensure the security of their personal information. The website's approach to generating secure passwords is a simple yet effective way for users to improve their online security and safeguard themselves against potential data breaches. By implementing this method, users can rest assured that their accounts are well-protected, and their sensitive information is kept secure. Overall, the website's contribution to enhancing password security serves as a crucial step towards improving the overall cybersecurity landscape.

When it comes to improving the presented website, displaying errors accordingly on the interface is on the list for future plans. In present, only the error when the user enters a word of 2 characters or less appears on the screen. Also, the error display is not well formatted and many other error cases should pop up. For example, when the user enters consecutive numbers or letters and when words such as passwords and user are chosen should not give the permission for creating a strong password.

References

[1]. Google. (n.d.). Manage your Google Account password. [Online]. Available: https://passwords.google.com/options?ep=1. [Accessed: Apr. 19, 2023].

[2]. J. Miller and Z. Snyder, "Innocuous Facebook Quizzes: Attacker Intel Goldmines," ZeroFOX, 2016. [Online]. Available: https://www.zerofox.com/blog/innocuous-facebook-quizzes-attacker-intel-goldmines/. [Accessed: May 06, 2023]

[3]. Sadat, S. E., Hedayathllah, H., & Ahamadzai, N. (2023). Highly Secure and Easy to Remember Password-Based Authentication Approach. Journal for Research in Applied Sciences and Biotechnology, 2(1), 18-25. DOI: 10.55544/jrasb.2.1.18. [Accessed: Apr. 19, 2023].

[4]. University Admissions. (n.d.). Log in. [Online]. Available: https://www.university admissions.se/intl/login. [Accessed: Mar. 20, 2023].

[5]. HackRead. (2019, Feb. 8). People use their names as passwords: Study. [Online]. Available: https://www.hackread.com/people-use-their-names-passwords-study/. [Accessed: Apr. 20, 2023].

[6]. Georgetown University. (2020, Oct. 27). Password size matters. [Online]. Available: https://security.georgetown.edu/csam-2020/password-size-matters/. [Accessed: Apr. 20, 2023].

[7]. Google. (n.d.). Create a strong password. [Online]. Available: https://support.google.com/accounts/answer/32040?hl=en. [Accessed: Apr. 23, 2023].

[8]. Cybersecurity and Infrastructure Security Agency. (2022, Mar. 10). Choosing and protecting passwords. [Online]. Available: https://www.cisa.gov/news-events/news/choosing-and-protecting-passwords. [Accessed: Apr. 23, 2023].

[9]. Reader's Digest. (n.d.). Passwords hackers can guess first. [Online]. Available: https://www.rd.com/article/passwords-hackers-guess-first/. [Accessed: Apr. 23, 2023].

[10]. Techzine. (2022, Mar. 2). Most people still use the same password for multiple accounts. [Online]. Available: https://www.techzine.eu/news/security/74094/most-people-still-use-the-same-password-for-multiple-accounts/. [Accessed: Apr. 23, 2023].

[11]. NordPass. (n.d.). Most common passwords list. [Online]. Available: https://nordpass.com/most-common-passwords-list/. [Accessed: Apr. 23, 2023].

[12]. Security.org. (2022, Mar. 18). How secure is my password? [Online]. Available: https://www.security.org/how-secure-is-my-password/. [Accessed: Apr. 23, 2023].

[13]. Capital One, "Facts 2019," Capital One, [Online]. Available: https://www.capitalone.com /digital/facts2019/. [Accessed: Apr. 24, 2023].

[14]. ERMProtect, "Top 10 Data Breaches So Far in 2022," ERMProtect, [Online]. Available: https://ermprotect.com/blog/top-10-data-breaches-so-far-in-2022/. [Accessed: Apr. 24, 2023].

Artificial Intelligence and its Impact on Cybercrime

Carla LOZONSCHI, Irina BAKHAYA, PhD
"Alexandru Ioan Cuza" Police Academy, Bucharest, Romania
carlalozonschi@gmail.com, i.bakhaya@gmail.com

Abstract

It is well known that technology is becoming increasingly prevalent among us, and that it is evolving at a quick pace. We're hearing more and more about artificial intelligence and how it affects our lives. Opinions on AI split the globe into two camps. Therefore, we choose to discuss what Artificial Intelligence is and how it marks our lives. Is it good to employ artificial intelligence? If so, how far should this be taken? Can it be used in a bad way? Sure, but this may also play a significant role in preventing and combatting cybercrime. All of these topics will be addressed in the next article.

Index terms: Artificial Intelligence, cybercrime, cybersecurity, deepfakes, bots

1. Introduction

Artificial intelligence is referred to as intelligent machines. This is in contrast to humans' and animals' innate intelligence. Machines use Artificial Intelligence to execute operations including learning, planning, reasoning, and problem solving. The most notable aspect of Artificial intellect is the replication of human intellect by machines. It is most likely the most rapidly developing advancement in the world of technology and innovation. Furthermore, many experts believe AI has the potential to tackle huge problems and crisis circumstances.

2. What is AI and how it marks our lives?

We have all heard about artificial intelligence in recent years, but few are aware that it has been developed since the twentieth century, i.e. after World War II and the name itself was coined in 1956. AI currently encompasses a huge variety of subfields, ranging from the general (learning and perception) to the specific, such as playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases. AI is relevant to any intellectual task; it is truly a universal field.

Artificial Intelligence can be seen from several perspectives [1]. The definitions on top are concerned with thought processes and reasoning, whereas the ones on the bottom address behavior. The definitions on the left measure success in terms of fidelity to human performance, whereas the ones on the right measure against an ideal performance measure, called rationality.

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think ... machines with minds, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

ACTING HUMANLY

The Turing Test, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of intelligence. A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer. Chapter 26 discusses the details of the test and whether a computer would really be intelligent if it passed. For now, we note that programming a computer to pass a rigorously applied test provides plenty to work on. The computer would need to possess the following capabilities:

- natural language processing to enable it to communicate successfully in English;
- knowledge representation to store what it knows or hears;
- automated reasoning to use the stored information to answer questions and to draw new conclusions;
- machine learning to adapt to new circumstances and to detect and extrapolate patterns.

Turing's test deliberately avoided direct physical interaction between the interrogator and the computer, because physical simulation of a person is unnecessary for intelligence. However, the so-called total Turing Test includes a video signal so that the interrogator can test the subject's perceptual abilities, as well as the opportunity for the interrogator to pass physical objects "through the hatch." To pass the total Turing Test, the computer will need:

- computer vision to perceive objects, and
- robotics to manipulate objects and move about.

These six disciplines compose most of AI, and Turing deserves credit for designing a test that remains relevant 60 years later. Yet AI researchers have devoted little effort to passing the Turing Test, believing that it is more important to study the underlying principles of intelligence than to duplicate an exemplar.

THINKING HUMANLY

The interdisciplinary field of cognitive science brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind. Cognitive science is a fascinating field in itself, worthy of several textbooks and at least one encyclopedia. In the early days of AI, there was often confusion between the approaches: an author would argue that an algorithm performs well on a task and that it is therefore a good model of human performance, or vice versa. Modern authors separate the two kinds of claims, allowing both AI and cognitive science to develop more rapidly. Computer vision, which incorporates neurophysiological evidence into computational models, is an example of this.

THINKING RATIONALLY

The Greek philosopher Aristotle was one of the first to attempt to codify "right thinking," that is, irrefutable reasoning processes. His syllogisms provided patterns for argument structures that

always yielded correct conclusions when given correct premises. This study initiated the field of logic, which developed a precise notation for statements about all kinds of objects in the world and the relations among them. By 1965, programs existed that could, in principle, solve any solvable problem described in logical notation. However, there are two main obstacles to this approach: it is not easy to take informal knowledge and state it in the formal terms required by logical notation, particularly when the knowledge is less than 100% certain, and there is a big difference between solving a problem "in principle" and solving it in practice. These obstacles appear first in the logicist tradition.

ACTING RATIONALLY

The "laws of thought" approach to AI emphasizes correct inferences, but there are also ways of acting rationally that do not involve inference. The rational-agent approach is more general than the "laws of thought" approach, as correct inference is just one of several possible mechanisms for achieving rationality. It has two advantages over the other approaches: it is more general and is more general than the "laws of thought" approach, and it is more general than the "laws of thought" approach because correct inference is just one of several possible mechanisms for achieving rationality. The rational-agent approach has two advantages over the other approaches: it is more.

The standard of rationality is more amenable to scientific development than approaches based on human behavior or human thought. It is mathematically well defined and completely general, and can be "unpacked" to generate agent designs that provably achieve it. Human behavior, on the other hand, is well adapted for one specific environment and is defined by the sum total of all the things that humans do. Despite the apparent simplicity of the problem, an enormous variety of issues come up when we try to solve it. To simplify the problem, perfect rationality is a good starting point for analysis, as it simplifies the problem and provides the appropriate setting for most of the foundational material in the field.

What is the significance of discussing artificial intelligence? We can all see how much technology affects our life, from the most basic to the most complicated activities a person may engage in. For this reason, we must understand how and when to apply Artificial Intelligence. If we don't know how to utilize it, it has a lot of power. Regarding cyber crime, it can be used in both directions, either to increase the rate of criminal crime, or by preventing and combating it.

How can artificial intelligence be used for negative purposes?

Cybercriminals are already using AI to make their attacks more effective and far-reaching. It will only grow more widespread.

In 2020, a study by European police agency Europol and security provider Trend Micro, identified how cybercriminals are already using AI to make their attacks more effective, and the many ways AI will power cybercrime in future [2].

While AI and ML can support businesses, critical infrastructures, and industries as well as help solve some of society's biggest challenges (including the Covid-19 pandemic), these technologies can also enable a wide range of digital, physical, and political threats to surface. For enterprises and individual users alike to remain protected from malicious actors who are out to misuse and abuse AI, the risks and potential malicious exploitations of AI systems need to be identified and understood.

DEEPFAKES

Deepfakes are a common kind of AI abuse in which audio and visual information is created or altered to seem authentic using AI algorithms. An purported deepfake video that purports to show a Malaysian political assistant engaging in sexual relations with a cabinet minister is one illustration of this, and it demands for the cabinet member to be looked into for potential wrongdoing. Another instance is a UK-based energy company that was tricked into sending over 200,000 British pounds (about $260,000 as of this writing) to a Hungarian bank account after a malevolent person impersonated the voice of the company's CEO using deepfake audio technology. Since Buzzfeed

collaborated with actor and filmmaker Jordan Peele on Deepfakes, it can serve as a beneficial tool for teaching people about their potential abuses.

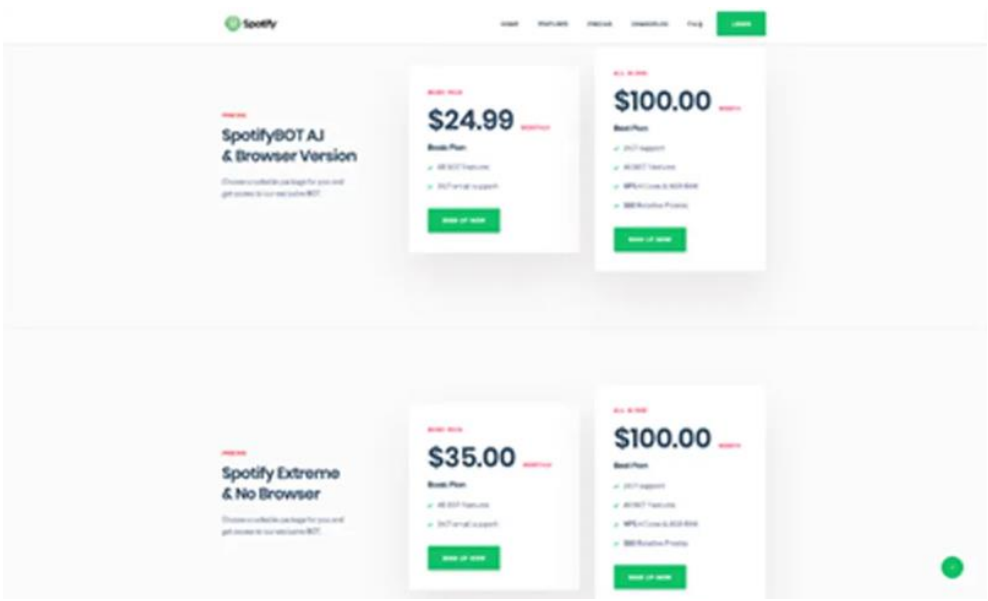
AI-SUPPORTED PASSWORD GUESSING

Cybercriminals are using ML to enhance password guessing algorithms. It is already possible to effectively determine the password that matches to the password hash using more conventional methods like HashCat and John the Ripper, which compare several permutations to the password hash. However, thieves would be able to examine big password datasets and develop password variants that suit the statistical distribution using neural networks and generative adversarial networks (GANs). This will eventually result in more precise and profitable password guesses as well as more opportunities for profit.

HUMAN IMPERSONATION ON SOCIAL NETWORKING PLATFORMS

AI is being used by cybercriminals to mimic human behavior. By imitating human-like usage patterns, they may, for instance, easily fool bot detection algorithms on social media networks like Spotify. Cybercriminals may then monetise the infected system to produce phony streams and traffic for a particular artist using this AI-supported impersonation.

An AI-supported Spotify [2] bot on a forum called nulled[.]to claims to have the capability to mimic several Spotify users simultaneously. It employs a number of proxies in order to evade discovery. This bot raises the number of streams (and therefore, revenue) for particular songs. It also produces playlists with other songs that reflect human-like musical preferences rather than playlists with random tracks, since the latter might suggest bot-like activity, to further elude discovery.



In the future, I envision criminals using AI in a variety of ways. It is quite possible that cybercriminals will use AI to increase the breadth and size of their assaults, avoid detection, and abuse AI as both an attack route and an attack surface.

Criminals will employ AI to carry out nefarious operations such as social engineering to victimize companies. Cybercriminals may utilize AI to automate the earliest phases of an attack by creating content, increase business information collecting, and accelerate the detection rate of potential victims and business operations. This can lead to faster and more accurate business fraud via different tactics such as phishing and business email compromise (BEC) schemes.

Artificial intelligence (AI) is playing a massive role in cyber attacks and is proving both a "double-edged sword" and a "huge challenge," according to NATO.

"Artificial intelligence allows defenders to scan networks more automatically, and fend off attacks rather than doing it manually. But the other way around, of course, it's the same game," David van Weel, NATO's Assistant Secretary-General for Emerging Security Challenges, told reporters earlier this month [3].

Since the Ukraine war, cyber assaults on national infrastructures and commercial organizations have increased tremendously and become a focus point. This year, NATO stated that a cyber assault on any of its member nations might trigger Article 5, which specifies that an attack on one member is considered an attack on all of them and may result in a collective reaction.

AI-based technologies may be used to better detect and fight against threats, but hackers can also utilize the technology for more sporadic attacks that are more difficult to defend against since there are so many of them at the same time.

AI cyber attacks can be used not just to shut down infrastructure but also to exploit information, said Alberto Domingo, technical director of cyberspace at NATO Allied Command Transformation [3].

"I think AI is a critical threat. The number of attacks is increasing exponentially all the time," he told Euronews Next, adding that at the moment the world is simply "living with these attacks" and needs more cybersecurity rules.

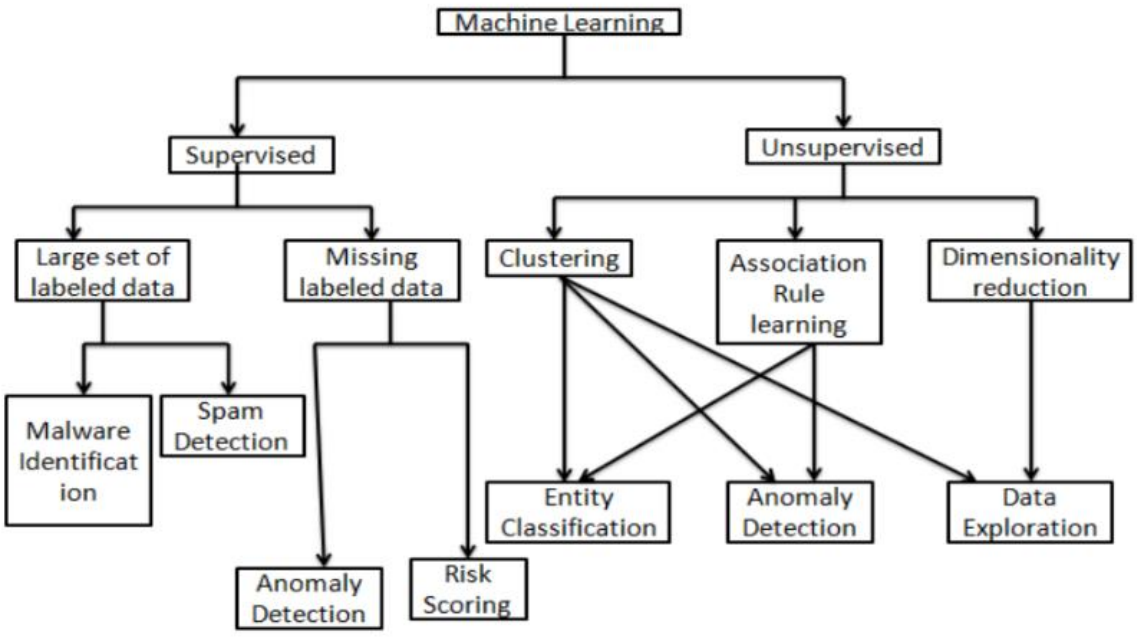
"We are not yet at a stage where we identify that this is simply not acceptable. These behaviours cannot be allowed in cyberspace," he said.

"It shows you that we still don't have a collective common approach to react to those things, but those things are simply not acceptable".

As a result, while Artificial Intelligence may mark crimes negatively by easing the labor of numerous hackers, it can also be utilized to prevent and combat cybercrime.

Advantages of AI in Cybersecurity

Because cyber security risks are always changing and evolving, a rapid and automated reaction is essential. As a result, machine learning techniques, particularly deep learning, that do not often require prior expertise or reliance on past expert classifications may be very useful in the application of cyber security AI systems. The research [4] Security examined the efficiency of machine learning methods for cyber security reasons. This study involves the use of machine learning technologies to detect intrusions, spam, and malware.



AI has several benefits and uses in a range of fields, one of which is cybersecurity. With today's rapidly developing cyberattacks and rapid device proliferation, AI and machine learning can assist in keeping up with cybercriminals, automating threat detection, and responding more efficiently than traditional software-driven or manual procedures.

Using complicated algorithms, AI systems are being trained to recognize malware, perform pattern recognition, and detect even the smallest features of malware or ransomware assaults before they reach the system. AI may deliver more predictive intelligence using natural language processing by scanning through articles, news, and studies on cyber dangers and selecting material on its own. According to Tech Republic, a mid-sized corporation receives alerts for over 200,000 cyber events per day. This volume of assaults would overwhelm a typical company's security personnel. As a result, some of these attacks may go undetected, causing considerable network damage. Security personnel require considerable assistance from intelligent machines and current technologies such as AI to function efficiently and safeguard their companies from cyber attacks.

DETECTING NEW THREATS

AI may be used to detect cyber dangers and potentially harmful behavior. Traditional software systems simply cannot keep up with the huge volume of new viruses developed each week, thus this is an area where AI may be really useful.

AI systems are being trained to identify malware, run pattern recognition, and detect even the smallest behaviors of malware or ransomware assaults before they reach the system using advanced algorithms. AI enables higher predictive intelligence through natural language processing, which curates material on its own by scraping articles, news, and cyber threat research. This can provide novel abnormalities, cyberattacks, and protection techniques. After all, hackers, like everyone else, follow trends, so what's popular with them changes all the time.

AI-based cybersecurity solutions may give the most up-to-date knowledge about global and industry-specific threats, allowing you to make more informed prioritizing decisions based not just on what could be used to attack your systems, but also on what is most likely to be used to attack your systems.

BATTLING BOTS

Bots account for a significant portion of internet traffic nowadays, and they may be deadly. Bots may be a serious threat, from account takeovers using stolen passwords to fraudulent account creation and data theft.

Manual answers alone will not suffice to combat automated threats. AI and machine learning assist in developing a comprehensive knowledge of website traffic and distinguishing between good bots (such as search engine crawlers), malicious bots, and humans.

AI helps us to evaluate massive amounts of data and allows cybersecurity teams to modify their approach to an ever-changing scenario.

“By looking at behavioral patterns, businesses will get answers to the questions ‘what does an average user journey look like’ and ‘what does a risky unusual journey look like’. From here, we can unpick the intent of their website traffic, getting and staying ahead of the bad bots,” explains Mark Greenwood, Chief Technical Architect & Head of Data Science at Netacea [5].

BREACH RISK PREDICTION

AI systems assist in determining the IT asset inventory, which is a precise and thorough record of all devices, users, and apps with varying levels of access to various systems.

Taking into account the asset inventory and threat exposure (described above), AI-based systems can forecast how and where you are most likely to be hacked, allowing you to plan and allocate resources to the most vulnerable regions.

AI-based analysis provides predictive insights that allow you to set and optimize policies and procedures to strengthen your cyber resilience.

What Cybersecurity Executives Think About AI?

Capgemini Research Institute analyzed the role of AI in cybersecurity and their report titled Reinventing Cybersecurity with Artificial Intelligence [6].

Respondents to a poll (850 executives from cybersecurity, IT information security, and IT operations from ten countries) agree that AI-enabled response is needed due to cybercriminals using AI technology to execute assaults.

The following are some of the report's primary takeaways:

- According to three out of four CEOs polled, AI enables their firm to respond to breaches more quickly.
- 69% of businesses believe AI is required to respond to threats.
- According to three out of every five companies, utilizing AI enhances the accuracy and efficiency of cyber analysts.
- AI gives better answers to an organization's cybersecurity demands as networks get larger and data grows more complicated. Simply said, humans are incapable of dealing with the increasing complexities on their own, and the employment of AI will become unavoidable sooner or later.

To summarize, artificial intelligence is well on its way to dominating the planet. While this may appear to be a frightening and worrisome future, there is no major or reliable evidence to suggest that the usage, application, implementation, and assimilation of artificial intelligence would harm the planet in any manner. So far, technology has only worked to better existing problems and living conditions. Artificial intelligence has helped both those who developed it and those who utilize it. Unlike in Hollywood sci-fi movies, the planet is not vulnerable to an invasion by renegade robots. For those who are still skeptical, experts believe that introducing one case where a hypothesis does not hold true can invalidate it.

Artificial Intelligence has thus far shown to be nothing but positive. Artificial intelligence is on its approach to being a game changer in ethics, social welfare, healthcare, and workforce. Because these are the most important components of life on a daily basis, technological developments in these domains will matter far more than in any other. Improving ethical standards, benefiting societal welfare through regulated surroundings, greatly increasing healthcare, and transforming the workforce; artificial intelligence is the unanticipated but much needed miracle. Artificial Intelligence supremacy will substantially enhance the quality of living and reshape the globe.

References

[1]. P. Norvig, S. Russell, Artificial Intelligence: A Modern Approach, Global Edition, pg. 2.

[2]. https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/exploiting-ai-how-cybercriminals-misuse-abuse-ai-and-ml.

[3]. https://www.euronews.com/next/2022/12/26/ai-cyber-attacks-are-a-critical-threat-this-is-how-nato-is-countering-them#:~:text=Artificial%20intelligence%20(AI)%20is%20playing,rather%20than%20doing%20it%20manually.

[4]. https://www.researchgate.net/publication/364126309_Artificial_Intelligence_in_Cyber_Security.

[5]. https://www.computer.org/publications/tech-news/trends/the-use-of-artificial-intelligence-in-cybersecurity.

[6]. https://www.capgemini.com/wp-content/uploads/2019/07/AI-in-Cybersecurity_Report_20190711_V06.pdf.

Protecting Your E-Commerce Business. Analysis on Cyber Security Threats

Georgiana ANDREIANU
Faculty of Electronics, Telecommunications, and Information Technology,
University POLITEHNICA of Bucharest, Romania
georgiana.andreianu@stud.etti.upb.ro

Abstract

This paper aims to gather complete information needed for a retailer running an e-commerce website, with the intention of presenting some of the most common cyber security threats, such as malware, ransomware, SQL injection, and phishing, as well as ways to prevent them from happening and ways to manage the aftermath of a full-scale attack being carried out. Some best practices will be noted as a process that should always be considered when setting up an e-commerce business, and a risk management strategy will be outlined. An analysis will be performed on a data breach with one of the biggest number of victims in the last decade, which affected the Microsoft Exchange Servers.

Index terms: attack, cyber security, e-commerce, threat, vulnerability

1. Introduction

Over the last decade, the e-commerce industry has exponentially grown, especially during the time of Covid and the popularization of remote work. Going online to shop at your favorite store has never been easier. For the retailers, this is a true blessing, providing immense business opportunities from small scale retailers and service providers to whole large-scale industries. A high sales volume can be achieved much easier with the help of aggressive online marketing and the cost decrease of not having to rent a physical location. However, at the same time, it is a burden regarding the high risk of cyber security threats that e-commerce websites have.

With the evolution of websites and mobile applications, cyber-attacks have become more complex and more frequent in the past years. As we can see in figure 1, businesses have fallen victim to an average of 340 million attacks every year for the last seven years, according to data provided by [1]. Given the growth of online presence ever since the start of the pandemic, we can see a spike in the number of attacks since the year 2020, with the highest number of attacks being in 2021. Every e-commerce business is a hot target for cyber-crime, considering the amount of personal data and financial information that it collects. The attacks range from ones directed at the customer, such as phishing and malware to attacks that are directed more towards the server and website, such as SQL injection and Cross Site Scripting. A weakness of this kind that results in a breach can greatly affect both the customer and the business, in terms of cost of revenue and of customer trust.

For an e-commerce business to be successful, important cyber security aspects must be adhered to and respected. Defending against cyber security threats has become a job for strong electronic security, rather than exclusive human foresight and conservation. Furthermore, it is very important to have a strict protocol for risk management in the event of a breach that can affect both the customer and the business.

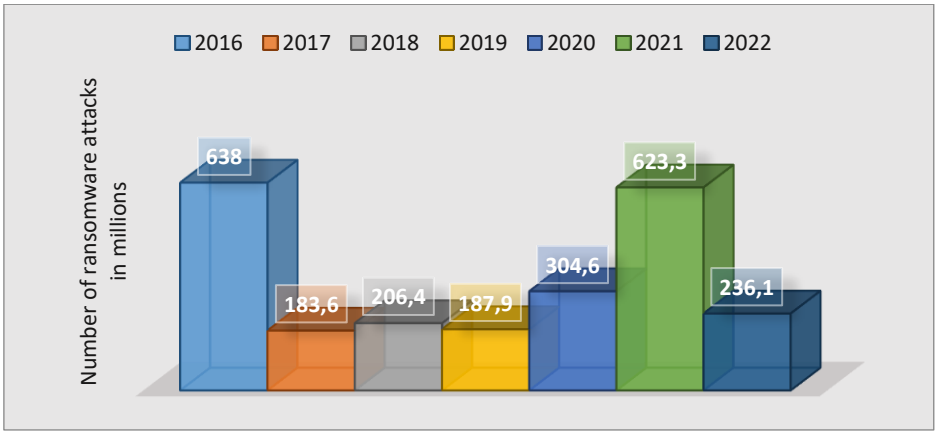


Fig. 1. Number of ransomware attacks in millions in the last 7 years

2. E-commerce Security

Electronic commerce is the business of buying, selling, or trading of goods or services using the internet as a means of communicating between the provider and the beneficiary. In other words, E-commerce is a transaction activity based on the media of information network. All business activities realized through, relying on, based on, or with the help of information network can be included in the category of electronic commerce [2].

In terms of engineering, e-commerce is a large system consisting of platform operators, providers, network consumers, suppliers, manufacturers, distributors, retailers, and many other groups collectively working together. Since e-commerce success as an industry depends on a system of entities, its security should also be based on the security of the whole system. The aim is to create a dynamic balance between all the components.

Cyber security protects computer systems from information disclosure, misdirection, damage, or theft of electronic data, software, or hardware [3]. In e-commerce, the main goal is to have a secure electronic behavior related to e-commerce activity. Even though businesses continuously invest in technologies to prepare against cyber threats, hackers are developing more and more complex ways to gain access to business systems and data, using cyber-attacks of different approaches and complexities, which will be reviewed in the following subchapters.

2.1. Phishing

Phishing is a type of social engineering and refers to methods used by attackers to trick victims - typically via email, text, or phone - into providing private information like passwords, account numbers, social security numbers, and more [4]. An attacker may try to send an email which seems to come from a reputable source, asking for information. This is an attempt at gathering personal and financial information about a person and using it in a malicious way. Websites are prone to becoming a casualty of this cyber security threat, with about 70% of the companies worldwide falling victims to phishing in 2020 [5].

2.2. Malware and Ransomware

Malware, or malicious software, is the most common cause of cyber-attacks. This software is programmed to handle control on your computer, and anything on it or entered into it, over to the cyber criminals without you even knowing it [6]. Malware is designed specifically for damaging, interrupting and obtaining unauthorized access to a system, while locking you out of your computer and denying you access to all important data.

Ransomware is a specific type of malware where the attacker holds access to sensitive files until the victim pays for them to be released. Both malware and ransomware are great threats to a

business, as it can cause inconveniences for the retailers, the employees, and the customers, all while having expensive costs for removing.

2.3. SQL Injection

Another important aspect of security is protecting databases. They can be put at risk by attacks such as SQL Injection, a method of cyber-attack where attackers insert malicious SQL commands into the forms or input fields of a web application to access or modify database information without authorization. This type of attack can have serious consequences, such as exposing sensitive data, losing, or altering data and even taking control of the system.

There are a wide range of vulnerabilities, attacks and SQL injection techniques that occur in different situations. Some common examples include obtaining hidden data, where a SQL query can be modified to return additional results; subverting application logic, where a query can be changed to interfere with application logic; database examination, where information about the version and structure of the database can be extracted.

2.4. Cross Site Scripting

Cross Site Scripting, or XSS, implicates inserting a malicious code, most commonly JavaScript, into a webpage. Basically, an attacker injects malicious executable scripts into the code of a trusted application or website by sending a malicious link to a user, who then accesses that link. An attacker can use XSS to send that malicious script to an unsuspecting user. The end user's browser has no way of knowing that the script is not to be trusted and will execute it. Because it believes the script comes from a trusted source, the malicious script can access any cookies, session tokens or other sensitive information kept by the browser and used with that site. These scripts can even rewrite the content of the HTML page. In this instance of a threat, the website itself is not in any danger, but the customers are at risk of being exposed to phishing, malware and other kinds of cyber security risks.

2.5. E-skimming

E-skimming is a method of stealing credit card information and personal data from payment processors on e-commerce sites. In this attack, hackers gain access to checkout pages and capture payment information as customers type it in real time [7]. The collected information is sent to an Internet-connected server with a domain that is controlled by the attacker, who will either sell the payment data or use it to make fraudulent purchases. E-skimming can result from XSS, phishing, brute force attacks or third-party compromise.

2.6. Brute Force Attacks

A brute force attack is a password-based attack where the cyber-criminal uses cracking tools to try all possible combinations of passwords to uncover valid passwords [8]. The aim of this attack is to duplicate a valid password for the online store’s administrator and gain otherwise unauthorized access. Another password-based attack is a dictionary attack, which aims for the same result, but instead of using scripts, the attacker manually tries all possible combinations of letters and numbers to guess the password. Given the time and labor required for the latter, a brute force attack is the more common approach.

3. Best Practices for Implementing Cyber-Security

Authentication and authorization are essential parts of basic security processes and are vital concepts that e-commerce business administrators should use to protect users' systems, information, and personal data. Although the two terms seem similar, they have different and necessary roles that, when combined, determine the security of the entire platform.

Authentication verifies the identity of a user or service. This is necessary to protect and secure access to the website, its data and all its content. For example, when we need to access a website or online service, we usually need to enter our username and password. Then, behind the scenes, it compares the username and password we entered with a record it has in its database. If the information entered matches, the system assumes we are a valid user and we are granted access.

Once the user's identity has been verified through authentication, it is up to the authorizer to determine the user's access rights. In other words, authorisation is the security process that determines the level of access a user or service has. As an example, we can think of ourselves as being employed in the marketing department of a company that uses an internal web application to store and process data. After successfully logging in, the content on the page must be specific to the department we belong to, without having access to the information of the accounting or human resources department, for example, for which we are not authorised.

In order to ensure these two processes, some simple, but at the same time crucial aspects of web application security must be ensured and checked. Authentication can be put at risk by attacks such as brute force attacks. Because of this, strict policies have been put in place and are being followed by more and more companies, such as fixed forms for passwords that can be used (a password must contain a minimum of 8 characters, a capital letter, a digit, and a special character).

Also, a good practice for maintaining authentication is to formulate the alert that appears when a wrong email or password is entered in such a way that it is not disclosed whether the email or username used exists in the system, so as not to provide an opportunity to gather information about web application users. If multiple unsuccessful login attempts have been made to an account, it should be automatically locked for a period of several minutes to provide protection against brute force attacks.

Another best practice that is progressively being adopted with web application development is the integration of two-factor authentication, called 2FA for short, which requires two different authentication methods. Using this method, a user is granted access to a website only after demonstrating two or more pieces of confirmation to an authentication mechanism: knowledge, such as knowing the account username and password; possession, such as an access card to an office building; and inherence, such as a fingerprint.

To ensure effective authorisation, it must be verified that each user has access only to data specific to the group to which they belong. Authorisation can be bypassed by simple actions such as adding a group- or user-specific extension in the access link to a web application (e.g., www.cybercon.ro/admin). If that page opens and the information is visible, this is considered a security vulnerability.

To ensure customer trust it is crucial to offer total transparency over personal data policies and ways of enforcing them. With the expanding demand for personal data handling by every website, customers are getting increasingly distrustful over how their information is stored and shared. Demonstrating transparency by giving details on the privacy policy that is used reassures customers that their information will be safeguarded and will lay a good base for customer trust and loyalty.

Cyber security threats such as phishing, malware and ransomware are often directed towards customers or employees, so a good way of protecting against this sort of attacks is having regular trainings with the employees about the dangers and effects, both on the company, and especially on the individual.

4. Risk Management

Cybersecurity risk management is an ongoing process of identifying, analyzing, evaluating, and addressing your organization’s cybersecurity threats [9]. The job of risk management involves everyone in the organization. For this, key actions must be considered, such as: developing robust

policies and tools to assess merchant risk; mitigation of IT risks, possibly through training programs, as mentioned before, or new policies; identification of internal weaknesses such as lack of 2FA; testing of the overall security process; documentation of merchant risk management and security; identifying emergent risks, like new regulations with business impact.

Generally, the strategy for managing risks involves following a four-step process which helps organizations have a better grasp and control over cyber security threats. The process commences with identifying the risks and assessing them based on potential impact and the possibility of attackers exploiting present vulnerabilities. After the analysis is completed, risks are prioritized based on each company's preferred mitigation strategy. The final step is monitoring risks, which concentrates on controlling the response, even in a constantly changing environment. This process can help e-commerce businesses to develop and adopt a cyber-security risk management plan.

Software approaches that should be used to protect against cyber-crime involve using cyber-security tools such as firewalls, encryption software, digital certificates, digital signatures, public key infrastructure and a strong password policy. Data plays an important role in any e-commerce business, so the best path to ensure data integrity is to use hashing and perform regular data back-ups, which will establish data availability and integrity in the event of a breach.

Organizations should follow the appropriate standards and frameworks for certifying a safe and secure online shopping experience. Some of the most important ones that also offer best practices and requirements for cyber risk management are ISO/IEC 27001:2022 and NIST Cybersecurity Framework Version 1.1.

Table 1. Cyber risk management according to standards and regulations [10][11]

Standard	Risk management strategy
ISO/IEC 27001:2022	Establishing and maintaining information security risk criteria
	Ensuring that repeated risk assessments produce consistent, valid and comparable results
	Identifying risks associated with the loss of confidentiality, integrity and availability for information withing the scope of the information security management system
	Identifying the owners of those risks
	Analyzing and evaluating information security risks according to the criteria established earlier
	Identifying and documenting asset vulnerabilities
NIST Cybersecurity Framework Version 1.1	Tuning into the latest cyber threat intelligence from information-sharing forums
	Identifying and documenting threats, both internal and external
	Identifying the potential business impacts and likelihood of risk events, utilize threats, vulnerabilities, likelihood, and impacts to determine risk
	Identifying and prioritizing risk responses

5. Effects of a Full-Scale Cyber-Attack

In January 2021, an attack with one of the largest number of victims happened on the Microsoft Exchange Servers, one of the largest email servers in the world, affecting over 60.000 businesses, companies, and organizations worldwide. The attackers took advantage of four security vulnerabilities that were not known to the provider, called zero-day vulnerabilities, which can be seen in figure 2, explained by Microsoft in [12]. Thus, they were able to gain unauthorized access to user emails, passwords, and administrator privileges to small businesses and even US governments.

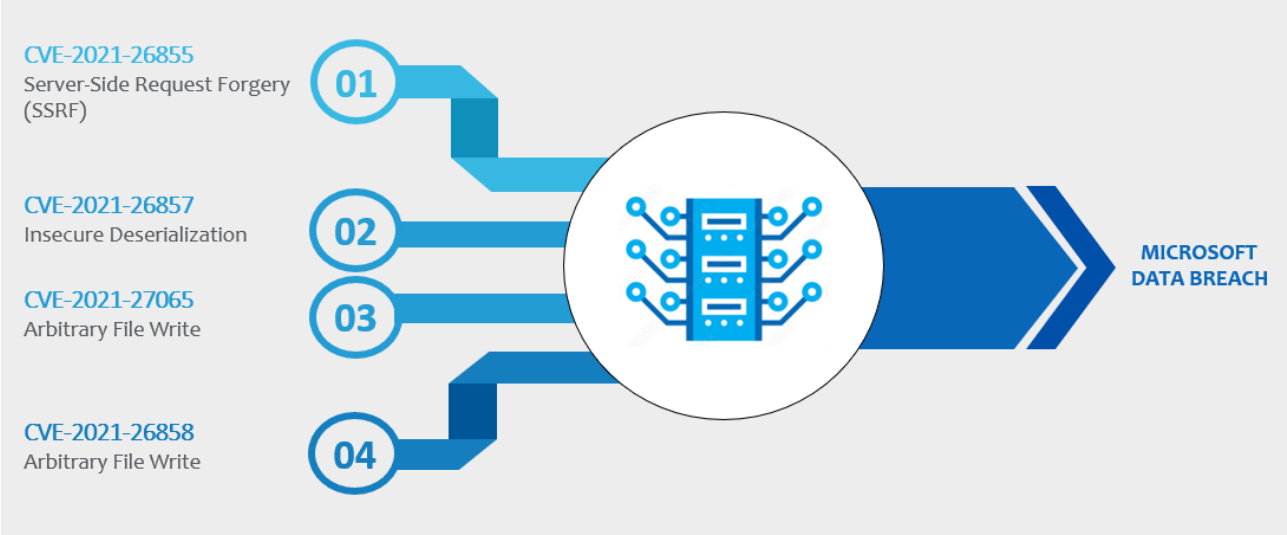


Fig. 2. Zero-day vulnerabilities exploited in the 2021 Microsoft data breach

The first of the zero-day vulnerabilities allowed attackers to connect to a server by creating session IDs and access tokens, to then authenticate as a standard user without actually owning those privileges. Afterward, a second vulnerability was exploited which gave administrator rights to the falsely authenticated user, when the last two unknown security weaknesses allowed attackers to upload code to the server in any location, using the administrator privileges.

The operation required three months in which the cyber-criminals carefully searched for coding errors which would allow them to gain control of the vulnerable systems. They then broke into each company's email server with only a connection to the internet and a locally managed system. The next step was to install malware, a web shell that provided a backdoor to the compromised servers, to access other systems and fundamentally take over the whole server.

The cyber-criminals used the web shell to run commands remotely and gather information such as passwords and email addresses, which was possible because Microsoft Exchanges doesn't use encryption to store them in memory. They also added users, added additional backdoors to other vulnerable systems and installed ransomware.

Since the hackers were able to access organizations' systems, the requests appeared to be coming from the MES, so Microsoft could not detect the malicious code and approved it. Eventually, they discovered the vulnerabilities and patched them. Microsoft released a total of 24 security updates for MES 2013, 2016, and 2019. Table 2 shows a timeline of the revisions performed on the servers.

Table 2. Microsoft Released Security Updates Timeline [13]

Version	Date released	Details
1.1	2 March 2021	CVSS scores were updated for the affected products
1.0	2 March 2021	Information was published
2.0	8 March 2021	Security updates for CVE-2021-27065, CVE-2021-26855, CVE-2021-26857, and CVE-2021-26858 for several Cumulative Updates that are out of support, including Exchange Server 2019 CU 6, CU 5, and CU 4 and Exchange Server 2016 CU 16, CU 15, and CU14
3.0	10 March 2021	Security updates for CVE-2021-27065, CVE-2021-26855, CVE-2021-26857, and CVE-2021-26858 for several Cumulative Updates that are out of support, including Exchange Server 2019 CU 3; and Exchange Server 2016 CU 17, CU 13, CU12; and Exchange Server 2013 CU 22, CU 21
4.0	11 March 2021	Final set of security updates for CVE-2021-27065, CVE-2021-26855, CVE-2021-26857, and CVE-2021-26858 for several Cumulative Updates that are out of support, including Exchange Server 2019, CU1 and CU2; and Exchange Server 2016 CU 8, CU 9, CU10, and CU11

Version	Date released	Details
5.0	16 March 2021	Security update for CVE-2021-27065, CVE-2021-26855, CVE-2021-26857, and CVE-2021-26858 for Microsoft Exchange Server 2013 Service Pack 1

Even if Microsoft released security updates for the vulnerabilities, the companies and organizations that were using the servers were still susceptible to attacks until they upgraded their systems as well, otherwise hackers would have still been able to exploit the CVEs. This would not have been necessary if the systems used a cloud-native infrastructure, because Microsoft could provide automated security by pushing the patch and immediately fix the issues.

As we've seen, no entity is completely safe against cyber-security threats and there is no definite template to keep hackers away, but it may be worth to consider using solutions that engage security into your development pipeline from the first release, rather than struggling to push patches after data has already been breached.

6. Conclusions

E-commerce is a flourishing industry, and it is set to only develop more over time. To establish a successful business, one needs to take into consideration all of the threats, vulnerabilities and risks, which is a harder job than ever seeing the on-going evolution of technology today.

This paper is a starting point in studying the cyber-security of the e-commerce industry. An analysis of cyber-security threats for e-commerce businesses was conducted, revealing the most common vulnerabilities and attacks that pose a risk for the business. The results were documented and explained for every person that may be interested in developing a successful and safe shopping experience for customers. For this purpose, best practices were underlined, and a risk management strategy was proposed.

The most common cyber threats include phishing, malware and ransomware, SQL injection, XSS, e-skimming. The best approach for protecting an e-commerce business is to set and apply a set of best practices which consist of using encryption software, firewalls, 2-factor authentication, password policy, transparency about data policy, tools to ensure data integrity and availability, and using cloud infrastructure to provide automated security. Important factors to consider for a safe environment for customers are authentication and authorization, which set the basis for effective cyber-security.

Future developments should contain exploring vulnerabilities represented by the vendor's implemented software, such as PHP version, databases, web servers or other such components. Other vulnerabilities might be represented by physical attacks or employee error (accidentally sharing sensitive business information).

In light of the constant evolution of software technology and cybersecurity, cyber criminals are becoming progressively more able to tackle complex web applications, with the use of hacking tools and online available documentation. Thus, vendors, retailers, providers, merchandisers, and any person aspiring to become a part of the e-commerce industry must be prepared for the most common cyber threats and have a well-established risk management strategy.

References

[1]. M. Mclean, 2023 Must-Know Cyber Attack Statistics and Trends, 2023, Available online at: https://www.embroker.com/blog/cyber-attack-statistics/. Accessed on 14.03.2023.

[2]. R. Zhang, L. Fang, X. He, C. Wei, E-commerce and E-commerce Security, in The Whole Process of E-commerce Security Management System: Design and Implementation, Singapore, 2023, pp 1-4.

[3]. Schatz, D., Bashroush, R., and Wall, J. (2017). Towards a more representative definition of cyber security. J. Digit. Forensics Secure. Law 12, 1558–7215.
[4]. BigCommerce, „What You Need to Know About Securing Your Ecommerce Site Against Cyber Threats", 2020 [Online]. Available online at: https://www.bigcommerce.com/articles/ecommerce/ecommerce-website-security/. Accessed on 05.03.2023.
[5]. Galov, N. (2022). 17+ sinister social engineering statistics for 2022. Available online at: https://webtribunal.net/blog/social-engineering-statistics/#gref. Accessed on 10.03.2023.
[6]. CH. Sireesha, V. Sowjanya, Dr K. Venkataramana, „Cyber security in E-commerce" in International Journal of Scientific & Engineering Research, 2017, pp 187-193.
[7]. Adobe Experience Cloud Blog, "Ecommerce security - what it means, common threats, and modern best practices". Available online at: https://business.adobe.com/blog/basics/learn-about-ecommerce-security#:~:text=Ecommerce%20security%20is%20a%20set, need%20to%20defend%20against%20cyberattacks. Accessed on 10.03.2023.
[8]. M. Abomhara, G. M. Køien, „Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks" in Journal of Cyber Security, 2015, pp 65-88.
[9]. Hyperproof, "Cybersecurity Risk Management: Frameworks, Plans, & Best Practices", 2023. Available online at: https://hyperproof.io/resource/cybersecurity-risk-management-process/#:~:text=and%20manage%20risk.-,What%20is%20Cybersecurity%20Risk%20Management%3F,has%20a%20role%20to%20play. Accessed on 13.03.2023.
[10]. Information security, cybersecurity and privacy protection — Information security management systems - Requirements, International Standard ISO/IEC 27001; Geneva, 2022. Available online at http://www.itref.ir/uploads/editor/2ef522.pdf. Accessed on 13.03.2023.
[11]. Framework for Improving Critical Infrastructure Cybersecurity, National Institute of Standards and Technology, 2016. Available online at https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf. Accessed on 13.03.2023.
[12]. Microsoft, Security Update Guide, 2023, Available online at: https://msrc.microsoft.com/update-guide/vulnerability. Accessed on 15.03.2023.
[13]. Microsoft, Microsoft Exchange Server Remote Code Execution Vulnerability, 2021. Available online at: https://msrc.microsoft.com/update-guide/vulnerability/CVE-2021-26855. Accessed on 15.03.2023.

Types of Attacks and Security Methods. Virtual Machines

Dorina-Luminița COPACI[1], Constantin-Alexandru COPACI[2]
[1] Associate Professor, University Politehnica Bucharest - ETTI, Bucharest, Romania
lcopaci@yahoo.com
[2] Student, Titu Maiorescu University, Faculty of Informatics, Bucharest, Romania
copacialexandru8@gmail.com

Abstract

Virtualization is a type of process used to create a virtual environment. Many organizations think about the security implications after implementing a new technology. Virtualization can be used in many ways and requires appropriate security controls in each situation. This paper presents the idea of using a virtual machine to share services and information over the Internet. In case of an attack, the resources of the virtual machine will be affected, while the resources of the real machine are safe. In this paper, we present the perspective of an attack by running malicious software on a virtual machine. We will show that although unauthorized control of the virtual machine is obtained, the real machine is not affected.

Index terms: attack, security, Virtualization, virtual machine, VMware Workstation

1. Introduction

The introduction of computers into virtually every dimension of society has significantly changed the way people and organizations obtain or disseminate information or conduct business, allowing for greater efficiency, increased operational control, and efficient access to information. Along with many benefits, however, computers and their interconnection also present negative aspects, such as the emergence of new types of crimes (for example, the distribution of computer viruses), as well as the possibility of committing traditional crimes through new technologies (to for example, fraud or forgery). Since attacks on information systems can produce a series of negative consequences - financial, operational, legal, or strategic - at an individual, organizational or even national level, the risk of an attack must be well understood in order to be mitigated or eliminated [1].

In this paper we propose to discuss the types of electronic attacks, as well as methods of securing computer systems. We present the idea of using a virtual machine [6], [7] to protect real machine resources after a network attack. Thus, an attacker will attack the virtual machine that is installed on top of the real machine. In section 2 we present the concept of virtual machine. In this part of the work, we describe the VMware Workstation. Section 3 presents some of the types of attacks. The security methods are presented in section 4, to be exemplified by a case study. In the last section, we present our conclusions as a result of the study done for the realization of the work.

2. The concept of virtual machine

A virtual machine (VM) is an operating system (OS) or application environment that is installed on software, which imitates dedicated hardware [5]. A virtual machine provides an isolated

environment for running its own OS and applications independently from the underlying host system or from other VMs on that host. The virtual machine depends on the physical resources of the host. These resources are virtualized and distributed on the virtual machine and can be reallocated as needed so that it is possible to run different environments simultaneously and adapt workloads.

Advantages of the virtual machine:

- Partitioning – multiple applications and operating systems in one machine;
- Isolation – each virtual machine works in isolation from the hosts and from the other virtual machines;
- Encapsulation – every state of a virtual machine is contained in software, with standard virtual hard drives guaranteeing compatibility.

Types of Virtual Machine [7]:

1. System Virtual Machines — Hardware Virtual Machines

This provides an environment with the execution of separate complete operating system. Examples of this type of virtual machines are: VMWare, VirtualBox

2. Process Virtual Machines — Application Virtual Machines

This provides platform independent programming environment that abstract away details of the underlying hardware from software. Ex: .NET Framework, Java Virtual Machine.

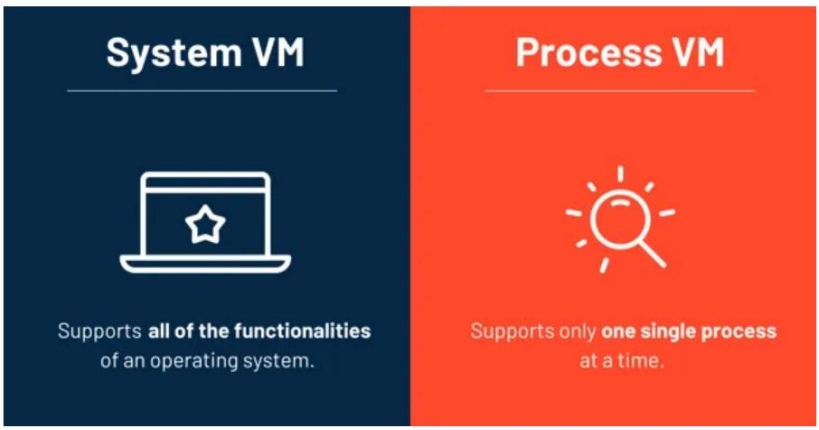


Fig. 1. Types of Virtual Machines [7]

2.1. VMware Workstation

Workstation is powerful desktop virtualization software for software developers/testers and enterprise IT professionals that runs multiple operating systems simultaneously on a single PC.

VMware [8], [10] refers to the computer and operating-system instance that executes the VMware Workstation process as the host machine, and identifies instances of operating systems running inside a virtual machine as guest virtual machines.

Like an emulator, VMware Workstation provides a completely virtualized set of hardware to the guest operating system — for example, regardless of make and model of the physical network adapter, the guest machine will see an AMD PC-net network adapter.

VMware [9], [10] virtualizes all devices within the virtual environment, including the video adapter, network adapter, and hard disk adapters. It also provides pass-through drivers for USB, serial, and parallel devices.

3. Types of attacks

Attacks on information in computer systems can take different forms.

A first classification of attacks can be made taking into account the place from where the attack is executed. We distinguish two categories of attacks: local and remote.

A second classification can be made according to the way the attacker interacts with the information resulting from a successful attack. Here two categories of attacks are distinguished: passive and active.

There are two main categories of attacks [14]: passive attacks (data interception) (Figure 2) and active attacks (data flow interruption, data modification and disinformation) (Figure 3).

a. The passive attacks are characterized by: they violate the confidentiality rules; they do not generate damages (do not delete or modify the data); transmitted data are intercepted using tapping wires, electromagnetic radiation interception, etc.
b. The active attacks are more dangerous, because they modify the status of data, computers or communication systems. There are the following main types of active attacks:
    - *Interruption* – uses the replay of a message or of a part of a message in or der to produce an unauthorized access.
    - *Modification* - represents an attack that modifies (through insertion and/or deletion of characters) a part or all transmitted data.
    - *Disinformation* – represents a type of attack where an unauthorized user pretends that is an authorized user.

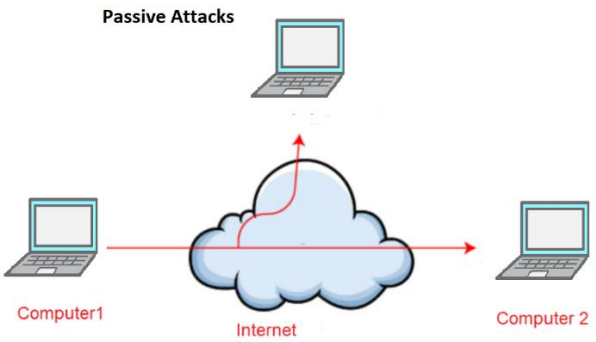


Fig. 2. Passive Attacks

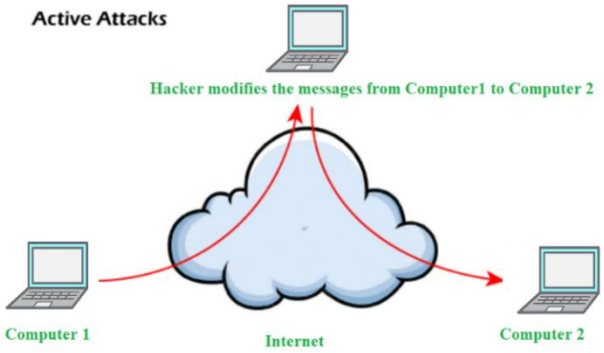


Fig. 3. Active Attacks

3.1. Examples of attacks

DOS attack

A denial-of-service attack (DoS attack) [11] is an attack on a computer system or network that causes a loss of service to users, typically the loss of network connectivity and services by consuming the bandwidth of the victim network or overloading the computational resources of the victim system. A DoS attack can be perpetrated in a number of ways.

There are three basic types of attack: consumption of computational resources, such as bandwidth, disk space, or CPU time; disruption of configuration information, such as routing information; disruption of physical network components. Examples of DOS attack: SYN flood attack, Fraggle attack, Ping of death attack, Distributed Denial of Service attack etc.

Viruses attack

In computer security, a computer virus is a self-replicating computer program that spreads by inserting copies of itself into other executable code or documents. A computer virus behaves in a way similar to a biological virus, which spreads by inserting itself into living cells. Extending the analogy, the insertion of a virus into the program is termed as an "infection", and the infected file, or executable code that is not part of a file, is called a "host".

Viruses are one of the several types of malicious software or malware. In common parlance, the term *virus* is often extended to refer to worms, trojan horses and other sorts of malware; viruses in the narrow sense of the word are less common than they used to be, compared to other forms of malware. Examples of viruses attack: Companion viruses, Resident viruses, Nonresident viruses etc.

Trojan attack

In the context of computer software, a Trojan horse is a malicious program that is disguised as or embedded within legitimate software. The term is derived from the classical myth of the Trojan Horse. They may look useful or interesting (or at the very least harmless) to an unsuspecting user, but are actually harmful when executed.

Backdoor attack

A backdoor [13] in a computer system (or cryptosystem or algorithm) is a method of bypassing normal authentication or securing remote access to a computer, while attempting to remain hidden from casual inspection. The backdoor may take the form of an installed program (e.g., Back Orifice) or could be a modification to a legitimate program.

Buffer overflow

In computer security and programming, a buffer overflow [1] is an anomalous condition where a process attempts to store data beyond the boundaries of a buffer. The result is that the extra data overwrites adjacent memory locations. The overwritten data may include other buffers, variables and program flow data.

E-mail spoofing

E-mail spoofing [12] is a technique commonly used for spam e-mail and phishing to hide the origin of an e-mail message. This involves changing certain properties of the e-mail, such as the *From*, *Return-Path* and *Reply-To* fields (which can be found in the message header) to make the e-mail appear to be from someone other than the actual sender.

4. Security methods

The security model for a system (a computer or a network of computers) can be seen as having several layers that represent the levels of security surrounding the subject to be protected. Each level isolates the subject and makes it more difficult to access it in any other way than it was intended.

Physical security represents the outer level of the security model and generally consists of locking computer equipment in an office or other premises as well as ensuring security and access control. Logical security consists of those logical methods (software) that ensure access control to system resources and services. It, in turn, has several levels divided into two large groups: access security levels and service security levels. Among the security methods of a system, we mention:

Virtual private networks

A Virtual Private Network (VPN) provides a way to establish secure communications over an otherwise insecure network. With the help of a VPN connection, the two sides of a connection can communicate under the same security conditions as those provided by a company's local network.

Firewall

Firewalls are used to protect/isolate segments of an extended network (eg the Internet), but especially to protect the private networks (Intranet) of a company/institution/bank/etc connected to the Internet. In what follows, by internal network we will refer to the network segment that must be

protected, respectively by external network we will refer to the network segment from where the threats can originate and over which we have no control (usually the rest of the Internet).

Antivirus programs

Antivirus programs must be chosen so that they have as large a database as possible with the definitions of known viruses in order to effectively protect the system, occupy as little memory as possible when monitoring computer activity and update themselves on the Internet as often as possible.

4.1. Case study. Practical example

We installed VMware Workstation on two operating systems: Windows and Linux. Then we installed the Windows 10 and Linux Red Hat 9.0 operating systems on the virtual machine [4].

We will try to attack both operating systems on the virtual machine to see if the local machine resources will be affected.

4.1.1. Security methods against a DOS attack

For a Denial-of-Service attack we use a program which was written for a Linux machine with a kernel patch in place to allow IP source address spoofing.

This program scans a host to determine which ports are open, or listening for connections. Once a list of receiving ports has been compiled, the program then floods each of them with the specified number of SYN packets.

When a TCP/IP stack receives a SYN packet, it responds with a SYN/ACK. At this point, it is waiting for an ACK. Now, if the source address in the SYN packet does not exist, but has a path to it in place, that SYN/ACK will never be answered with an ACK, and the TCP/IP stack will wait forever for that packet (actually until a certain amount of time has passed which is implementation-dependent). If a whole bunch of those faked SYN packets are received simultaneously, the connection queue of the target machine will he filled.

We can set the network connection to a virtual machine in three ways:

- Bridged – connected directly to the physical network;
- NAT – used to share the host's IP address;
- Host-only – a private network shared with the host.

If we chose to set up the network connection bridged, the virtual machine we'll use a different IP address from the real machine, and the bandwidth will be partly affected in case of a DOS attack over the virtual machine.

So, if we want that the local machine bandwidth not to be affected, we'll have to install a net limiter to divide the bandwidth in two: one for the virtual machine IP and one for the real machine IP. When the virtual machine IP will be flooded, half of the bandwidth will be affected, but the second half of the bandwidth will not be affected. In this way the real machine resources will not be affected.

If we chose to set up the network connection NAT, the virtual machine will use the same IP address as the real machine. In this case, if a DOS attack will affect the virtual machine, the real machine will be affected too. The real machine bandwidth will be affected.

So, if we want that the real machine resources not to be affected after a DOS attack, we must set the network connection bridged and then to use a net limiter to limit the bandwidth for the virtual machine IP address.

Otherwise, we can create a strong firewall on the virtual OS. We must set the firewall to drop the packets in case of a flood. In a Linux operating system [3], we can use the *iptables* command to drop the packets in case of a DOS attack.

4.1.2. Security methods against a buffer overflow attack

To create a buffer overflow attack [1], we first accessed the virtual machine with a Trojan horse. We inserted programs [3] into the virtual operating system with the intention of creating buffer overflows and then executed them.

We observed that this attack can destroy certain services or affect the memory allocated for the virtual machine, but the memory of the real machine remains unaffected. Therefore, the buffer overflow attack on the virtual machine has no effect on the resources of the real machine.

5. Conclusions

In this paper we wanted to study the situation in which, having installed a virtual machine on a real machine and attacking this machine, what is the probability that the resources of the real machine will be affected.

We analyzed some of the important types of attacks in a network, such as: DoS attacks, viruses, Trojans, backdoors, buffer overflow. After this analysis it was concluded that the probability of the real machine resources being infected in the event of an attack is very small. If the attack on the virtual machine represented a threat to the resources of the real machine, we presented in the paper solutions to remedy the problem.

Therefore, the resources of the real machine will be safe after an attack on the virtual machine if the instructions in the work are used.

References

[1]. Bernaschi, M., Grabrielli, E., Mancini, L. (2000) “Operating System Enhancements to Prevent the Misuse of System Calls”, Proceedings of the ACM Conference on Computer and Communications Security.

[2]. Bernaschi, M., Grabrielli, E., Mancini, L. (2002) “REMUS: A Security-Enhanced Operating System”,ACM Transactions on Information and System Security, Vol 5, 2001,

[3]. Blunden, B. (2002) “Virtual Machine Design and Implementation in C/C++”, Wordware Publ. Plano,Texas – USA.

[4]. Chen, P., Noble, B. (2001) “When Virtual Is Better Than Real”, Proceedings of the 2001 Workshop on Hot Topics in Operating Systems (HotOS).

[5]. Goldberg, R. (1973) “Architecture of Virtual Machines”, AFIPS National Computer Conference. New York – NY– USA.

[6]. *Oliphant,P., "Virtual Machines". Virtual Computing. Archived from the original on 2016-07-29*. Retrieved 2015-09-23.

[7]. Randika, Y., “Virtual Machines”, Feb 28, 2021, https://yasirurandika.medium.com/virtual-machines-937c99156ca5.

[8]. Sugerman, J., Ganesh, V., Beng-Hong L. (2001). Virtualizing I/O Devices on VMware Workstation’s Hosted Virtual Machine Monitor. Proceedings of the 2001 USENIX Annual Technical Conference.

[9]. VMware Inc. (1999) “VMware Technical White Paper”, Palo Alto – CA - USA.

[10]. Vmware Emulator. http://www.vmware.com.

[11]. https://www.fortinet.com/resources/cyberglossary/types-of-cyber-attacks

[12]. https://www.proofpoint.com/us/threat-reference/email-spoofing

[13]. https://www.educative.io/answers/what-is-a-backdoor-attack

[14]. https://www.javatpoint.com/active-attack-vs-passive-attack

A Signal Theory Model for Security Monitoring using CheckMK

Iliuță-Alexandru IONEL
Faculty of Electronics, Telecommunications, and Information Technology,
University POLITEHNICA of Bucharest, Romania
iliuta.ionel@stud.etti.upb.ro

Abstract

Continuous monitoring of intelligent systems is used to analyze data and text from various sources. They usually monitor things such as risk, controls, opportunities, competition, and other concerns. While there exists literature that provides information on the capabilities of this kind of system, there has been a limited theoretical development in this field. The information sources monitored by these systems provide signals related to events, activities, or issues. However, selecting the appropriate information sources is not a simple task, because it is influenced by factors such as time, cost, redundancy, reliability, or weak signals. Furthermore, for the monitored signals, it is recommended to generate some analytics to study the flow and have a traceability of the issue we are dealing with. In this paper, a signal theory model is introduced and applied to address some of these issues regarding the SSH brute-force attacks. I will use a tool called CheckMK and its capabilities to implement a signal theory model used for monitoring security of a system [1].

Index terms: Brute force, Monitoring, Security, Signal, SSH

1. Introduction

Performing security audits on large-scale networks is a complex task, as the configurations of networked devices are constantly changing in the ever-evolving environment of today's technology. Many critical devices use the Secure Socket Shell (SSH) protocol to expose their remote access interfaces to the internet, and default usernames and passwords are often used. The issue has been exacerbated by the widespread availability of stolen credentials, which enables attackers to pose as authorized users and gain unauthorized access to internal networks, compromising sensitive data and misusing computational resources to gain some form of leverage from the legitimate users or administrators. Although the success rate of such attempts is low, they have resulted in significant consequences for 51% of the 1,800 organizations surveyed, with a financial impact of up to $500,000 per organization [2]. Brute force attacks are among the most common types of attacks on machines connected to the internet. Brute force attacks are essentially a trial-and-error method of guessing login credentials by attempting numerous combinations of usernames and passwords until the correct combination is discovered.

To protect against these attack vectors, several defense techniques have been developed. One of the most common techniques is to implement a password policy that mandates the use of complex passwords or passphrases, which are more difficult to guess. This can involve requiring a certain length or complexity, and even enforcing regular password changes. Another technique is to implement account lockout policies that temporarily lock an account after a certain number of failed login attempts, which can prevent brute force attacks from being successful [3].

2. Related Research

2.1. Signal Theory

Professor Emeritus Michael Spence from Harvard and Stanford Universities observed, "If the incentives for veracity in reporting anything by means of a conventional signaling code are weak, then one must look for other means by which information transfers take place." This implies that signals provide information about individuals, events, activities, and more, when conventional signaling codes fail to provide accurate information. For example, a college degree can serve as a signal of an individual's capabilities. Professor Brian Connelly et al. provided a definition of signaling theory, stating, "Signaling theory is useful for describing behavior when two parties (individuals or organizations) have access to different information." Information asymmetries can arise due to differences in experiences, opinions, observations, and even incentives, resulting in different information being available from different sources [1][4][5].

Signal theory has been applied in various fields of business research, including accounting, economics, finance, and management. For example, education serves as a signal for potential employees to showcase their market value in human resources. Similarly, faculty members may use information about their research publications and teaching experiences to signal their value to other universities.

In the context of continuous monitoring systems (CMS) for security purposes, signal theory is especially relevant due to the multiple information sources used and the need to choose appropriate analytics. These sources can vary significantly in terms of availability, reliability, informativeness, and veracity, and often contain unstructured data that requires structuring for analysis. Signal theory is useful for aligning analytics with the relevant signals, such as counting the number of relevant topics identified [1].

2.2. SSH (Secure Shell)

Secure Shell (SSH) is a critical remote access protocol that allows users to log into or execute commands on a remote computer, copy files, and perform other functions. Compared to plain-text communication protocols like Telnet, which lack encryption, SSH provides secure management and confidentiality of communications by supporting remote system communications through strong authentication and encryption.

While direct console connections are recommended for initial network system setup, SSH is commonly used to enable remote access. However, allowing SSH service from all origins may expose a system to threats, as not all remote access requests come from managers. Therefore, it is important to establish policies that only permit remote access to SSH from manager IPs and block the others IPs, and to define appropriate firewall start and end points [6].

2.3. Brute-force attack

Brute-force attacks involve submitting all possible account inputs to access the system's account information. These attacks use dictionary or random sequence methods, with the former trying all strings in a pre-arranged listing and the latter trying all possible string combinations in a sequence. To counter brute-force attacks, access can be controlled when an incorrect password is entered beyond a set number of times. For servers, the account lock policy restricts access once the login failure threshold is exceeded. Password policies also help to protect the system and its accounts by setting the password complexity, usage duration, and minimum length. Network systems prevent unauthorized access through the ACL(Access Control Lists) and track unauthorized attempts for reference in the access control policy. While network systems do not block access following login failures exceeding the threshold, logs of many attempts of unauthorized external access on Internet routers can be a valuable reference [6].

2.4. Server Virtualization

A virtual machine (VM) acts as an intermediary layer between the hardware components and the user, enabling the execution of an operating system within a virtual environment. Virtual machines are sometimes referred to as virtual servers and can be hosted on a physical server that shares its hardware resources such as CPUs, memory, and I/O (input/output ports) among the virtual machines. In contrast, a "real machine" or a bare metal system refers to the host operating system and hardware components such as the memory, CPU, motherboard, and network interface. Virtual machines are built on top of the core components of a real machine, and the communication between the virtual and real machines is facilitated by hypervisors or virtual machine monitors (VMMs). Hypervisors provide a layer of abstraction that allows different virtual machine operating systems and configurations to run on the same real machine hardware components. While hypervisors and emulators are similar, hypervisors are more efficient as they have specialized management functions that enable multiple virtual machines to co-exist peacefully and share real machine resources. Therefore, hypervisors and emulators are different primarily in terms of their semantic meanings rather than their functionalities [7].

2.5. Network Security Monitoring

A network monitoring system is utilized to monitor the internal network for identifying slow or failing system components, and reporting and resolving problems. Continuous monitoring of the network helps maintain high-performance networks with little downtime, regardless of whether it is a small business or a large enterprise. Monitoring reports cater to different levels of audiences, including the network and systems administrators and management. Therefore, a monitoring system should have basic reporting and drill-down functionalities, but it should also be easy to understand and use, and not too complex so basic users can understand and properly use the functionalities of the monitoring system. Network monitoring covers every aspect of a networked system, including response time, availability, uptime, and security, making it a difficult and demanding task. To maintain smooth operation, network administrators are constantly striving to optimize data flow and access in a complex and changing environment. The aim of network monitoring for security management is to protect sensitive information on devices connected to a data network by controlling access points to that information, preventing, or denying different forms of malicious attacks. There are various network monitoring approaches to ensure network security. By identifying specific activities and performance metrics, these systems can generate results that help address a range of needs, including compliance requirements, internal security threats, and operational visibility [8].

3. Implementation

This paper proposes to implement a Signal Theory Model for Security Monitoring by provisioning locally two Ubuntu 20.04 Virtual Machines that can communicate with each other and are monitored, mainly for security reasons using the tool called CheckMK.

CheckMK is a comprehensive monitoring solution that provides a unified dashboard to monitor the performance and health of various systems and applications on the network. It allows the administrator to monitor everything from servers and network devices to cloud services and containers. It provides various alerting options, including email, SMS, and mobile push notifications, so the administrator can be notified immediately if any issues arise. Additionally, it offers support for multiple notification channels to ensure that alerts reach the right people at the right time as in the continuous monitoring world any second can make the difference between a fiasco and a good functionality of the system.

CheckMK is a powerful and reliable monitoring solution that is well-suited for businesses and organizations of all sizes. Its comprehensive monitoring capabilities, automatic discovery and

configuration, and flexible alerting options make it an ideal choice for monitoring any network and keeping any systems running smoothly [9].

Ubuntu Server 20.04 (Focal Fossa) is a server-focused operating system that is designed to provide a stable and secure platform for running a wide range of server applications. It is based on the Linux kernel version 5.4 and offers long-term support (LTS) for five years, making it a reliable choice for businesses and organizations. Also, it is very easy to jump from a LTS version to another and all the applications installed on the Ubuntu Server have dependencies that are compatible with the new LTS releases. It is a powerful and versatile server operating system that is well-suited for a wide range of server deployments, from small businesses to large-scale cloud environments. Its extensive software library, comprehensive security features, and long-term support make it a popular choice for businesses and organizations of all sizes [10].

3.1. System Architecture

To simulate this Signal Theory Model, I will use 2 Virtual Machines with Ubuntu 20.04 as Operating System. Both are using static IPs from my internal network so they can communicate with each other. One machine is named "Attacker" with the job of making brute force attacks using SSH towards the other virtual machine. The other one is named "Monitoring Site". On this virtual machine I have installed CheckMK to mainly monitor the security of the server, as shown in Figure 1.

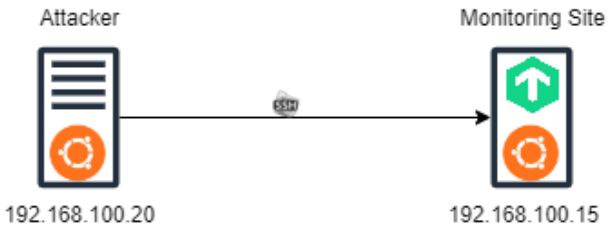


Fig. 1. System Architecture

The CheckMK instance on the "Monitoring Site" is used to count the number of SSH connection attempts that are made and to act depending on the that number. I will explain more when I define the flowchart of the system. The main role of the "Attacker" server is to test the security of the other server by making SSH brute force attacks, while the other server tries to deny the brute force attacks, notify the administrator about the intentions of that server and, if no action is taken by the administrator, to blacklist the IP of the attacker.

3.2. Flowchart

The Flowchart of the system is quite simple. The "Attacker" server will start to make SSH requests towards the "Monitoring Site" Server with the scope of brute force its way into the server. But the server is prepared to combat this kind of action by using the security utilities of CheckMK. So, if the attacker tries to connect via SSH as the root user 3 times in a minute, the server will notify the admin with a warning notification. If the attacker tries to connect via SSH 5 times in a minute, the server will automatically blacklist the IP of the attacker to protect itself from a brute force attack. This is a preventive method and it's mainly used for two reasons. The first reason is not to overwhelm the server with the SSH requests and cause a DDoS (Distributed Denial of Service) attack. The second reason implies human error: maybe the administrator was caught up in something else, did not see the notification involving the possible SSH brute force attack and as a result the server is down as it was overwhelmed by the SSH requests or, worse, the SSH brute force attack was a success.

As shown in Figure 2, the server tries to defend against brute-force attacks if the number of SSH requests it's too big and there is no administrator to take security measures. This practice is a

common one, as daily, if a server is opened to the internet, it will be bombed with SSH requests for a malicious connection on the server and if the resources of the server aren't good enough, the server might get down for some time. So, to save time and money, it is better to invest in a protection system that tries to prevent instead of treating the problem.

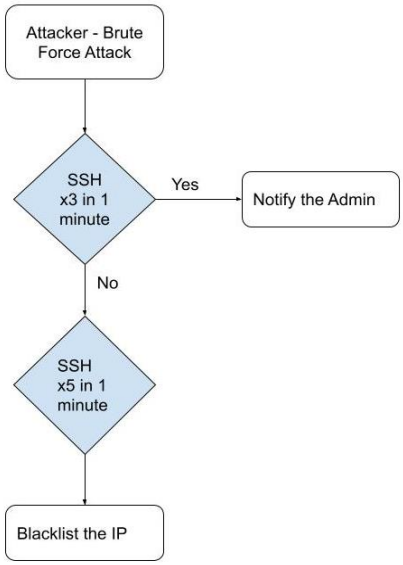


Fig. 2. Flowchart of the System

3.3. Deployment

The deployment of the Virtual Machines was made with the help of VirtualBox. VirtualBox is a powerful and feature-rich virtualization software that allows you to run multiple operating systems on a single computer. It is free and open source. It allows you to create virtual machines (VMs) that run these operating systems, each with its own set of virtual hardware components, such as virtual CPUs, memory, and disk space [11].

I used an Ubuntu 20.04.6 live server image to deploy both the machines. I chose this version because it's the Long-Term Support one, it's well-integrated with CheckMK and it is mature (first released in 2020) so it has a wide and detailed documentation that I can use in my research.

On the virtual machine named "Monitoring Site" I have installed CheckMK 2.1.0p25. This virtual machine has 2 GB of RAM, 2 CPUs and 20 GB of disk memory, and the machine running the "Attacker" server environment has 1 GB of RAM, 1 CPU and 20 GB of disk memory. The first one needs more resources because it runs the CheckMK tool and monitors both machines, just to be sure that the deployment was done correctly, and everything works properly.

3.4. Configuration

On both machines I had to configure the static IPs for the to communicate more easily between them. As an example of an IP configuration, let's look over the next figure, which shows the configuration made for the Monitoring Site.

As shown in Figure 3, I assigned the 192.168.100.15/24 IP to the "Monitoring Site" server with the gateway of 192.168.100.1, which is the default gateway of my internal network, and as a DNS server I've also used the default gateway. Now the machine is open to communicate with any other machine that has allocated an IP from my internal network. The same was done for the "Attacker" server, but with a different IP address (192.168.100.20/24).

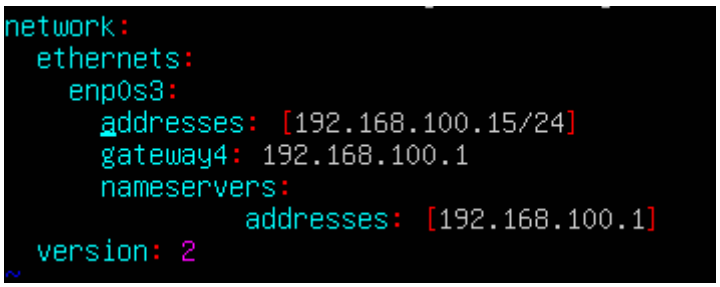


Fig. 3. IP Configuration for "Monitoring Site" Server

3.5. Defining the Rules

CheckMK is a rule orientated tool. To achieve the scope of this paper, I used a feature named Event Console. Event Console in CheckMK is a feature that allows users to view a chronological record of events that have occurred on a monitored system. These events can include system events, application events, and user-defined events. It provides a valuable tool for system administrators to monitor the health and status of their systems, troubleshoot issues, and take appropriate actions to prevent downtime and ensure system availability.

Using the Event Console, I have created two rules that will apply when the server notices that someone tries to SSH brute-force. For example, let's analyze the rule created for blocking the IP when it tries to SSH 5 times as the root user in a minute without having the password.

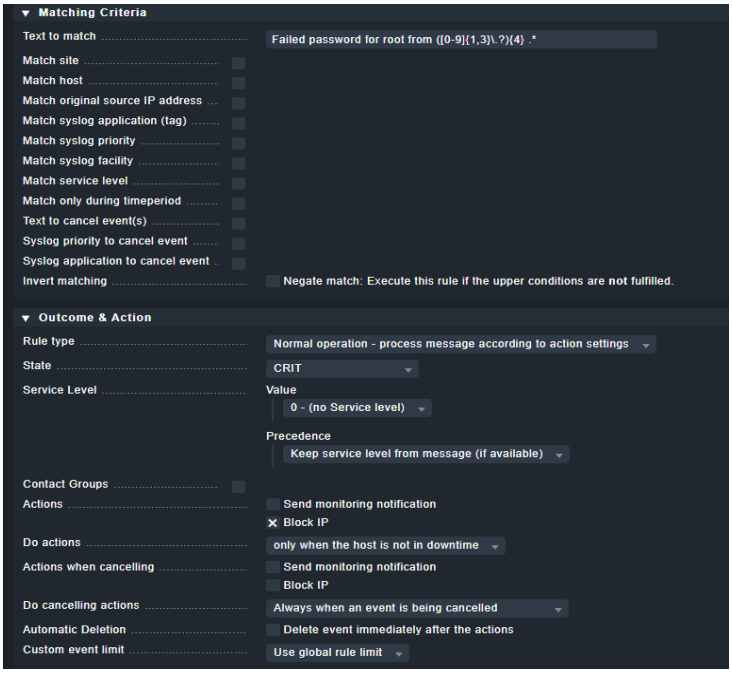


Fig. 4. Rule for Blocking the IP

I have defined a text to match which was taken from the journalctl logs of the "Monitoring Site" Server, I have added those logs for monitoring in CheckMK and I have used a Regex in order to define any possible IP that might try to SSH it's way into the server. I chose a State of CRIT, so the server administrator will know that an IP address was blocked when he reviews the Event dashboard of the server. As Actions I have created one that blocks the IP address using the Linux *iptables* command. The dashboard of the site and the action "Blocked IP" are not the scope of this paper, so I won't delve into explaining them more. I chose to use a Regex to define all the possible IP addresses, instead of specifying the IP address of the attacker (I know it in this case, because I have assigned it to the other server), but in a normal environment the administrator won't know the IP address of an attacker, so using a Regex that understands all the possible IP addresses is better and more proactive.

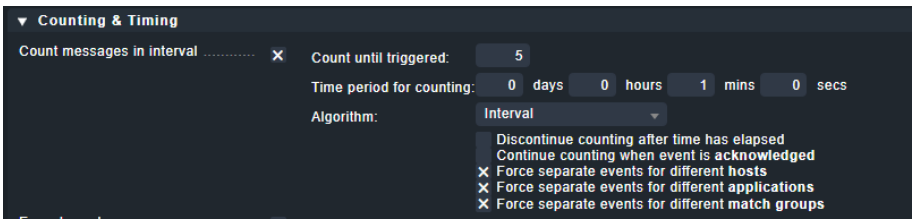


Fig. 5. Counting before blocking the IP

As I have specified earlier, CheckMK will count the SSH log in tries and if in a minute, there will be 3 failed log in tries, the server administrator will be notify. If the count reaches 5(as shown in figure 5) or more in minute, the server will try to protect itself alone by blocking the IP address from whom the SSH requestes came from. This is a proactive tactic that wants to treat the cause, not the effect of an event.

4. Tests and Results

To test the system, I am going to log in as root on the "Attacker" server and start to try to brute-force by SSH my way into the "Monitoring Site" Server and see what happens.

I am going to try to log as root 3 times from the "Attacker" Server to the "Monitoring Site" Server to see if I am going to be notified.

Fig. 6. SSH 3 times as root with the wrong password

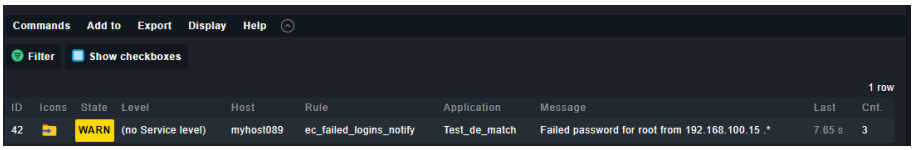


Fig. 7. The Admin was notified

Now I am going to SSH 5 times with the wrong password to see if the IP is going to be blocked:

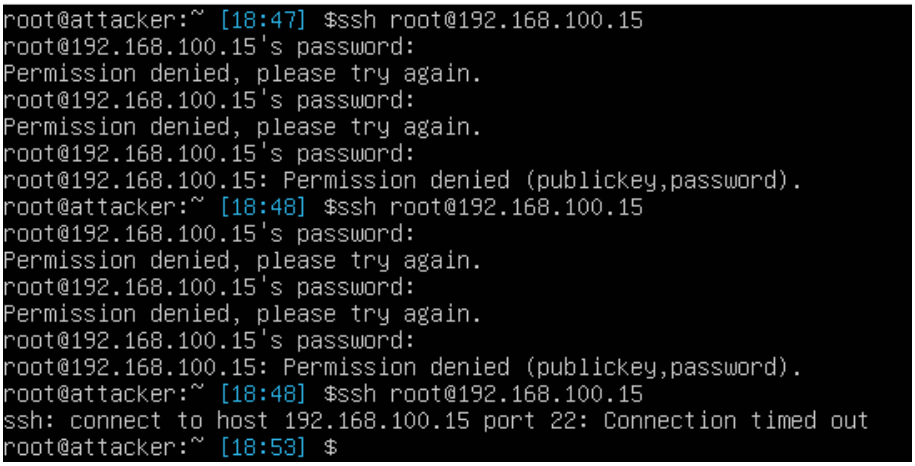


Fig. 8. "Monitoring Site" Server blocks the IP of the "Attacker" Server

As seen in Figure 8, the connection timed out, because the IP was blocked by the other server. The server is now protected against the brute force attacks coming from the "Attacker" server and the server's administrator was notified about the SSH attempts and the blocking of the IP.

5. Conclusion

In conclusion, the use of a Signal Theory model in security monitoring, combined with tools like CheckMK, can provide an effective means of detecting and responding to potential security threats. In this study, I have intended to demonstrate the practical application of this approach by setting up two servers: one with the role of the attacker, and the other one with the role of the defender. On the defender server, I have installed CheckMK to detect and respond promptly to SSH brute force attacks coming from the attacker server.

By applying signal theory concepts to network security monitoring, I was able to identify abnormal patterns of activity and trigger alerts based on the rules I have set. Specifically, I set up rules in CheckMK to detect repeated failed login attempts by an attacker and notify the administrator accordingly. If the attacker persisted in their attempts, the attacked server could automatically block their IP address to prevent further brute force SSH attempts and mitigate unauthorized access.

Overall, my experiment demonstrated the effectiveness of using a Signal Theory model with CheckMK for security monitoring. By leveraging the power of data analysis and pattern recognition, we were able to quickly identify potential threats and take proactive steps to mitigate them. As the threat landscape continues to evolve, this approach will become increasingly important for organizations looking to protect their valuable data and infrastructure.

References

[1]. O'Leary, Daniel E., A Signal Theory Model for Continuous Monitoring and Intelligence Systems (September 9, 2020). Available at SSRN: https://ssrn.com/abstract=3746001 or http://dx.doi.org/10.2139/ssrn.3746001J.

[2]. Phoung M. Cao et al, CAUDIT: Continuous Auditing of SSH Servers To Mitigate Brute-Force Attacks. Available at https://www.usenix.org/conference/nsdi19/presentation/cao.

[3]. Faust, Joshua, "Distributed Analysis of SSH Brute Force and Dictionary Based Attacks" (2018). Culminating Projects in Information Assurance. 56.

[4]. Spence, M., Job Market Signaling, The Quarterly Journal of Economics, Vol. 87, No. 3. (Aug., 1973), pp. 355-374.

[5]. Connelly, B., Certo, S., Ireland, R., Reutzel, C., (2011) "Signaling Theory: A Review and Assessment," Journal of Management, (37.2), January 2011, pp. 39-67.

[6]. Jeonghoon Park, "Network Log-Based SSH Brute-Force Attack Detection Model", Tech Science Press, 2021.

[7]. Daniels, Jeff, "Server virtualization architecture and implementation" XRDS: Crossroads, The ACM Magazine for Students Volume 16Issue 1September 2009 pp 8–12 https://doi.org/10.1145/1618588.1618592.

[8]. Ghafir, I., Prenosil, V., Svoboda, J., & Hammoudeh, M. (2016). A Survey on Network Security Monitoring Systems. 2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW). doi:10.1109/w-ficloud.2016.30.

[9]. Official Documentation of CheckMK: https://docs.checkmk.com/latest/en/.

[10]. Official Documentation of Ubuntu: https://releases.ubuntu.com/.

[11]. Official Documentation of VirtualBox: https://www.virtualbox.org/wiki/Documentation.

Digitalization of Finance: Effect or Cause of Programmed Chaos?

Ruxandra RÎMNICEANU, PhD
National Bank of Romania, Bucharest, Romania
ruxandra.rimniceanu@bnro.ro

Abstract

The actual "permacrisis" marks the five transitions that are unfolding simultaneously: a transition in the planet's climate regime, an energy transition, a geopolitical transition, a technological transition and a demographic transition. In this context, all the risks that are around show us that we are dealing with a programmed chaos that might affect the financial ecosystem, also. In this respect, such to avoid a collapse and to strengthen the banking and financial sector, the European entities appreciate that there it is necessary to strengthen the leadership of the EU in the digital domain by promoting inclusive and sustainable digital policies, serving citizens and businesses. Taking into account that the risks of increased exposure to potential cybercrime, operational resilience failures and data protection and privacy issues could have an important impact, the digital transformation must be in line with EU values - the 2030 policy program entitled "The Path to the Digital Decade" and "The Declaration on Digital Rights and Principles in the E.U.".

Index terms: cybercrime, data protection and privacy issues, digitalization of the financial services sector, financial ecosystem, operational resilience

1. Introduction

The overlap, interdependence and the impact of the multiple crises that we are going through today, in this "***permacrisis***"[1], reflect the profound transformations that we are witnessing and participating in, in equal measure, and **mark the five transitions** [1] that are unfolding simultaneously: *a transition in the planet's climate regime, an energy transition, a geopolitical transition, a technological transition and a demographic transition*.

The existing and foreshadowed risks, the disturbances and turbulences manifested, as well as the surrounding vicissitudes indicate that we are dealing with a **programmed chaos**[2] and it is increasingly necessary to develop ***catastrophic scenarios, for making prompt and optimal decisions***, based on **Decision Support Systems (DSS)**[3].

[1] "***Permacrisis***" is one of the ten terms of the year 2022, declared by the Collins English Dictionary as a word denoting ***an extended period of instability and insecurity*** - marked by war, pandemic, inflation, flood, drought or fire, recession, hunger, protests and political instability etc. - and ***which reflects the reality that will dominate economic, social and political life for more than a decade*** and which reveals that the world we live in seems increasingly chaotic.

[2] **Chaos theory** was formulated by Edward Lorenz in 1960. The scientist said, "*A phenomenon that appears to unfold at random actually has an element of regularity that could be described mathematically*." In simpler terms, there is a hidden order in any apparently chaotic evolution of any complex dynamic system, [Online]. Available: https://ro.wikipedia.org/wiki/Teoria_haosului.

[3] The history of **Decision Support Systems - DSS** ("Sistem Suport de Decizie" - SSD) began around 1965, when building large IT systems was quite expensive, [Online]. Available: https://en.wikipedia.org/wiki/Decision_support_system.

2. The digital age of finance

Over the past three years, digital technologies have reshaped the financial services industry and paved the way for innovative consumer financial products and services. They have also transformed traditional value chains and have given birth to new business models and new actors or business models in the financial-banking market. Through digitization, funding has become more competitive, accessible and inclusive.

For the purpose of *financial inclusion*, however, a distinction should be made between the two concepts to reveal the revolutionary possibilities of *open finance*. With *open banking*, data accumulated about bank customers can also be accessed by external financial service providers. But, in order for these providers to offer more individualized and tailored IT solutions to the situations for which they need to be used, banks use this method to share data about their customers' transactions with other external parties, too. However, this method is considered quite limited as the data sharing would not go beyond the operations of the bank itself.

Open finance, on the other hand, has the potential to open new horizons for both consumers and businesses, being a concept of collecting all of a user's financial information in one location, including but not limited to banking transactions (also, includes purchases made with digital wallets, payments made with insurance and pension accounts, investments, money transfers, and cryptocurrency transactions).

So, challenges remain and providing more data to product providers, start-ups, scale-ups and SMEs could lead to the implementation of innovative services and products in the internal market of each EU member state. Subsequently, the expansion of data could bring an additional diversity of innovative products that better reflect the needs of customers.

From this perspective, the adoption in the European Parliament of the *Report on the legislative initiative in digital finance* [2], marked the main areas considered important to be legislated at the European level, to support the growing digitization in the financial sector, as a result of the initiatives proposed by the European Commission in 2020, namely ***Regulation cryptocurrency markets***, the ***Digital Operational Resilience Act*** and the ***Distributed Ledger Technology (DLT) Pilot Regime***. They address different areas - digital asset trading, cyber resilience in the financial sector and the further development of new technologies that could bring efficiency to the sector.

At the same time, the European Union foresaw and accepted the need to develop innovative political considerations, in the sense that it intends that the adopted measures allow sufficient flexibility to encourage innovation and global competitiveness, at the same time addressing the risks in the financial-banking sector, in order to regulate the specific activity in a technologically and consumer protective, neutral manner.

Thus, on ***8 December 2022***, the Council of the European Union adopted **the 2030 policy program** entitled ***"The Path to the Digital Decade"*** [3][4], which ensures that the EU meets its objectives for a digital transformation in line with EU values, with **the aim of strengthening the leadership of the EU in the digital domain** by promoting inclusive and sustainable digital policies, serving citizens and businesses.

With this aim in mind, on ***December 15, 2022***, the President of the European Parliament, Roberta Metsola, the Council of the European Union - through the Prime Minister of the Czech Republic, Petr Fiala (whose country held the EU Presidency until the end of 2022) and the President of the EU Commission, Ursula von der Leyen - signed the ***"Declaration on Digital Rights and Principles in the European Union"*** [4], which presents the Union's commitment to *a safe,*

[4] The ***2030 Policy Program on the Digital Decade*** (Decision (EU) 2022/2481 of the European Parliament and of the Council of 14 December 2022), which establishes a monitoring and cooperation mechanism to achieve common objectives for Europe's digital transformation by 2030, entered into force on ***9 January 2023***.

sustainable and durable digital transformation and which indicates to citizens that *European values, as well as their rights and freedoms must be respected in the online environment as well*.

The *Declaration* has 6 chapters with ***rights and principles*** which, in short, mean:

• affordable and high-speed digital connectivity everywhere and for everyone;

• well-equipped classrooms and digitally qualified teachers;

• problem-free access to online public services;

• a safe digital environment for children, disconnection after work and class hours;

• obtaining easy-to-understand information regarding the impact of digital products on the environment, and

• control over how personal data is used and to whom it is sent.

At the same time, this *Declaration* establishes ***the concrete digital objectives*** that the EU and its member states aim to achieve by the end of the decade **in four areas**, namely:

• strengthening digital **skills** and digital education;

• the development of secure and sustainable **digital infrastructures**;

• digital transformation of **enterprises**;

• digitization of **public services**.

The Policy Program:

• introduces a new form of governance based on **cooperation between Member States and the Commission, to pool in common the EU, national and private resources** and to make **progress in digital capabilities and technologies** that no Member State would otherwise not be able to make on its own;

• will facilitate **investments** in areas such as high-performance computing, shared data infrastructure and services, blockchain technology, low-power processors, pan-European development of 5G corridors, high-tech partnership for digital skills, secure quantum infrastructure and hub network of cyber security, digital public administration, test facilities and digital innovation centres.

In this sense, Member States will develop draft of **national trajectories** and **strategic roadmaps** to achieve these objectives until their ***expected revision in 2026***. Progress will be monitored based on the **Digital Economy and Society Index (DESI)** [5][5] and will be evaluated in ***the Commission's annual report*** on the progress of the digital decade.

This policy resides from the **"Compass for the Digital Dimension 2030: The European Model for the Digital Decade"** of **9 March 2021** and comes more than 2 years after that the financial organizations accelerated their ***digitization of the financial services sector*** by adopting **"Cloud Computing"** [6][6] which improved the ability to store data and record data, enabling better access and management of customer information. In this context, the **Cloud Computing technology has helped these organizations** through the global pandemic, economic crisis and other emerging challenges, with the aim of:

• expanding digital services in the financial services sector (such as online banking and instant and contactless payments);

• offering personalized, relevant and responsible solutions to consumers, to the requests made through electronic channels, so that the interaction with them is carried out without interruptions and,

[5] Since 2014, the European Commission has monitored Member States' digital progress and published annual reports on the Digital Economy and Society Index (DESI). Romania ranks 27th out of the 27 EU member states and in the 2022 edition also.

[6] The term **"cloud computing"** refers to the storage, processing and use over the Internet of data that resides on remotely located computers. Many people today use the cloud without even realizing it. Existing services such as Internet e-mail or social networks have their technological base stored in the cloud. "*Cloud computing*" offers professional IT users a high degree of flexibility in terms of computing power required. For example, if the use of a service increases, it is very simple to add additional capacity - an operation that would require increased additional costs and much more time if the company had to physically install a new computer in its own data center.

at the same time, to remain competitive on the market and to optimize the activity in the relationship with consumers.

In other words, the **Cloud offers increased agility** and the ability ***to create resilient, real-time applications that are available and able to scale quickly on demand***, which is **an advantage for the digitalization of the financial sector**. At the same time, enterprises can achieve higher levels of security of the platforms used by ***automating secure infrastructure and newer technologies***, such as for continuous monitoring of activities, ***based on security and compliance controls, while reducing human configuration errors***.

This trend will help financial institutions maintain the confidentiality and integrity required by their customers, while ensuring ***the timely and accurate reporting required*** by financial services industry regulators and supervisors, as well as ***managing operational risks*** in their cloud environment and ensuring that they have identified sufficient security processes and measures ***to support the encryption, authentication and reporting of collected data***.

This is also crucial, especially now at a time when **cyber threats are on the rise**, including from hostile states, **with potential implications for financial stability**. As the European regulatory framework evolves, ***integrated network services*** (e.g. *Amazon Web Services - AWS*) will continue to provide the financial services industry with the most advanced ***data control, security and privacy capabilities***, empowering the EU financial sector with additional options to meet its needs, bringing ***innovation, security and resilience benefits to the sector*** and being ***essential for growth, economic development and global competitiveness***.

Thus, the **era of digital finance**:

• offers, in turn, the opportunity to create ***a competitive digital economy*** of the European Union and an open society, ***centered on innovative financial products and services*** for consumers;

• transforms traditional value chains and gives rise to ***new business models and new enterprises***;

• ensures more competitive, more accessible and more inclusive financing, facilitating the emergence of ***new FinTechs*** and ***tempting applications*** that offer consumers cheaper, simplified and faster access to financing;

• marks advances in ***Artificial Intelligence (AI)*** that increase the speed and ability to analyze and evaluate the multitude of data collected to improve human decision-making and reduce the risks arising from the increased use of digital finance;

• foreshadows a significant impact on payment services through ***Distributed Ledger Technology (DLT)*** which is a distributed technology, intended to decentralize the recording of transactions and reduce the need for expensive intermediate chains, and in its mode of operation and resilience is ***similar to the technology of Internet file transfer torrent*** [7][7]. Thus, the same information is stored in multiple locations, and if one location is affected, the other locations separately certify active transactions/data. In this way, the technology is resilient to fraud and computer system malfunctions. At the same time, ***blockchain technology ensures the confidentiality of transactions without human intervention to protect personal data and thus minimize possible legal consequences***.

[7] ***BitTorrent*** is a peer-to-peer data transfer technology, one of the best known and most used technologies of this type.
Peer-to-peer (P2P), loosely translated from peer to peer, is a network architecture for distributed applications that divides tasks among multiple partners.
Peer-to-peer networking allows computers to connect directly to each other for mutual file exchange (file sharing). There is no theoretical limit to the size of a peer-to-peer network, those can be consist of two or hundreds of computers. Examples of P2P networks: *BitTorrent, eDonkey, Gnutella, FastTrack, ANts, Kazaa, BearShare, Direct Connect or Limewire.*

3. Accelerating Digitalization of the Financial Services Sector: Trend or Opportunity?

The war, the energy crisis, the disinformation, the food crisis, the climate changes, the underestimation of the role of industry are the ***main factors that completely undermine*** the foundations of the democratic world and ***the free market economy***.

To these is added **the impact of the informational warfare** (visible through the public outing and the mobilization made on social networks, ***affecting democracy as a system***), as well as, resuming the idea from the first paragraph of this paper, **the impact of overlapping and interdependent crises** that is beginning to be seen, especially from ***the perspective of daily costs***, which are increasingly high and ***reflect the reduced resilience of consumers to shortages and to a harder life.***

In this context [8], innovation facilitates the emergence of ***new products, processes or business models that are possible thanks to digital technologies***, while the **information technology systems**, combined with the corresponding software, have **become a central pillar of economic activities** for many enterprises. According to assessments carried out at the level of the European Union, this would be due to the fact that ***digitization offers substantial new opportunities***, as *digital networks and data services generally facilitate economies of scale* [9][8], allowing *the provision of better quality services at a lower cost*.

Thus, **innovation cycles accelerate**, becoming more open and collaborative. At the same time, digital technologies and applications are increasingly being built in a modular manner, communicating between them through application programming interfaces (APIs) [10][9]. Those allow ***a better adaptation of services to the client's demand*** and also offer ***more opportunities for experimentation and collaboration*** between various actors. This can have a number of consequences for the way that the financial services are provided.

On the other hand, it is equally important to pay increased attention to ***the risks and challenges that digitization brings***, among which we mention:

• the risks of increased exposure to **potential cybercrime**, **operational resilience** failures and **data protection and privacy issues**;

• the challenges of addressing issues of **accountability and transparency**, as well as those of market concentration with **a potential over-reliance on third-party suppliers**;

• **the adoption/adaptation of the financial regulatory and supervisory framework** from the perspective of the digitalization of the financial sector and the development of EU public policy, in accordance with *the European Commission's 2020 Digital Finance Strategy* [11] *and the new key legislative initiatives on crypto-assets (MiCA)* [12] and the strengthening of standards of resilience of digital operations (DORA) [13], adopted in recent months at the level of the European Union.

The draft of ***Regulation on crypto-asset markets and amending Directive (EU) 2019/1937 (MiCA***), for which, on **October 5, 2022**, the Committee of Permanent Representatives (Coreper) approved the provisional agreement, thus **starting the formal adoption process** will regulate the conditions for authorization and operation of crypto-asset issuers and crypto-asset service providers on the EU single market and will be the first step towards updating the European legal framework in this area. In strict correlation with MiCA, should also be considered the proposal for a pilot regime for market infrastructures based on distributed ledger technology near the *Regulation (EU) 2022/858 of the European Parliament and of the Council on a pilot regime for market infrastructures based on*

[8] **Economies of scale** is a term that describes what happens when the quantities of factors used in production increase. More specifically, an enterprise reduces its unit costs by producing more goods or services and, as production increases, average costs decrease by spreading fixed costs over a larger output.

[9] An **application programming interface (API)** is a set of programming code that queries data, parses the answers, and sends instructions between one software platform to another. APIs are widely used in the provision of data services in a wide range of domains and contexts.

distributed ledger technology and as well amending the Regulations (EU) no.600/2014 and (EU) no. 909/2014 and Directive 2014/65/EU (referred to as **the DLT Regulation**) [14].

4. Digital Financial Legislation: Framework for Balancing Innovation and Consumer Protection - RegTech, SupTech and RiskTech

The new legislative proposals issued at the level of the European Union are part of ***a set of future directions regarding the reformulation of public policy***, according to a new approach determined by the tendencies to "*open*" the public policy process. This tendency to open up this process is associated with ***the new developments that are based on the emergence of blockchain technology***[10], **to increase the level of trust between the citizen and the public authorities**.

The possible **causes of the unsatisfactory results of the reforms** regarding the formulation of public policies of the European Union, including in our country, can be associated with ***the existence of a very low level of communication and collaboration*** between different actors involved in the process of public policies, respectively ***a low degree of trust given to the administrative center***, in terms of its ability to fulfill its role of managing the public policy process.

Taking these aspects into account, the question that persists today is: *what* ***kind of change could this technology bring about***, especially in ***the provision of public services*** and *the development of public policies*?

This way of organization - ***the distributive organization*** - shows the relationship that is established between different actors in the process of services and public policies and the way in which the interaction between them **no longer depends on the involvement of a center with full power, nor on secondary local centers**. The distributive organization, thus, offers ***the possibility of a horizontal collaboration between these actors***, respectively of the citizens/consumers between them, changing the role and objectives of any old type of center and, as a consequence, its attributions and functions.

Blockchain technology can offer **the possibility of forming links** of this type, by ***developing a viable solution to the problem of trust*** between the different parts of some collaborations carried out ***in order to obtain policy results or public services***. In addition, this type of organization allows obtaining a high level of quality of these results, respectively of public services, so that centralized authorities (central or local level) could become less relevant, and their role could move towards ***to providing/managing a platform in order to facilitate distributive organization***, rather than being more or less at the center of any initiative.

Without going into technical details, the assessments carried out by the World Bank's specialist analysts [15] reveal that **the regulatory approaches** observed in different jurisdictions can be broadly grouped into:

• ***applying existing regulatory frameworks*** to new business models, focusing on the core economic function (e.g. regulating digital currency exchanges as businesses or money service exchanges);

• ***adjusting existing regulatory frameworks*** to accommodate the redesign of existing processes and enable the adoption of new technologies (e.g. minor changes to allow banks to only operate

[10] The technical concept of ***blockchain*** can be understood as a growing list of records (information) that form a block. The blocks are linked together forming a "chain" which actually represents accumulated information. Links between records are made using cryptography. According to it, each block contains a cryptographic hash (algorithm, model) of the previous block, a time stamp and the transaction data. The main advantage of this technology is that, thanks to the technical characteristics offered by the blocks and the links between the different records that form them, it becomes impossible to fraudulently change the security of the information once recorded. This advantage is essential, because it technically solves a problem similar to the one that underpinned the emergence of contracts, including the one that underpins the older perspective of the citizen-state relationship as one of client-supplier.

digitally - digital banks or neo-banks, use of digital forms of identification to open accounts and enable the adoption of cloud computing for outsourced banking services);

• ***creating new regulations*** to expand the regulatory perimeters and introduce specific requirements for the new class of players in the ecosystem (for example, creating a new class of regulated entities for electronic money and lending platforms, requiring bank providers to provide interfaces of application programming - API - to allow other institutions to directly access information and provide services to customers - open banking);

• ***adopting new frameworks to promote innovation and experimentation*** in areas where the regulatory framework is unclear or omitted. These frameworks include developments such as regulatory sandboxes, innovation, hubs and accelerators. Regulatory sandboxes are structured to allow experimentation with restrictions imposed on a large scale, duration and scope, to reduce risk while enabling the deployment of new technologies and approaches. Experience gained from regulatory sandboxes can then be used to structure the regulatory framework.

Innovation hubs created around the world seek to expand on this latter practice, ***allowing innovators to interact directly with regulators and the financial services industry*** through experts **to help innovation** in general. Business accelerators also seek to direct funding funds **to help develop and systematize research-development-innovation (RDI)** activities and bring new innovations to market. Regulatory sandboxes have captured the attention of several jurisdictions, including income economies (e.g. Australia, Hong Kong, Japan, the UK and the US) and World Bank client countries (e.g. China, Colombia, India, Indonesia, Jordan, Mexico and Morocco).

4.1. Regulatory Technology (RegTech)

The World Economic Forum calls our current environment "***The Fourth Industrial Revolution***" [16][11].

The new technologies emerge and evolve, creating **new challenges for regulatory structures** that strive ***to protect customers while embracing innovation***.

The digital world still seems far away, and regulators are working hard and hard **to create agile legislation and effectively set the rules to follow**. Organizations are also working further to implement processes and systems that enable them to monitor and respond to these changes in the regulatory framework.

Thus, **regulatory technology (RegTech)** [17][12] is ***a technological solution designed to streamline the regulatory compliance process*** within a bank, credit union or other financial institution. Financial institutions, and not only them, are often overwhelmed by the large volume of laws, rules and regulations (national and European) that they have to implement, comply with and supervise, given the complex spectrum of their own governance and business continuity, supplier management, fair lending and/or cyber security, etc.

RegTech solutions, depending on the complexity of processes and activities, take many forms, and can be ***designed to render compliance and risk*** to those in real time, while ***compiling a specific rule targets a specific domain***.

It is well known that the regulations do not only provide "black and white" rules.

Financial supervisors give institutions the freedom **to develop risk management and compliance programs** that are appropriate for their size and complexity, but without these ***programs being one-size-fits-all***.

---

[11] On **October 10, 2016**, the World Economic Forum (WEF) announced the opening of its new ***Center for the Fourth Industrial Revolution*** in San Francisco that will "*serve as a platform for interaction, insight and impact on the scientific and technological changes that are changing the way we live, work and relate*".

[12] **RegTech** is defined as "*the application of various new technology solutions that assist highly regulated industry stakeholders, including regulators, in establishing, performing and fulfilling governance, reporting, compliance and regulatory risk management obligations".*

From this perspective, **the best RegTech solutions**:

• combine automated, cloud-based software with the expertise and services of information and communication technology (ICT) experts, being an advantage if they also have years of experience in interpreting the nuances and subtleties of regulations and their implementation in internal processes, focusing on the "*overall picture*", identifying and examining the interaction of different types of risks throughout the institution, in order to increase institutional efficiency;

• allow an institution to better understand the challenges regarding the regulatory framework and organization of the basic activity, so that it can exercise the assigned powers and implement the necessary resources (human, material, financial) more efficiently, with predilection in critical areas, instead of use a dispersed approach, inconsistent and dissociated from the object of activity and the interests of the institution.

In this context, ***regulatory technology (RegTech)*** is an emerging technology that **involves the implementation of digital tools and processes** that improve the way organizations manage their growing regulatory compliance commitments. As it is a new technology area, its development embraces cutting-edge technology elements including big data analytics, machine learning, blockchain (distributed ledger technology), Internet of Things (IoT), artificial intelligence (AI) and more others, providing a wide range of capabilities such as:

• monitoring the regulatory framework;
• assessment of risks of (over)regulation and inconsistencies;
• compliance monitoring;
• knowing and monitoring clients;
• monitoring against money laundering;
• tracking and reporting;
• transaction aggregation and reporting.

The benefits of RegTech and the optimized results that can be obtained through this technology, aim at:

• reporting outsourcing;
• high data quality;
• operational resilience;
• privacy versus surveillance;
• prevention of money laundering.

The objectives that organizations must meet to implement RegTech are:

• increasing internal efficiency;
• effectiveness and applicability of regulations;
• reducing non-conformities by regulating processes and activities.

4.2. Surveillance Technology (SupTech)

Just as financial institutions are responsible for complying with thousands of rules and regulations, supervisors are tasked with ensuring that **all of these rules and regulations are followed** by financial institutions.

Supervisory Technology (SupTech) [18][13] is ***the technological solution designed to help financial supervisory authorities*** to ensure regulatory compliance and adequately manage risks identified in the financial system and regulatory enforcement by adopting some innovative

[13] **SupTech** "*refers to the use of technology to facilitate and improve supervisory processes from the perspective of supervisory authorities*". ***Regtech*** and ***Suptech*** solutions are emerging for a wide range of regulatory areas including regulatory change tracking, fraud detection, know your customer (KYC), countering the financing of terrorism (CFT), conduct and prudential risk management, regulatory reporting and associated audit trail.

technologies (such as Artificial Intelligence - AI and Machine Learning – ML) by regulatory authorities, called supervisory agencies, to support supervision.

SupTech **helps regulators**, i.e. supervisors and representatives of other regulators of financial institutions and other industries (including financial services), become ***more efficient, automated and reduce costs and errors***. Like RegTech solutions, SupTech focuses on **maximizing the efficiency of supervisors** by ***automating*** processes, ***optimizing*** operational and administrative operations and ***digitizing*** work tools and data, with the aim of **reducing the reporting burden of enterprises** and **promoting improved reporting, more prompt monitoring and overall compliance** with existing regulations and resources to oversee.

In **the case of data analysis**, SupTech solutions can minimize specific problems by ***automating the process of collection and storage***, but only if they are transmitted in ***defined reporting formats***, to facilitate the collection and evaluation of data and allow their ***unified interpretation*** (including aspects such as market surveillance, deviation analysis and macro and micro prudential supervision).

In **the case of automated reporting**, activities include ***data push*** **vs** ***data pull*** where entities being supervised push data through M2M APIs[14] or regulators can programmatically pull data directly from supervised entities and also include ***real-time monitoring***.

In **the case of virtual assistance**, AI ChatBots (chatbots using Artificial Intelligence) are used to ***address consumer complaints*** and ***assist regulatory bodies*** in reviewing and improving regulations.

While SupTech presents a set of new ways to make regulators' operations more accurate and faster, that doesn't mean that the technology does not come with its own set of challenges, namely: cyber risk, legal risk, operational risk, IT risk.

4.3. Risk Technology (RiskTech)

Bankers and credit union executives often **worry about not complying with regulatory requirements** and envision how regulators find specific violations of statutory rules, ***worrying about the cost of non-compliance***. Thus, they want to solve the problem as simply (and cheaply) as possible.

In this sense, **RiskTech** [19], as a subset of **Insurance Technology (InsurTech)**[15], constitutes another new technological innovation in insurance that is intended to support risk professionals to make better and more efficient decisions, based on data analysis.

RiskTech solutions focus on **the big picture**, identifying and examining ***the interaction of different types of risk*** across the entire enterprise. Those confirm that risk remains an integral part of all discussions, regardless of the segment/area of supervision. Those, also, enable an institution **to better understand and prioritize its risk management needs** so that it can more effectively ***deploy resources in the most critical areas***, rather than using a scattered approach.

[14] **M2M** (Machine 2 Machine) **API** (Application Programming Interface) refers to a **passive interface** for electronic communication between two remote devices or remote software. Unless otherwise specified (for example, in the cloud-facing MQTT protocol), the device waits until it receives a request for a value from an active element (a master system).

[15] **RiskTech** consists of **InsurTech**-related tools that are specifically designed for risk management professionals. As risks continue to evolve and multiply, risk managers will come to rely on RiskTech tools and bigger data-driven insights to chart a risk's trajectory.

InsurTech refers to technological innovations that are created and implemented to improve the efficiency of the insurance industry and encourage the creation, distribution and administration of insurance business as the insurance industry is well poised to take advantage of disruptive and innovative technology. InsurTech helps large insurance companies explore new options outside of traditional human endeavours. This could include dynamically priced insurance policies, small business insurance and social insurance options. InsurTech also gives insurance companies access to data streams from IoT (Internet of Thinks) devices. This creates a dynamic pricing system based on market conditions and customer behavior.

5. Epilogue

While the "***technological transition***" improves the professional experience and confers undoubted benefits in terms of ***accuracy and processing time of data and information collected and processed***, both for providers, regulators or supervisors and for consumers, there are still massive restraints in the industry domestic financial services that minimizes **the importance and consequences of a pro-active attitude towards the digitization of finances**.

Although the slowest process is manifested at the level of regulatory agencies and supervisory authorities, and even there are still conservative, ***the technological solutions envisioned denote that automation, streamlining and increasing the efficiency of automation are increasingly necessary***. The sheer volume of regulations (some of which are unenforceable, out-of-date or inconsistent), remote working environments and the need for ongoing risk management add to **the challenge for supervisors and regulators** to exercise their powers and modernize and optimize their way of working.

On the other hand, in the spirit of the predictions of the executive director of the World Economic Forum, Klaus Schwab (creating the conditions for a "*stakeholder economy*", building in a more "*resilient, equitable and sustainable*" way, using environmental, social and of governance (ESG) parameters and "*harnessing the innovations of the fourth industrial revolution*") [20], **to reduce the time gap with economically and technologically advanced countries**, we appreciate that quick and ***effective solutions should be implemented***, through specialized working groups, to outline ***realistic measures to adapt the behavior of financial businesses and consumers*** to the new conditions and to manage the evolving changes as much as possible.

References

[1]. Aurelian Dochia, “The five transitions that change the world (and produce chaos along the way)”, [Online]. Available: https://www.contributors.ro/cele-cinci-tranzitii-care-schimba-lumea-si-produc-haos-pe-parcurs/.

[2]. REPORT on the proposal for a decision of the European Parliament and of the Council establishing the policy program for 2030 "The Road to the Digital Decade", [Online]. Available: https://www.europarl.europa.eu/doceo/document/A-9-2022-0159_RO.html.

[3]. "The Path to the Digital Decade", [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:32022D2481&from=EN#d1e40-4-1.

[4]. "Declaration on Digital Rights and Principles in the European Union", [Online]. Available: https://digital-strategy.ec.europa.eu/ro/library/european-declaration-digital-rights-and-principles.

[5]. “Digital Economy and Society Index (DESI)”, [Online]. Available: https://ec.europa.eu/newsroom/dae/redirection/document/88758, p.3.

[6]. “Harnessing the potential of cloud computing in Europe - what is it and what does it mean for me?”, [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/ro/MEMO_12_713.

[7]. Bittorrent Technology, [Online]. Available: https://www.techtorials.ro/2010/05/05/tehnologia-bittorrent/.

[8]. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the EU Strategy on Digital Finance, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri= CELEX:52020DC0591&from=EN.

[9]. Economies of scale, [Online]. Available: https://ro.wikipedia.org/wiki/Economie_de_scar%C4%83.

[10]. An application programming interface (API), [Online]. Available: https://www.techtarget.com/searchapparchitecture/definition/application-program-interface-API.

[11]. The entire package include:
1) EU Cyber Security Strategy for the Digital Decade, JOIN(2020) 18 final, 16.12.2020, [Online]. Available: http://www.cdep.ro/afaceri_europene/CE/2020/JOIN_2020_18_RO_ACTE_f.pdf.
2) Directive related the Critical Infrastructure Resilience, 16.12.2020, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:52020PC0829&from=EN.
3) Directive (EU) 2016/1148 on the Security of Networks and Information Systems, [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=CELEX:32016L1148&from=IT.

[12]. "The Regulation on markets for crypto-assets (MiCA)", [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0593.

[13]. Regulation on increasing the digital operational resilience of financial institutions (DORA), adopted by the European Commission on 10 November 2022 and intended to strengthen the operational resilience of the financial sector, [Online]. Available: https://www.europarl.europa.eu/doceo/document/TA-9-2022-0381_RO.html, in addition to NIS2, https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52020PC0595.

[14]. Regulation (EU) 2022/858 of the European Parliament and of the Council on a pilot regime for market infrastructures based on distributed ledger technology and as well amending the Regulations (EU) no.600/2014 and (EU) no. 909/2014 and Directive 2014/65/EU (referred to as the DLT Regulation), [Online]. Available: https://eur-lex.europa.eu/legal-content/RO/TXT/HTML/?uri=CELEX:32022R0858&from=RO and https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52020PC0593.

[15]. The World Bank, Fintech and the Future of Finance (18.05.2022), [Online]. Available: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099450005162250110/p17300600228b70070914b0b5edf26e2f9f.

[16]. World Economic Forum, The new forum center for promoting global cooperation on the fourth industrial revolution, [Online]. Available: https://www.weforum.org/press/2016/10/new-forum-center-to-advance-global-cooperation-on-fourth-industrial-revolution/.

[17]. What is RegTech and what does it mean for policy makers?, [Online]. Available: https://www.weforum.org/agenda/2022/06/what-is-regtech-and-what-does-it-mean-for-policymakers/.

[18]. World Bank,The Next Wave of Suptech Innovation: Suptech Solutions for Market Behavior Surveillance, [Online]. Available: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/735871616428497205/the-next-wave-of-suptech-innovation-suptech-solutions-for-market-conduct-supervision and https://documents1.worldbank.org/curated/en/735871616428497205/pdf/The-Next-Wave-of-Suptech-Innovation-Suptech-Solutions-for-Market-Conduct-Supervision.pdf.

[19]. The future of risk management will be determined by risktech, [Online]. Available: https://www.propertycasualty360.com/2019/12/10/rims-the-future-of-risk-management-will-be-determined-by-risktech/.

[20]. The Great Reset, [Online]. Available: https://en.wikipedia.org/wiki/Great_Reset.

A FMEA Analysis on Web Applications

Gabriel PETRICĂ[1], Costel CIUCHI[2]
[1] EUROQUALROM, Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania
gabriel.petrica@upb.ro
[2] Associate Professor, University POLITEHNICA of Bucharest, Romania
costel.ciuchi@upb.ro

Abstract

Based on the Failure Mode and Effects Analysis (FMEA) method, this paper identifies the potential causes that lead to the failure of a Web application built on the WordPress platform. Both software vulnerabilities identified in the U.S. National Vulnerability Database (NVD) and other platform administration and configuration processes that can be exploited in cyber-attacks against the Web application are considered. Finally, measures to eliminate potential security breaches are proposed in the form of a best practice guide for managing sensitive data and increasing the level of security for this type of application.

Index terms: cybersecurity, FMEA analysis, software vulnerabilities, WordPress

1. Introduction. Content Management Systems (CMS)

Currently and conventionally, the term "*content*" is used to refer to information in the category of text (documents in various formats), audio, video, binary files or any other type of media, transmitted electronically through traditional systems or via the Internet. In the latter case, specifications for media types - Multipurpose Internet Mail Extensions (MIME) - are managed by the Internet Assigned Numbers Authority (IANA), the official authority for standardizing and publishing these specifications. The content can be produced, modified, transmitted, consumed, or traded in parts or in its entirety, being available on demand, possibly under certain conditions, and accessible permanently or during certain periods. In the context of the media industry, a team representing the Society of Motion Picture and Television Engineers (SMPTE) and the European Broadcasting Union (EBU) defined the term "*content*" in 1998 and identified its two components: *the essence* and *metadata* [1] (Figure 1).

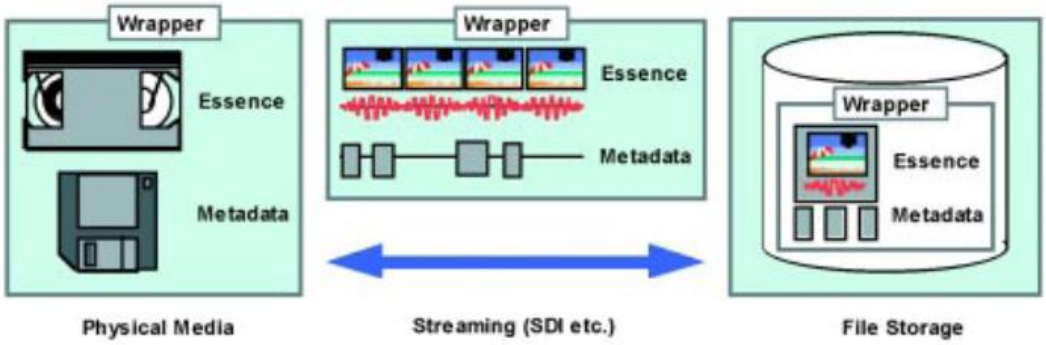


Fig. 1. The "content" elements: essence and metadata [1]

The *essence* refers to the raw material of the program itself - represented by text, images, sounds, video and others. The *metadata* is the part that characterizes the essence and other attributes of the content. Metadata describes the actual content or subject matter, material (available formats, encoding parameters, and specific recording information) or location.

A system that deals with content and metadata management is called a *Content Management System* (CMS). It allows an organization to manage information in real time, provides up-to-date content and respond to changing consumer demands. CMS provides automated control of information collection, management and distribution [2]. The need for a Content Management System within an organization is supported by the following factors:

- large amount of content.
- multiple access.
- finding information on different sites of the same organization, in a format adapted to the type of channel.
- varied content, which changes in a dynamic way.
- personalized content, reflecting the way each organization interacts with its consumers.
- multiple authors, contributors, and publishers.
- systems for recording workflows and managing tasks between teams.
- the need for flexibility in entering and processing data.

There are many CMSs and the most popular of them is WordPress (according to W3Techs statistics from April 2023, 63.3% of monitored websites used the WordPress platform, which represented 43% of all websites [3] - Figure 2).

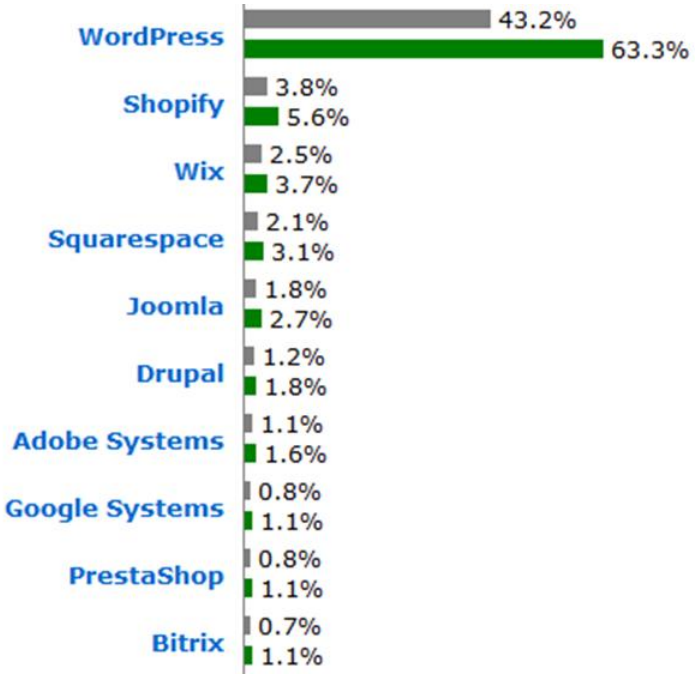


Fig. 2. Content Management System usage statistics - April 2023 [3]

Launched in 2003, WordPress is the most widely used open-source CMS worldwide, with approximately 18 million installations. Started as a blogging system, WordPress has now become a fully functional Content Management System [4]. WordPress has many free and varied templates, easy installation using a wizard, easy-to-search URLs, and management and publishing tools for mobile solutions. As the most popular CMS, WordPress is the most common target of attackers. However, WordPress is built on a very secure code and responds quickly to security vulnerabilities. It also has an auto-update mechanism that allows the system to automatically update when there are new versions.

2. Analysis of failure modes for a Web application based on the WordPress platform

Failure Modes and Effects Analysis (FMEA), an effective method in systems reliability, maintainability, security, and testability studies, involves the exhaustive enumeration of possible failure modes for all system components and highlights the effects of these failures at the component or (sub)system level [5]. FMEA is a tool that helps deliver products or processes that are reliable, acceptable to the customer and, above all, safe to use. Because FMEA helps the designer identify potential critical product/process defects, it is used to:

- develop product or process requirements that minimize the probability of failures.
- evaluate the requirements obtained from the clients or other participants in the design process to ensure that these requirements do not introduce potential defects.
- identify those design features that contribute to the occurrence of critical defects and minimize the resulting effects.
- designs methods and procedures to develop and test the product / process so that there is certainty that defects have been successfully eliminated.
- track and manage potential risks in design.
- ensure that any defects that may occur will not affect or will not have a serious impact on the user of that product / process [5].

In the next chapters we developed a FMEA analysis by identifying and describing the main components of WordPress (Figure 3): *core* (the main code), *theme* (which represents a collection of files that change the appearance of the website, and the way information is presented, keeping its content unchanged), and *plugins* (PHP code sequences that extend the default functionalities of the platform).

Fig. 3. The software components of a WordPress platform

2.1. Exploitation of the software vulnerabilities within the core of WordPress

Considering recent statistics (April 2023) showing that the most used versions of WordPress are currently v.6 and v.5, with over 90% of installations [6] (Figure 4), we can consider vulnerabilities in these versions as the most likely to be exploited by the current potentially cyber-attacks. Released in December 2018, WordPress 5.0 “Bebo” introduced a new editor based on blocks (abstract elements that make up the layout of a page) and a new default theme called “Twenty Nineteen” [7]. The new WordPress version 6, called “Arturo”, was released to the public in May 2022 and enhanced several aspects like performance, accessibility, or design tools and added a better writing experience [8].

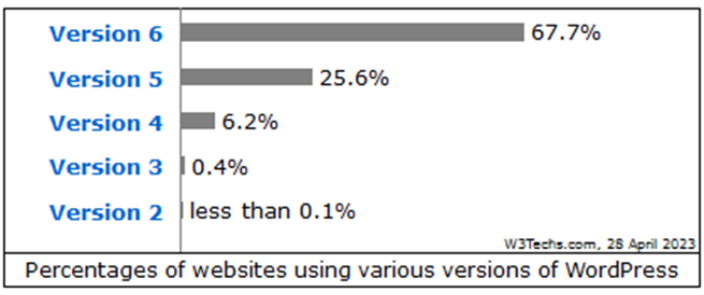


Fig. 4. WordPress versions usage statistics [6]

The existence of different ways of identifying and scoring vulnerabilities and the lack of interoperability between databases and tools that referred to the same vulnerabilities led to the emergence of the Common Vulnerabilities and Exposures (CVE) project in 1999. Designed and coordinated by the MITRE Corporation, CVE is a tool for monitoring and standardizing the most known vulnerabilities, which ensures trust between parties when used to discuss or share information about a unique vulnerability of an application, operating system, service, or firmware [9]. Each known vulnerability is uniquely identified in the National Vulnerability Database (NVD) and is also available in the CVE list, with detailed descriptions for each vulnerability, as well as a system - the Common Vulnerability Scoring System (CVSS) - that quantifies (on a scale from 0 to 10 - maximum severity) the impact of that vulnerability [10]. Software vulnerabilities reported for the WordPress platform from 2019 to 2022 are summarized in Table 1.

Table 1. WordPress vulnerabilities reported in 2019 - 2022 [11]

Year	# of Vulnerabilities	DoS	Code Execution	SQL Injection	XSS	Directory Traversal	Bypass something	Gain Information	Gain Privileges	CSRF
2019	23		4		12	1	2	2		2
2020	21	1	2		7				2	1
2021	8		1		2		2	2		
2022	9			2	3		1			

The 61 vulnerabilities of WordPress version 5.x and 6.x (from 2019 to 2022) have the CVSS score distribution shown in Table 2, having an average score of 5.3 [12].

Table 2. CVSS vulnerability score for WordPress v5.x and 6.x (2019 - 2022) [12]

CVSS score	Number of vulnerabilities	Percent
0-1	4	6.60%
1-2		0
2-3		0
3-4	13	21.30%
4-5	18	29.50%
5-6	11	18%
6-7	6	9.80%
7-8	9	14.80%
8-9		0
9-10		0

2.2. Exploitation of the vulnerabilities in themes and plugins

The two components of a WordPress platform, *themes* and *plugins* implemented by third parties, bring different ways of displaying information, respectively implement new functionality starting from the core code. However, these components are also the most vulnerable, being most of the time exploitable resources by cyber attackers. They typically use automated scripts to scan the Internet for websites that contain known software vulnerabilities. When a target is identified, malware code is executed or can be injected to gain unauthorized access to the compromised environment. The attacker then deploys software tools, depending on available resources, to launch new attacks on other targets.

According to SUCURI statistics from 2022, the ten most frequent vulnerable software components are indicated in Table 3. It is noted that “36% of all compromised websites had at least

one vulnerable component present in the environment at the point of remediation" [13]. However, the data does not necessarily indicate that these plugins were attack vectors, but instead contributed to an overall insecure environment.

Table 3. Top software with vulnerabilities in 2022 [13]

Software Component	Percent
Contact-Form-7	27.44%
Fremius Library	20.85%
WooCommerce	14.51%
UpdraftPlus Free	5.35%
Gutenberg Temp. Library & Redux Framework	3.83%
Advanced Custom Fields	3.23%
WP Fastest Cache	3.21%
Essential Addons for Elementor	3.04%
PageBuilder by SiteOrigin	2.22%
File Manager	1.89%

The tools used by SUCURI in responding to various security incidents detected signatures of the malware application categories shown in Figure 5.

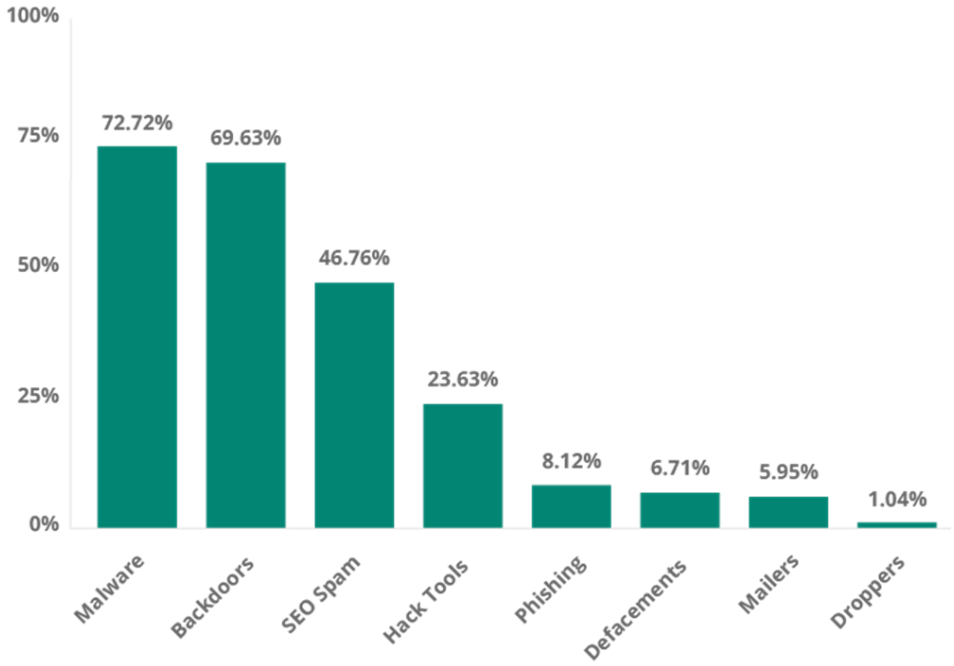


Fig. 5. Distribution of malware applications in 2022 (according to SUCURI) [13]

2.3. Exploitation of PHP or DBMS vulnerabilities

Another sensitive subject is represented by the versions of the PHP language installed on Web servers. Although SUCURI statistics indicate that over 65% of the analyzed websites used in 2021 PHP v7.x or a higher version (Figure 6), there remains a considerable number of those still using outdated versions of the 5.x series [14]. Among the motivations behind this situation, we can mention:

- the added code, themes and plugins used are incompatible with new PHP versions.
- some codes require partial / total rewriting, involving additional time and funds.
- many websites depend on the hosting company and the owners may have limited or no control over the PHP version.
- some owners simply neglect or do not want to update.

PHP versions that have reached EOL (End-of-Life) no longer receive regular security updates, their use being a vulnerability often exploited by attackers.

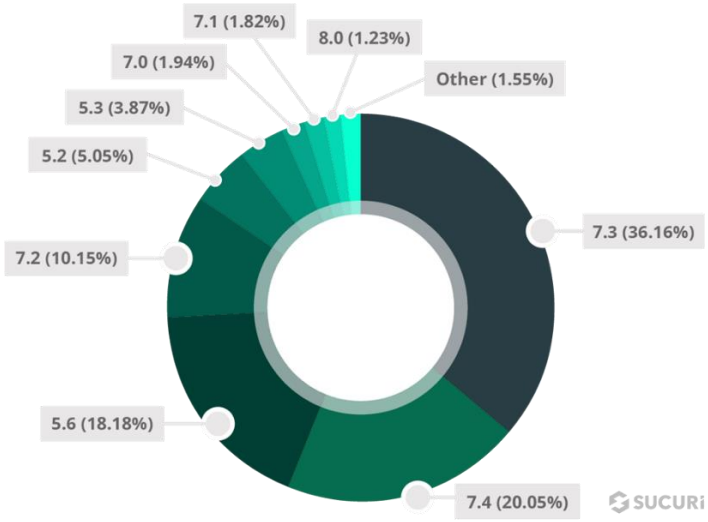


Fig. 6. PHP versions used in 2021 (according to SUCURI) [14]

Regarding Database Management Systems (DBMS), according to Statista, the most popular systems worldwide in February 2023 were Oracle, MySQL and Microsoft SQL Server (Figure 7) [15]. However, it has been found that DBMS systems are generally not delivered as a security-safe package, leaving administrators the task of configuring them optimally from this point of view. Typical vulnerabilities may refer to:

- vulnerable credentials (empty passwords, weak username / password combinations).
- activating some functions that are not necessary.
- insecure configurations, enabled for the convenience of administrators or developers.
- sensitive data stored or transferred in clear text format (not encrypted).

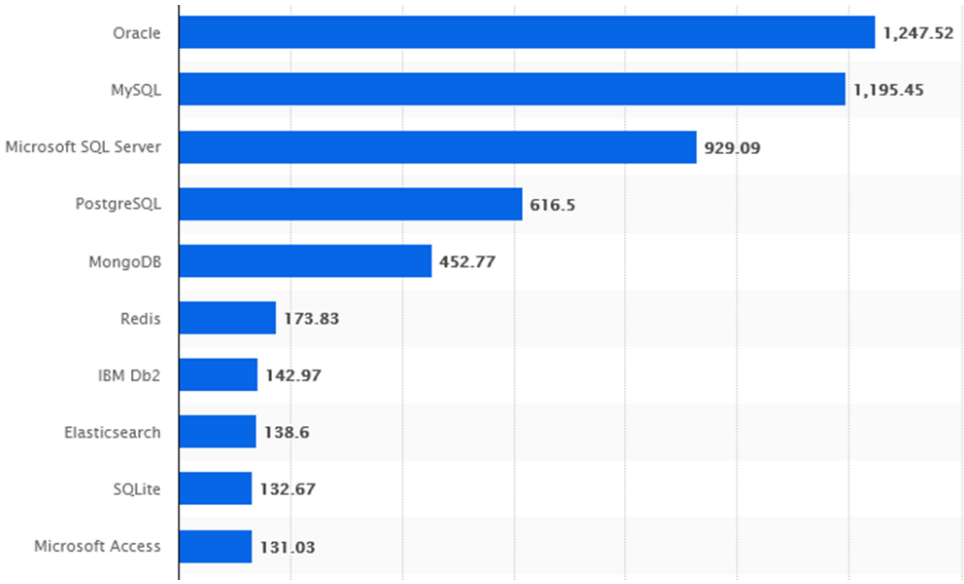


Fig. 7. Top 10 DBMS in February 2023, according to Statista [15]

2.4. Performing the FMEA analysis

Starting from the typical stages of a FMEA analysis and based on the aspects presented in previous chapters, we developed a FMEA analysis that addresses several types of processes (classes of actions) associated with a WordPress platform. In Table 4 the potential causes for each type of process are identified and the modes of manifestation / failure are described. The proposal of solutions that mitigate the impact and the attack surface will be made in chapter 3 as a good practice guide.

Table 4. FMEA analysis for a WordPress platform

Process	Potential causes	Effect description / Failure mode
Exploitation of the core components		
Exploitation of software vulnerabilities specific to the WordPress platform	Vulnerabilities in code (core)	- 61 vulnerabilities identified in the most used current versions (WordPress v.5.x and v.6.x) - the exploitation of these vulnerabilities (Table 1) allows the launch of some types of cyber-attacks with various unwanted effects on the platform
Exploitation of third-party software vulnerabilities (plugins / themes)	Use of vulnerable plugins/ themes	- Stealth code - BASE64 encoding function calls - Functions for displaying some information - e.g. phpinfo() - Running system functions (fopen, chown, chmod, exec...)
Exploiting specific vulnerabilities of the PHP language version	Using an outdated version of the PHP language (e.g. v.5.x)	- Execution of malicious PHP code - Exploitation of known PHP vulnerabilities
Exploitation of specific vulnerabilities in the database management systems	Using an outdated / insecure version of DBMS	- Corruption or loss of data - Loss of right of access - Extraction of sensitive data - passwords, Personally Identifiable Information (PII) - Disclosure of data to unauthorized parties - Interruption of the provision of some services
Other exploits		
Brute-force attacks	There are no mechanisms to limit brute-force attacks	- Valid username / password pairs are obtained (guessed) - Decreases the processing power of the server running the application
Access to sensitive files	Unsecured special file locations	Access to configuration files, installation scripts, or documentation files is allowed
Admin account exploit	The admin account (“admin”) is the default account	Obtaining the admin account password by brute-force attacks on a supposedly existing account (“admin”)
Table prefix	Default Table prefix (“wp_”)	The “wp_” default table prefix is an advantage for an attacker
Password exploitation	Weak passwords - insufficient password length	Illegal application access
	Weak passwords - insufficient password complexity	- Illegitimate access to user / application data - Leaks of personal data and confidential information
Password storage	- Unencrypted password in source code - Unencrypted password in properties / configuration files	Illegally obtaining admin password
	Unencrypted passwords in database	Illegally obtaining user passwords
Password management	- Remember password option is checked - Save password in the browser for subsequent logins	Illegal access to the application if run on a public computer
	Display error messages on login	Disclosure of incorrect elements (username or password) on an incorrect login

3. Proposed corrective actions. Good practice guide

Next, for each of the categories of processes identified in Table 5, corrective actions are proposed, which will lead to the reduction of the effect or the elimination of potential causes. Thus, Table 5 presents recommended actions for increasing the level of cyber security of a Web application based on the WordPress platform.

Table 5. FMEA analysis - proposed solutions

Process	Recommended corrective actions
Exploitation of the core components of the WordPress platform	
Exploitation of WordPress platform-specific software vulnerabilities	Update platform (core)
Exploitation of third-party software vulnerabilities (plugins / themes)	- Update plugins / themes - Deleting unused plugins / themes - Installing trusted plugins from safe sources - View user comments and ratings before installation
Exploiting version-specific vulnerabilities of the PHP language	Using the latest versions of the PHP language
Exploitation of specific vulnerabilities of the database management system	Regular application of patches
Other exploits of the WordPress platform	
Password storage and management	- Controls for a minimum password length - Warnings for weak passwords - Random password generator - Controls for a complex password - Reject typical content of passwords - Password strength indicator - Encryption of stored passwords - Re-authentication request before changing sensitive settings - Transmission of passwords using secure protocols (HTTPS) - Disconnecting inactive (idle) accounts - Mechanism for resetting the password if it has been forgotten - Imposing a periodic change of the password
Brute-force attacks	- Full validation of inputs - Limiting responses to queries (useful in case of a cyber- attack) - Limitation of login attempts for a user / IP address - CAPTCHA mechanisms - Deactivation of the XML-RPC protocol - Use of WAF (Web Application Firewall) - Implementation of Two-Factor Authentication (2FA) - Showing minimal information if login fail (for example: “Incorrect credentials”) - Remove WordPress or PHP version info - phpinfo() - Disabling or removing verbose debugging or error messages
Access to sensitive files	Rules in .htaccess (file used by the Apache Web server)
Admin account exploit	Creating an admin account with an unpredictable name, followed by deleting the default admin account (“admin”)
Prefix for tables	Changing the default prefix (“wp_”), at installation or later (manually or via a plugin)
Admin interface access	Restrict access to the CMS admin interface from approved or internal IP addresses

Other measures consist of “scanning the application to discover any vulnerabilities and fixing them as quickly as possible, but also applying security procedures in the development and maintenance life cycle of Web applications” [16]. For this purpose, dedicated tools can be used to scan for security vulnerabilities at the setup of WordPress (for example, WPScan) or later perform vulnerability assessments of code or plugin modules with third-party applications (such as RIPS, a static code analysis tool for automatically detecting security vulnerabilities in PHP applications).

A starting point in building secure, vulnerability-free web applications is the *OWASP Top 10*, an awareness document that has also been adopted as an industry standard to ensure cybersecurity. The Open Web Application Security Project (OWASP) community makes it easier for organizations

to develop, acquire, and maintain trusted applications and APIs (Application Programming Interfaces) by periodically publishing a Top 10 cybersecurity risks that developers need to be aware of. The latest OWASP statistics on security risks are displayed in Figure 8.

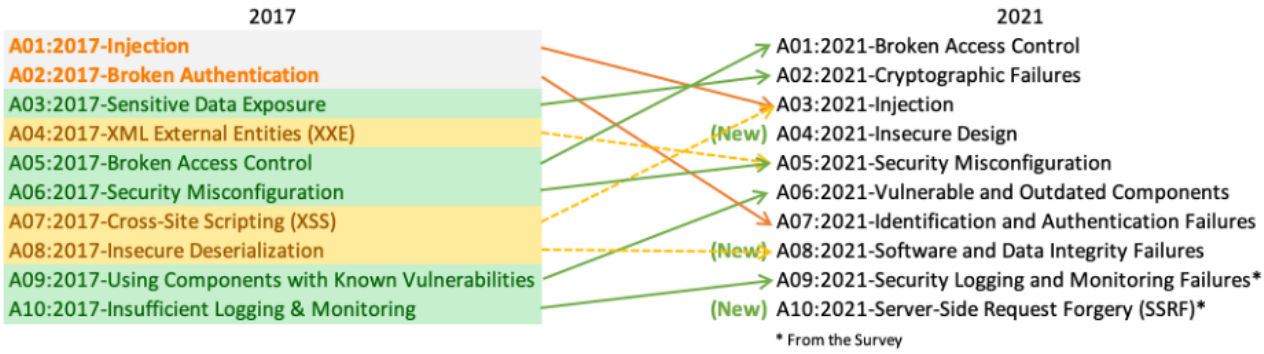


Fig. 8. The latest Top 10 Web Application Security Risks by OWASP [17]

4. Conclusions

Content management systems facilitate task management across teams and provide the context an organization needs to effectively create, update, and publish content. This paper has classified and identified the main causes (*failure modes*) that can cause a WordPress platform to stop working or that can compromise its security. The aspects and conclusions presented in the FMEA analysis are valid for any typical CMS structure, keeping in mind, however, that some elements (for example, vulnerabilities) depend on its type and version.

This analysis is a starting point in securing a Web application built on WordPress platform. Further developments of the study consist of extending the analysis to other software components of a complex Web application. Cyber-attacks and vulnerabilities can be related to the operating system, Web server or other applications components. Other situations that can disrupt the operation of a Web application may refer to the possibility of a physical attack in order to steal or destroy the equipment (access to the premises where the application runs or to the interconnection equipment).

Defending against various security threats is a continuous process that must involve awareness of the latest techniques and tools, correlated with the hardware and software environment in which the application operates, but also with user training, the weakest link in ensuring cyber security in an organization.

References

[1]. "Final Report of the EBU / SMPTE Task Force for Harmonized Standards for the Exchange of Television Programme Material as Bitstreams," 1998. Accessed: Mar. 15, 2023. [Online]. Available: https://tech.ebu.ch/docs/techreview/ebu-smpte-tf-bitstreams.pdf.

[2]. C. Benevolo, *Evaluation of Content Management Systems (CMS): a Supply Analysis*, 2017.

[3]. "Usage statistics of content management systems." W3Techs. https://w3techs.com/technologies/overview/content_management (accessed Apr. 5, 2023).

[4]. "CMS comparison 2022: The most popular content management systems." IONOS Digital Guide. https://www.ionos.com/digitalguide/hosting/cms/cms-comparison-a-review-of-the-best-platforms/ (accessed Apr. 2, 2023).

[5]. V.M. Cătuneanu and I.C. Bacivarov, *Fiabilitatea sistemelor de telecomunicații*, Ed. Militară, București, 1985.
[6]. "Usage statistics and market share of WordPress." W3Techs. https://w3techs.com/technologies/details/cm-wordpress (accessed Apr. 28, 2023).
[7]. "WordPress 5.0 Bebo." WordPress. https://wordpress.org/news/2018/12/bebo/ (accessed Mar. 20, 2023).
[8]. "WordPress 6.0 Arturo." WordPress. https://wordpress.org/news/2022/05/arturo/ (accessed Mar. 20, 2023).
[9]. "CVE - Common Vulnerabilities and Exposures." http://cve.mitre.org/ (accessed Mar. 10, 2023).
[10]. "NIST Common Vulnerability Scoring System Calculator Version 3." https://nvd.nist.gov/vuln-metrics/cvss/v3-calculator (accessed Mar. 10, 2023).
[11]. "WordPress: Vulnerability Statistics." CVE Details. https://www.cvedetails.com/product/4096/Wordpress-Wordpress.html (accessed Mar. 10, 2023).
[12]. "CVSS scores for WordPress between 2019 and 2022." CVE Details. https://www.cvedetails.com/cvss-score-charts.php?fromform=1&vendor_id=&product_id=4096&startdate=2019-01-01&enddate=2022-12-31 (accessed Mar. 10, 2023).
[13]. "2022 Website Threat Research Report." SUCURI. https://sucuri.net/reports/2022-hacked-website-report/ (accessed Apr. 10, 2023).
[14]. "2021 Website Threat Research Report." SUCURI. https://sucuri.net/reports/2021-hacked-website-report/ (accessed Apr. 25, 2023).
[15]. "Ranking of the most popular database management systems worldwide, as of February 2023." Statista. https://www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems/ (accessed Apr. 20, 2023).
[16]. C. Ciuchi, G. Petrică, a.o. *Cybersecurity Guide*. (2021). Accessed Apr. 10, 2023. [Online]. Available: https://dnsc.ro/vezi/document/ghid-securitate-cibernetica-2021.
[17]. "Top 10 Web Application Security Risks." OWASP. https://owasp.org/www-project-top-ten/ (accessed Apr. 1, 2023).

The Implications and Effects of Data Leaks

Paul-Andrei PREDESCU, Dragoș BĂLAN
Faculty of Law, "Alexandru Ioan Cuza" Police Academy, Bucharest, Romania
predescu.paul48@yahoo.ro, balan.dragos99@gmail.com

Abstract

In the following article we will present how data theft can have serious effects on the personal life of citizens and users of certain applications, and in general on public institutions and countries. In the following we will find out how these data can end up in the hands of hackers, for what purpose they are used and what are the legal implications. In the end we will analyze how the authorities try to limit this phenomenon and how each of us can take protective measures for this purpose.

Index terms: cybercrime, cybercriminal, data breach, data leak, malware

1. Introduction

In order to be able to understand how data leakage occurs, we must have a well-structured system in which information is formed and stored. Thus, cyberspace is the virtual environment, generated by the informational content processed, stored or transmitted, as well as by the processes and operations carried out by the users of the virtual environment, the human resource produces data that passes through different applications having the storage center in several places (Datacenters, PCs, Laptops, Mobile, Cloud).

Another concept related to data leakage is that of cybersecurity. It represents the state of normality resulting from the application of a set of proactive and reactive measures that ensure the confidentiality, availability, integrity, non-repudiation, authenticity of information in electronic format, of public or private resources and services, in cyberspace. When there is no timely response against threats to cyber infrastructures or human errors occur, data beaches can occur.

A Personal Data Breach is any breach of security that results in the accidental or unlawful destruction, loss, alteration or unauthorized disclosure of Personal Data or access to Personal Data. This includes violations due to accidental and intentional causes. It also means that a breach involves more than just the loss of personal data. A personal data breach can be broadly defined as a security incident that compromises the confidentiality, integrity, or availability of personal data. In other words, a personal data breach occurs whenever personal data is accidentally lost, destroyed, corrupted, or disclosed. If someone accesses or discloses your data without your permission or where the data is not available and the unavailability would have a material adverse effect on the individual [1].

2. How data leaks can occur

When the attacker creates a threat by exploiting vulnerabilities, it leads to risks. These risks can affect the assets causing exposure and thus the data breach has a high chance of occurring. Meanwhile a data leak is caused when an internal source exposes information. Criminals can use a variety of methods to try and break into a network, for example DDoS, Trojans, Malware, disruption via servers/network. Data leaks occur because of an internal problem. They don't usually happen because

of a cyberattack. This is encouraging news for organizations since they can proactively detect and remediate data leaks before they are discovered by criminals [2].

Let's review some of the most common causes of data leaks.

- **Bad infrastructure:** Misconfigured or unpatched infrastructure can unintentionally expose data. Having the wrong settings or permissions, or an outdated software version may seem innocent, but it can potentially expose data. Organizations should ensure that all infrastructure is carefully configured to protect data.
- **Social engineering scams:** While data breaches are the result of a cyberattack, criminals often use similar methods to create a data leak. Then the criminal will exploit the data leak to launch other cyberattacks. For example, phishing emails may successfully gain access to a person's login credentials, which could result in a bigger data breach.
- **Poor password policies:** People tend to use the same password for multiple accounts because it's easier to remember it. But if a credential stuffing attack happens, it could expose several accounts. Even something as simple as having login credentials written in a notebook could lead to a data leak.
- **Lost devices:** If an employee loses a device with a company's sensitive information, it qualifies as a potential data breach. If a criminal gains access to the device's content, it could lead to identity theft or a data breach.
- **Software vulnerabilities:** Software vulnerabilities can easily turn into a huge cybersecurity issue for organizations. It's possible for criminals to take advantage of outdated software or zero-day exploits and turn it into a variety of security threats.
- **Old data:** As businesses grow and employees come and go, companies can lose track of data. System updates and infrastructure changes can accidentally expose that old data [2].

3. The information and interests of hackers

For us to understand the complexity of the hacking world, we have to begin with the beginning and that is, to understand *what a hacker is, how does hacking work, the type of hackers that are currently navigating the web* and *what are the targets of these so-called "cyberpunks" or hackers* [3].

3.1. So what is a hacker?

A definition for this word would be: "A hacker is an individual who uses computer, networking or other skills to overcome a technical problem. The term also may refer to anyone who uses their abilities to gain unauthorized access to systems or networks in order to commit crimes. A hacker may, for example, steal information to hurt people via identity theft or bring down a system and, often, hold it hostage in order to collect a ransom."

- To continue the discussion, we have to understand the ***process of hacking***, how does it work, and an answer might be that "hackers use technical skills to exploit cybersecurity defenses. Ethical hackers test for cybersecurity vulnerabilities and may take up hacking as a profession -- for example, a penetration tester (pen tester) -- or as a hobby. The end goal is often to gain unauthorized access to computers, networks, computing systems, mobile devices or internet of things systems. Many professional hackers use their skills to determine security holes in enterprise systems and then advise where companies should boost their security defenses to keep threat actors out. Results can also be deleterious: Malicious hackers may steal login credentials, financial information and other types of sensitive information.

- Many hackers aim to exploit either technical or social weaknesses to breach defenses. Technical weaknesses may include vulnerabilities in software or other exploitable weak spots. To exploit social weaknesses, hackers may attempt to manipulate social outcomes through false pretenses, such as impersonating a co-worker or other individual to gain financial or login information. Hackers may also use their technical skills to install dangerous malware, steal or destroy data, or disrupt an organization's services.
- Hackers of all types participate in forums to exchange hacking information and tradecraft. There are numerous hacker forums where ethical hackers can discuss or ask questions about hacking. Many of these hacker forums offer technical guides with step-by-step instructions on hacking.
- In contrast, dark web sites often host forums and markets for threat actors or criminal hackers, which serve as a means of offering, trading and seeking out unlawful hacking services.

Scripts, and even specially tailored software programs, are frequently used by criminals who don't usually have the technical skills to penetrate corporate networks. For the purpose of obtaining information on the functioning of the target system, this software can have access to network data. These scripts can be found on the Internet, for anyone who is typically an entry level hacker. Hackers with limited skills are sometimes called *script kiddies*, referring to their need to use malicious scripts and their inability to create their own code. Advanced malicious hackers might study these scripts and then modify them to develop new methods [3].

3.2. What are the types of hackers navigating the web?

In the past, the security community informally used references to hat color as a way to identify different types of hackers, usually divided into five main types. A few of these terms have been replaced to reflect cultural changes.

- Ethical hackers or authorized hackers -- previously known as white hat hackers -- strive to operate in the public's best interest rather than to create turmoil. Many ethical hackers who work doing pen testing were hired to attempt to break into the company's networks to find and report on security vulnerabilities. The security firms then help their customers mitigate security issues before criminal hackers can exploit them.
- Threat actors or unauthorized hackers -- previously known as black hat hackers -- intentionally gain unauthorized access to networks and systems with malicious intent. This includes stealing data, spreading malware or profiting from ransomware, vandalizing or otherwise damaging systems, often in an attempt to gain notoriety. Threat actors are criminals by definition because they violate laws against accessing systems without authorization, but they may also engage in other illegal activity, including corporate espionage, identity theft and distributed denial-of-service (DDoS) attacks.
- Gray hat hackers fall somewhere between ethical hackers and threat actors. While their motives may be similar to those two groups, gray hats are more likely than ethical hackers to access systems without authorization; at the same time, they are more likely than threat actors to avoid doing unnecessary damage to the systems they hack. Gray hat hackers may offer to repair vulnerabilities they have found via their own unauthorized actions rather than using their expertise to exploit flaws for illicit profit, even though they aren't typically - or primarily - motivated by money.
- Red hat hackers, also called eagle-eyed or vigilante hackers, are similar to ethical hackers. Red hat hackers intend to stop unethical attacks by threat actors. While red hat hackers may have a similar intent to ethical hackers, they differ in methodology, as red

hat hackers may use illegal or extreme courses of action. Often, red hat hackers will deploy cyber-attacks toward the systems of threat actors.

- Blue hat hackers, also known as vengeful hackers, use hacking as a social weapon. Frequently, it is used as a means for revenge against a person, employer or other organization. Hackers who post personal and confidential data online to ruin reputations or attempt to gain unauthorized access to email and social media accounts are classified as blue hats.
- Script kiddies are amateur, inexperienced hackers who attempt to use pre-written scripts in their hacking efforts. Often, these are fledgling hacking enthusiasts who cause little damage.
- Hacktivists are organizations of hackers that use cyber-attacks to affect politically motivated change. The purpose is to bring public attention to something the hacktivist believes might be a violation of ethics or human rights. Hacktivism attacks may attempt to reveal evidence of wrongdoing by publicizing private communications, images or information [3].

3.3. What type of information do hackers look for?

There are various types of information that hackers can steal from your business. Make sure you're protecting these in particular:

- **Personal data**
  This includes Social Security numbers, financial information, birth dates, and other sensitive personal data. To hackers, these are quite valuable; in 2019 alone, there were 13 million recorded identity theft incidents. While passport information sells for the most amount of money, Social Security numbers are the most valuable to hackers, as these can be used for tax fraud, opening credit accounts, and other malicious activities. Your business may not collect Social Security numbers from your clients, but their financial data may be easily stolen.
- **Digital infrastructure**
  Hackers are aware of the high costs of a proper IT infrastructure, so they will resort to stealing another business's IT system to save money. Potential indicators of such an attack include network slowdowns, rapid decrease of storage space, and unknown devices connecting to your network. Over time, this will result in additional costs and lower business productivity.
- **Corporate accounts**
  Hackers can also steal your employees' corporate account data through phishing and malware attacks. They can use the information to solicit personal and financial information from your customers, conduct business email compromise attacks, disrupt your operations, or steal.
- **Intellectual property (IP)**
  Your IP is one of the most important aspects of your business. Without it, you won't be able to offer something unique to your customers and stand out from the competition. This is exactly why hackers might want to steal your IP. If they get their hands on your confidential data, they might sell it in the black market and expose your company's business plans, product ideas, and the like. For instance, a hacking group called the Advanced Persistent Threat 10 attacked the networks of more than 45 technology companies and government agencies in the USA to steal sensitive information regarding new and developing technologies. Two hackers from the group were indicted for conspiracy to commit computer intrusion, wire fraud, and aggravated identity theft [4].

4. The set of international norms

Because cybercrimes have become more and more present in every person's life and they can even affect states, international organizations are trying to regulate this problem.

Thus, in 1997, G8 released a Ministers' Communiqué that includes an action plan and principles to combat cybercrime and protect data and systems from unauthorized impairment. G8 also mandates that all law enforcement personnel must be trained and equipped to address cybercrime, and designates all member countries to have a point of contact on a 24 hours a day/7 days a week basis [5].

In 1990 the UN General Assembly adopted a resolution dealing with computer crime legislation. In 2000 the UN GA adopted a resolution on combating the criminal misuse of information technology. In 2002 the UN GA adopted a second resolution on the criminal misuse of information technology [5].

The International Telecommunication Union (ITU), as a specialized agency within the United Nations, plays a leading role in the standardization and development of telecommunications and cybersecurity issues. The ITU was the lead agency of the World Summit on the Information Society (WSIS). In 2003, Geneva Declaration of Principles and the Geneva Plan of Action were released, which highlights the importance of measures in the fight against cybercrime. In 2005, the Tunis Commitment and the Tunis Agenda were adopted for the Information Society [5].

The Council of Europe is an international organisation focusing on the development of human rights and democracy in its 47 European member states.

In 2001, the Convention on Cybercrime, the first international convention aimed at Internet criminal behaviors, was co-drafted by the Council of Europe with the addition of USA, Canada, and Japan and signed by its 46 member states. But only 25 countries ratified later. It aims at providing the basis of an effective legal framework for fighting cybercrime, through harmonization of cybercriminal offenses qualification, provision for laws empowering law enforcement and enabling international cooperation [5].

General Data Protection Regulation (GDPR) is applicable as of May 25th, 2018, in all member states to harmonize data privacy laws across Europe.[6] GDPR puts the individual as the central element and obliges to protect their data through appropriate measures.

4.1. The GDPR principles:

- Legality, fairness and transparency - data should be processed legally and fairly to the data subject. Explanations should be given to the person in a language they can understand, without legal jargon.
- Purpose limitation - the data will not be used in any other way than that presented to the individual.
- Data minimization - only necessary data will be processed.
- Accuracy - updated data will be kept.
- Integrity and confidentiality - the data will be protected by appropriate measures.
- Responsibility - processes will be documented and compliance with the above principles will be demonstrated [6].

4.2. The rights of the natural person:

- The right to information - the person must be informed, among other things, about what data is processed, why, for what purposes, to whom it is transmitted and what rights he has.

- Right of access - the individual has the right to access their own processed personal information.
- The right to rectification - the person has the right to obtain the rectification of incomplete and inaccurate information concerning him.
- The right to erasure - in some situations, the individual has the right to request the deletion of data that is no longer needed.
- The right to restriction of processing - restriction of processing when there are grounds.
- The right to portability - the right of the person to request data portability from one operator to another.
- The right to object - the right of the person to object to the processing, when there are grounds.
- The right not to be subject to automated decision-making, including profiling - the person has the right to human intervention in the case of important decisions concerning him.
- The right to lodge a complaint with the Supervisory Authority - when she is dissatisfied with the way in which her data is processed or when her rights have not been respected.
- The right to go to court - to obtain material and/or moral damages if damage has resulted [6].

5. Legal effects of data theft

Just like any crime, cybercrimes produce certain legal effects and involve the responsibility of the people. On the one hand, we have the criminal liability of the person or persons who stole or tried to steal the data, and on the other hand, we have the responsibility on the companies towards the users because they hat to do all the diligence to protect their information.

The legal ramifications of a data leak can be government fines, penalties, and in extreme circumstance, jail time, are some of the consequences of not protecting personally identifiable information adequately.

One ramification many don't consider is the cost of litigation associated with a breach. Many of the associated lawsuits can end up as class-action lawsuits, potentially multiplying the total cost of the breach exponentially [7].

Settlements can be harsh - depending on the judge or jury. For large breaches, settlements over $100 million are not out of the question, especially when dealing with healthcare information. Another cost of a breach includes having to pay the plaintiff's legal bills, which can be extremely high [7].

A cyber-attack on your business that exposes personal or confidential data could have several nasty consequences for your business, including:

- financial loss from stolen funds or a loss of income from an inability to operate your business as usual.
- claims being made by customers, for example where you have not complied with your privacy policy.
- claims for breach of contract if you do not meet your contractual obligations to comply with data protection legislation.
- regulatory fines for non-compliance with GDPR or the Data Protection Act 2018.
- reputational damage as consumers lose faith in your ability to securely process their data [8].

6. How we can protect our data

In order for us to have our data protected while we use our devices on the Internet we can use some safety precautions so as for our personal information not to end up in the wrong hands. Some advice that is widely used is for us to:

- **Create strong passwords**: For example, a strong password should contain at least 12 characters and contain a combination of lower and upper case letters, numbers and if possible symbols.
- **Never use the same password on multiple accounts**: Having multiple passwords makes it harder for hackers to gain access to your personal information.
- **Don't log in on personal account on free or public Wi-Fi**: Open networks make it really accessible for people to look into your activity and accounts.
- **Install an antivirus and keep it updated**: New viruses are created all the time and so to have and extra layer of protection is always good to have an antivirus installed and up-to-date.
- **Don't click on pop-ups and virus warnings**: These warnings and now called "scareware" which are fake security alerts that when you click them, they guide you to install a program to remove the virus in your computer, but the link contains viruses.
- **Be wary of phishing email**: These emails are sent to thousands of people, pretending to be from banks, companies, online shops, that try to send you on their website where you are asked to write down your personal information.
- **Store personal and financial information securely**: Never access such information in internet cafes or public computers [9].

7. Conclusion

Cybercrimes as we have seen in recent years have become more and more frequent and pose a real threat to our personal and financial information. Attacks can vary in many different ways from simple emails that try to insert malware in your personal devices if you click on them to full scale attacks on websites owned by enterprises.

This paper wanted to show the problem of the damage that those attacks do is in most cases is quite substantial not once for example did people lose their identity, credit cards information or even social security numbers to data breaches by hackers. This is why we have to be extra careful with our presence on the internet and take extra steps of precaution when navigating the web. The governments took note of the risks that can occur while handling this type of information and so adopted the well-known GDPR that protects our personal information on the Internet in a way that the data should be processed legally and fairly to the data subject. Even explanations should be given to the person in a language they can understand, without legal jargon.

Other steps that we can take to protect ourselves on the internet is to use different passwords for the accounts we have, minimalize the information we share on social media, never click on links or pop-ups that warn us that we have been infected with viruses and use an antivirus and keep it up to date.

References

[1] Information Commissioner's Office 'Personal data breaches'. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/personal-data-breaches/.

[2] Glossary ‘What Is a Data Leak? How They Happen and How To Prevent Them’ [Online] Available: https://abnormalsecurity.com/glossary/data-leak.
[3] Wesley Chai and Linda Rosencrance, “What is a hacker“. May 2021 [Online]. Available: https://www.techtarget.com/searchsecurity/definition/hacker.
[4] Ron Samson Jr. “Data Stealing; What Information is a Priority for Hackers?”. 2022 [Online]: https://www.clearnetwork.com/why-do-hackers-keep-stealing-the-same-consumer-data/.
[5] International cybercrime. 20 June 2022. Wikipedia. Available at: https://en.wikipedia.org/wiki/International_cybercrime.
[6] Intersoft consulting: https://gdpr-info.eu/.
[7] The legal ramifications of a data breach: https://www.ironmountain.com/resources/general-articles/t/the-legal-ramifications-of-a-data- breach.
[8] Clive Mackintosh, “Legal consequences of a cyber-attack”: Date: 9 March 2022. Available at: https://harperjames.co.uk/article/legal-consequences-of-a-cyber-attack/.
[9] ‘30 ways to love yourself online – A beginner‘s guide to Personal Data Privacy’. Available at: https://www.privacy.gov.ph/30-ways/.

Security by Design

Elena-Denisa STROE
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
elena_denisa.stroe@stud.etti.upb.ro

Abstract

The security should be an area that can cover multiple technical disciplines that needs to be focused on customers and to try protecting against different threats. There can be multiple disciplines that can be part of the security and those can be: assurance, anti-tamper and information assurance and cybersecurity. Security must be taken into consideration throughout the entire product lifecycle in order to maximize the protection of a system. The purpose of this article is to highlight design security flaws which should always be considered as part of the design flow for an application or a product. The recommendations can be applied in combination with different methodologies, depending on what the company chooses to use, wheatear it is Agile or Waterfall. Principle of security by design will be tackled within the article.

Index terms: design, OWASP, security, web application

1. Introduction

In software engineering, security by design represents a huge impact on the projects of an organization, being incorporated into the product from the beginning, the purpose remaining to create products functionally secure. Security by design is increasingly becoming the most desired development approach to ensure security and privacy in software systems. Companies or any parties which are involved in developing different projects consider and build security into the system at every layer using a robust architectural design. When the security design is taken into account, there is a full architectural design made for this in order to take the best decisions on well-known security strategies, tactics and patterns, defined as reusable techniques for achieving specific quality level.

Companies often expose themselves to risks while experimenting new or advanced technologies. Software development is touching new heights every day, hackers also develop cutting-edge methods to breach cyberspace defenses. Thus, traditional approaches like Vulnerability Assessment and Penetration Testing are insufficient to address the security of the cyber system. It is essential to use ground-breaking methods like security by design, which provides to developers the knowledge to manage the delivery operations and the development testing at any moment for potential flaws. In order to make a system robust when it comes to safety, the security by design is an approach to software and hardware development that seeks to make systems with no critical vulnerabilities and less prone to attacks, through measures such as continuous testing, authentication and best programming practices.

2. Security by Design (SbD)

Security by design is a methodology to strengthen the cybersecurity of an organization by automating its data security controls and developing a robust IT infrastructure. This approach focuses

on implementing the security protocols from the basic building blocks of the entire IT infrastructure design [1].

Although it's not a new concept, the expansion of public cloud usage has made security by design far simpler to be applied. In practice, security by design is about standardized coding, reusable, automated architectures so that your security and audit standards remain consistent across multiple environments.

When a software project is built, the focus has to be on the absence of vulnerabilities, otherwise the production deadlines might not be met or customers might be affected. With contextual knowledge, it is easier to choose the right components, for example, that can reduce risks and mitigate cyberthreats. Below are some objectives that should be taken into consideration when using SbD:

Table 1. Data & Purpose of the system

Objectives	Data System	Purpose of the System
Information to be considered	• What data are you're trying to use? • Where is that data going to go? • Who will use the data or interact with it?	• How are different components connected? • What are the requirements and implications if something goes wrong

3. Principles of Security by Design

- Minimizing Attack Surface Area: this means that in order to minimize attack surfaces, developers should ask what their system is supposed to do, as well as what it should not. It's also important to anticipate the attack vector or any other vulnerability that might potentially compromise your system.
- Least Privilege: users context is only providing the necessary information for their level of access, by establishing certain permissions, in order to prevent unauthorized access to sensitive data.
- Least Common Mechanism: advises against sharing system mechanisms among users or programs that do not require them, in order to function according to initial specifications.
- Separation of Duties: to prevent conflict of interest, wrongful acts, fraud, abuse or errors.
- Defense in Depth: any developed security system is prone to failure and therefore the best approach is to layer the security measures and deliberately overlapping their coverage in order to keep the system safe even when one security measure has failed. Additionally, a notification system should be in place in order be informed when a mechanism was deceived and the system is at risk.
- Failing Securely: depends on the eventuality that the system will fail, hence the need to design an architecture that allows failing without leaving any exposure [4].
- Open Design: The principle of open design states that the security of a mechanism should not depend on the secrecy of its design or implementation. A system that relies on a novel language or method so unusual that no one can currently understand it can and may still be open, in which case the attacker has immediate and full access to the system.

4. Example of security by design

4.1. Failing Securely

- Wrong email log in: In this case the email introduced is incorrect so an error is thrown. The email can be tried for 5 times.

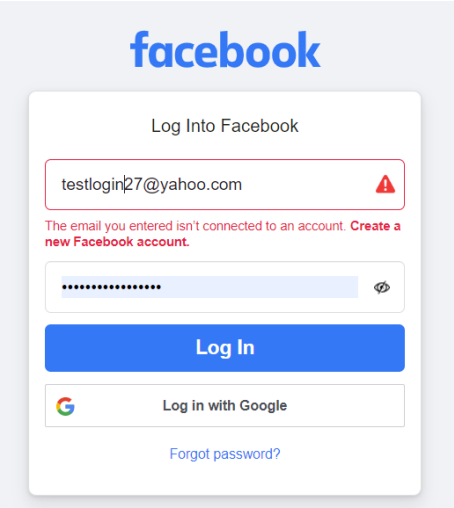


Fig. 1. Wrong email

- Wrong password log in: In this case the password introduced is incorrect so an error is thrown. The password can be tried for 4 times.

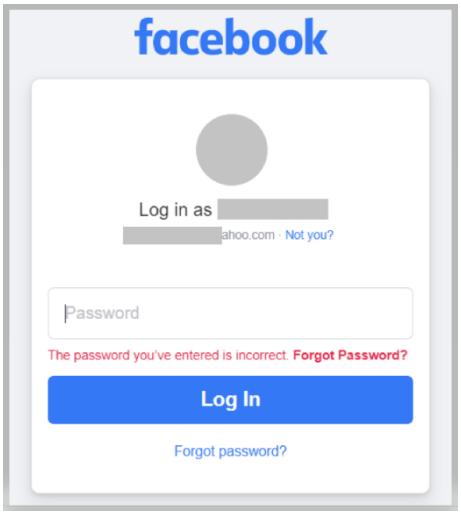


Fig. 2. Wrong password

In this case it is considered the log in page for an application. Sometimes it happens to introduce the wrong data: the email or the password. In case a wrong email will be introduced, meaning that the email doesn't exist at all, then the applications will return an error to highlight that the email is not in the database. Attackers might attempt to guess the email 5 times, so at the 6th attempt the page will be blocked and in that session used, cannot connect to any other email address.

In case attackers fill in an existing email signed up for Facebook, then the next step is to introduce the password. It can be guessed 4 times and at 5th try an error message will be displayed for wrong password. This kind of behavior is displayed in Figure 2.

Regardless of the reason of failure, sensitive user information and system errors should not be exposed. This principle states that only limited information should be shown when errors are encountered by the system. As underlined in Figure 2, when the password is not matching previously inserted user, a dedicated message is being prompted. The problem is that attackers can gather information related to existing accounts by user enumeration due to the fact that if the provided user name does not have an associated account, their messages are different. Attackers might attempt guessing the password 4 times.

4.2. Minimize Third-Party Access

For web applications, making use of the services of third-parties can be convenient for additional functions or data. However, these external parties have different security measures that may or may not be more secure. When an organization agrees to collaborate with other party, then it should take into account that those parties might have different cyber security measures, so it can lead to vulnerability to cybercriminals who might gain access.

4.3. Keep security simple

Contrary to popular belief, keeping the application's security simple is a better option than having complex designs. When organisations use complex systems, then those are very hard to maintain and correct in case of an undesired event. Troubleshooting can be time-consuming which puts the application at further risk.

5. OWASP recommendations

OWASP (Open Web Application Security Project) is an online community that produces free tools, documentation, articles, and technologies to help people secure their websites, web applications, and network resources. That was created for helping developers building highly secure web applications [6].

OWASP suggests that programmers create security controls that are appropriate for managed data value. For example, an application processing financial information must have binding restrictions, compared to a blog or a web forum.

6. Conclusions

As cybersecurity becomes an increasing area of concern for critical infrastructure providers, governments, and private enterprises, it requires greater attention from both management and development team members. Successful implementation will require action from multiple groups and at multiple levels. Security features should be designed into a system so that both human and software vulnerabilities are minimized. In addition, each component of a system should also be secured separately so that if a breach does occur, any damage is going to be limited, and it won't impact and spread through the entire environment.

The principles that guide the security by design approach could differ from one organization to another. But the OWASP listed some principles that programmers should take into account, the one presented in the article. With these in mind, they can design secure products. Below are four security by design principles.

References

[1] https://blog.unguess.io/what-is-security-by-design-the-best-approach-to-cybersecurity
[2] https://www.spiceworks.com/it-security/cyber-risk-management/articles/what-is-security-by-design/
[3] https://www.logicworks.com/blog/2017/01/what-is-security-by-design/
[4] https://www.techslang.com/definition/what-is-security-by-design/
[5] https://learn.microsoft.com/en-us/azure/architecture/guide/security/security-start-here
[6] https://patchstack.com/articles/security-design-principles-owasp/

Enhancing the Security of Cryptographic Systems by Pseudo-Random Number Generation Algorithms

Evelyn ENESCU
Faculty of Electronics, Telecommunications and Information Technology,
University Politehnica of Bucharest, Romania
evelyn.enescu@stud.etti.upb.ro

Abstract

Pseudo-random numbers play an indispensable role in the design of encryption systems, such as public and private key flow. The efficiency of crypto systems is directly proportional to the quality of the secret key generated using a random number generation algorithm. In this paper, the efficiency and applicability of a modified Linear Congruential Generator (LCG) type algorithm will be presented to increase the rate of occurrence of numbers and tend as much as possible to a truly random number. Moreover, it will be integrated into a graphical interface, which can later be integrated into the security of a larger application or even a website.

Index terms: cybersecurity, encryption systems, hazard, Linear Congruential Generator, token.

1. Concepts within cybersecurity

Cybersecurity is the discipline of preventing theft, damage, and unauthorized access to computer systems, networks, and data. Along with cybersecurity, "safety in operation" refers to making sure that computer systems and networks are run securely and safely, and that any potential risks or hazards are found and eliminated [1].

Reliability is another important concept in cybersecurity and safety in operation. It describes a system's capacity to consistently and flawlessly carry out its intended purpose over time. A trustworthy system is one that one can rely on to perform as intended, without unanticipated downtime or other problems.

Other concepts that are relevant to cybersecurity and safety in operation include resilience which refers to the ability of the system to recover fast after a security breach, but also the technique of encryption which protects data and communications from unauthorized access or interception, and other more. All these concepts are critical to ensuring the security, safety, and reliability of computer systems and networks, and to protecting against cyber threats and other risks.

To prevent risks there are some assessments that can be done, such as Fault Tree Analysis (FTA), Failure Mode Effect Analysis, Attack Tree Analysis (ATA) or Probability Risk Assessment (PRA) [2].

Depending on what the functionality is each method has its purpose. PRA is a method of assessing the probability and consequences of potential risks, assessing the likelihood of each hazard and developing strategies to manage them. Attack Tree Analysis (ATA) is a method of analyzing potential security threats by identifying the various attack paths that an attacker might use to gain unauthorized access to a system. FTA is a method of analyzing potential failures in a system identifying the potential causes of a failure, analyzing the likelihood and severity of each potential

cause, and developing strategies to prevent them. FMEA involves identifying the potential failure modes, analyzing the effects of each potential failure mode [2].

In this paper is used an LCG - Linear Congruential Generator algorithm to generate a 6-digit token with the goal of integrating it into the security of a larger software application.

2. Random number generators

A pseudo-random number generator (PRNG) is an algorithm that produces a series of numbers that mimic the properties of random numbers. These numbers are not completely random and are determined by an initial value called the seed.

Random number generators (PRNGs) are widely used in various applications such as simulations, electronic games, and cryptography. For cryptographic purposes, PRNGs must generate random outputs that cannot be derived from previous outputs.

Random number generators (PRNGs) are important to ensure that the result is reliable and accurate. Despite this, John von Neumann warned of the dangers of misinterpreting PRNGs as true random generators, saying "Anyone who relies on mathematical methods to generate random numbers commits a sin" [3]. A thorough analysis of the mathematical properties of the PRNG must be done to ensure that the output is sufficiently close to random for its intended use.

It is important to remember that even the best PRNGs cannot perfectly mimic "true" randomness. As such, for applications that require an extremely high degree of randomness, such as cryptography or scientific studies, hardware random number generators should be used.

A linear congruential generator (LCG) is an algorithm that produces a series of pseudo-random numbers based on basic arithmetic operations. The result is a sequence of pseudo-random numbers calculated with a recursive linear equation [4].

$$X_{n+1} = aX_n + c \bmod m \quad (1)$$

Hence the following conditions must always be respected:

$m > 0$ (the modulo number)

$0 < a < m$ (the multiplicator)

$0 \leq c < m$ (the increment)

$0 \leq X_0 < m$ (the seed)

LCG is defined by the following parameters: an initial base value (also known as a start value), a multiplier, an increment, and a modulus. The seed is the starting point of the sequence, the multiplier and the increment are two values that affect the output, and the modulus is used to terminate the output if it exceeds a certain maximum value. Once the initial values have been set, the LCG algorithm works by taking the current number in the sequence and multiplying it by the multiplier; then adding the increment to it; and finally taking the remainder of the result when divided by the modulus. This new number then becomes the next in the sequence. The same process is repeated until all the desired numbers in the sequence have been generated.

If the increment will take the value 0 then it will be called Multiplicative Congruential Generator (MCG) or Lehmer RNG [5].

The power of random generation depends on the selection of the parameters, so if the parameters have been chosen with very small values it is relatively easy to deduce the resulting sequence. For example, if a=1 and c=2 are chosen, the result will be a two-step numerator with a fairly large length, but of course it will not be random.

There are three options for choosing the parameters [6]:

1. c=0 and m a prime number

In this case, if the multiplier "a" is chosen to be an integer primitive modulo "m", then this is the original construction of the Lehmer random number generator. Its period is equal to "m-1" and the initial state must be an integer between 1 and "m-1".

2. c=0 and m a power of his

The most used values for the multiplier in this case are either 232 or 264, because the result is calculated very efficiently by simply truncation of the binary representation, thus it is no longer necessary to calculate the most significant bits. However, the approach comes with its drawbacks, as the least significant bits have a smaller period than the higher bits, so the resulting values will be able to sweep within a very small range of values. For example, if a=5 and m=8, the first bit will never change and the second bit will alternate between two states.

3. c≠0

If the values for all 3 parameters are chosen correctly, one can repeat the value of the sequence with a period m. The realization of this event will happen if and only if:

1. c and m are prime to each other
2. a-1 is divisible by all prime factors of m
3. a-1 is divisible by 4 if m is also divisible by 4

These 3 requirements are known as the Hull-Dobell theorem and work great when m is a prime number to a high power like 282, but if m were a number that is not squared or at least once, it would allow a single value of 1 and would lead to a very poor result. Although the Hull–Dobell theorem gives a maximum period, it is not sufficient to guarantee a good generator. For example, it is desirable that a − 1 is not divisible by more prime factors of m than necessary. Thus, if m is a power of 2, then a − 1 should be divisible by 4 but not divisible by 8, $a = 5 \; mod \; 8$.

Example values for LCG parameters and results obtained after 10 repetition cycles:

Table 1. LCG example

a	c	m	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	3	10	0	3	6	9	2	5	8	1	4	7	0
2	1	10	0	1	3	7	5	1	3	7	5	1	3
22	1	72	0	1	23	3	67	35	51	43	11	27	19
11	37	100	1	48	65	52	9	36	33	0	37	44	21

These sequences are not random, but (for the right choice of parameters) exhibit many properties of random sequences. For small parameter values, the non-randomness is particularly obvious: whenever we generate a value we've seen before, we enter a cycle where the same subsequence is continuously generated. In the first two examples above, the cycles are 0-3-6-9-2-5-8-1-4-7-0 and 1-3-7-5-1, of lengths 10 and 4 respectively. Since there are only M different possibilities, one must always enter such a cycle, but one wants to avoid short cycles.

Congruential linear generators are a popular and efficient type of pseudo-random number generator, but they have problems such as weak randomness and potential periodicity.

3. Hash map

Data requires a number of ways to be stored and accessed. One of the most important implementations includes Hash Tables. In Python, these Hash tables are implemented through the built-in data type i.e. dictionary. Hash maps are structured by indexed data. A hash map uses a hash function to calculate an index with a key in an array of slots. The key is unique. The analogy can be made with a cabinet with drawers with labels for the things stored in them.

In computer science, a Hash table or HashMap is a type of data structure that maps keys to its value pairs (implementing abstract array data types). Basically, it uses a function that calculates an

index value which in turn holds the elements to be searched for, inserted, removed, etc. This makes accessing data easy and fast. Generally, hash table stores key-value pairs and the key is generated using a hash function.

An example dictionary might be a mapping of employee names and their employee IDs, or student names and their IDs.

A hash map typically includes the following functions [6]:

1. set_val(key,value): Inserts a key-value pair into the hash table. If the value already exists in the hash table, the value will be updated.
2. get_val(key): Returns the value to which the specified key is mapped, or "No record found" if this map contains no mapping for the key.
3. delete_val(key): Removes the mapping for the specified key if the HashMap contains the mapping for the key.

The hash() method returns the hash value of an object if it has one. Hash values are just integers that are used to quickly compare dictionary keys during a quick lookup.

The syntax of the hash() method is: hash(object), where “object” is the object whose hash value is to be returned (integer, string, float) [7].

4. Software implementation

To create this program, the Python IDE development environment was used. The program is one of the most popular now and was used for the purpose of generating a token and making a user-friendly GUI.

Within the project, a series of steps were followed which are represented in the block diagram in Figure 1. The first step is to enter from the keyboard the email address on which you want to generate a token, this option was chosen because a user can have many more email addresses. If the address is not valid, the program will return an error, otherwise it will continue. With the correct email address, a hashing will be done on the email address from which an integer value is returned. The modulus function is applied to this and will be used as the seed for the LCG algorithm. The output of the algorithm is a vector, so it has been converted to a string to be displayed as an integer suitable for a valid token.

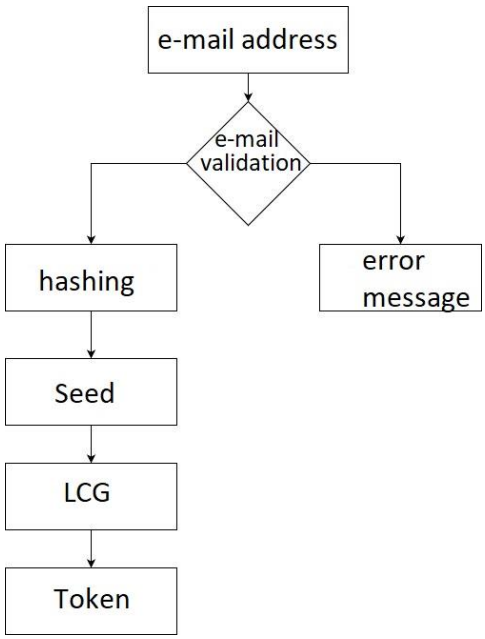


Fig. 1. Block diagram

To be able to determine if the email address is a valid one, a regexp was used. This is a regular or regular expression made up of a string of characters that specifies a pattern, a certain pattern. Which then, is to be used by another algorithm.

The following regexp was used for the email address:

Regex'([A-Za-z0-9]+[.-_])[A-Za-z0-9]+@[A-Za-z0-9-]+(\.[A-Z|a-z]{2,})+'*

Most of the time this is used inside an if structure because it returns the value 0 or 1 depending on whether the pattern was followed. Through the hashing process, a fairly large integer is generated at the output. For example, if an e-mail address is entered, it will return:

```
e-mail address: evelyn.enescu@stud.etti.upb.ro
e-mail address hashed 3650674425931386077
```

Fig. 2. The seed of an e-mail address

This resulting value will be used as the seed for the Linear Congruential Generator algorithm.

The seed value was taken and entered into the LCG algorithm where a was chosen a=1140671485 (multiplicator), c=12820116 (increment) and m=224 (module). With the help of the basic formula of the algorithm, this type of generator was implemented.

By means of a loop repeating 6 times, basically 6 numbers of variable length will be generated, so the last element of each number that resulted from the algorithm will be chosen. The resulting numbers will be concatenated to produce the token.

After testing with the email address *evelyn.enescu@gmail.com*, the following values for the LCG numbers and the token value were obtained:

```
3201383
10224022
7996681
7440048
417211
4572314
Token value 321814
```

Fig. 3. Token generation

Because the creation of an application in which you work with the command line is not very user-friendly, in this project it was also chosen to create a graphical interface (GUI) - Figure 4. This is designed to make interacting with our software as easy and intuitive as possible. It has a modern, clean design that is easy on the eyes and easy to navigate. The layout is organized in a logical manner, with all the main functionalities easily accessible from the main menu. Overall, this GUI is designed to increase productivity and make the user experience as smooth as possible. Moreover, since it is designed for login and Security use, a simple and accessible design has been implemented.

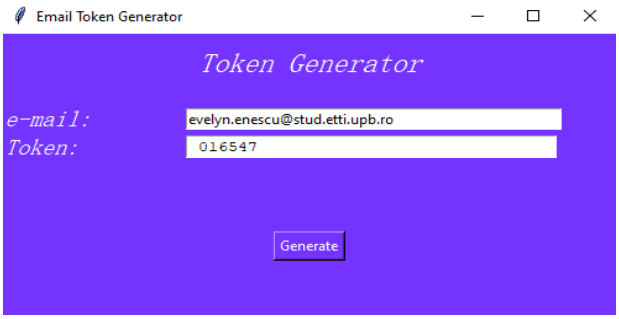


Fig. 4. Application interface

The GUI was implemented using the Tkinter package in Python, based on the TCL language. It is intuitive and easy to understand, making it an ideal choice for both novice and experienced users.

The interface is organized into several sections, with the labels, input field, button, and text field located in the center of the screen. Labels are prominently displayed to the left of the corresponding fields and are written in a clear, easy-to-read font. The input field is designed to accept user input such as text, numbers, or selections from a drop-down menu. The button is labeled with a clear and concise action such as "generate" and is positioned at the bottom of the page for easy access.

The text field sits above the button and is used to display the result of the action initiated by the button. The text field is also capable of displaying multiple lines of text, allowing the user to view detailed information or error messages. The GUI also includes a scroll bar to navigate through the text field if the text is too long to display on one screen. In addition, the GUI includes keyboard shortcuts for quick access to various functions, such as pressing the "Enter" key after entering data in the input field that will trigger the button's action.

5. Results

The results section of this article presents the findings of the study. Through a thorough analysis of the collected data, a number of significant conclusions could be drawn. Results are presented in a clear and organized manner, with graphs and figures used to enhance understanding of the data.

This study evaluated the performance of several different models on a common input data set. At the same email address, evelyn.enescu@stud.etti.upb.ro, 50 tokens were generated in order to determine if any token is repeated and the frequency of occurrence of random numbers at each position. The statistics can be found in the graphs below.

In order to make data testing more efficient, a .py file (Figure 5) exports the generated data to a .csv file (Figure 6). Quantitative and qualitative analysis will be performed on this data set.

```
com = "python main.py"
list = []
for i in range(50):
    out = os.popen(com).readlines()[0]
    out = out[:-1]
    list.append(out)
print(list)

f =  open("data.csv", "w")
for l in list:
    line = ""
    for c in l:
        line = line + "," + c
    f.write(line)
    f.write("\n")
f.close()
```

Fig. 5. Testing code

In this picture there is the code that creates the necessary files, calls for 50 times the main procedure that does the algorithms. Its purpose is to generate a .csv file to which the data will be transferred. The .csv has 50 rows representing all the output of the procedure, but the numbers are in different cells so as to see if a token was repeated and to have a graphic classification of the frequency of the numbers on each position.

Therefore, in Figure 7, the actual value of the number on the first position is entered on the X axis, and the frequency with which it appears on the first position, during 50 iterations of the same e-mail address, is marked on the Y axis.

Position 1	Position 2	Position 3	Position 4	Position 5	Position 6
9	2	9	6	3	8
6	3	8	9	0	7
3	2	3	6	1	6
0	1	0	3	4	7
6	1	0	7	0	9
7	8	5	0	1	0
9	2	1	4	5	8
7	8	5	0	5	2
1	4	1	4	7	2
9	2	5	6	3	4
4	3	2	1	2	3
7	8	5	4	3	0
6	1	4	3	2	1
5	4	9	8	1	4
1	2	3	4	9	8
7	8	1	6	9	0
7	2	7	2	7	2

Fig. 6. Exported data in .csv file

For the following figures, from Figure 8 to Figure 12, the same values corresponding to position indices 2 to 6 are illustrated. For the 1st position, the most frequent numbers are 9 and 6 followed by 5 and 7. For the 2nd position, the most frequent are 2 and 3, however for the 3rd position, the most used are 0 and 5. In the 4th position, the most used are 0 and 7, the 5th position uses the numbers 5 and 8 very often, and in the last the most frequent are 2, 7 followed by 1 and 0.

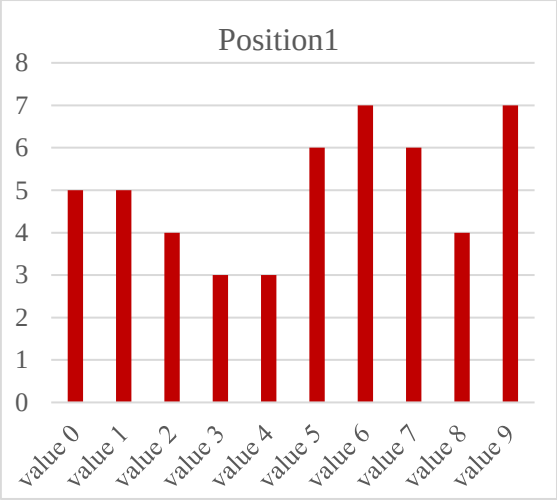


Fig. 7. The frequency on 1st position

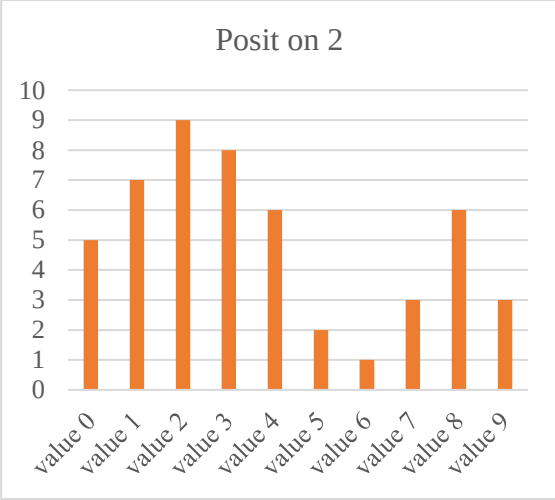


Fig. 8. The frequency on 2nd position

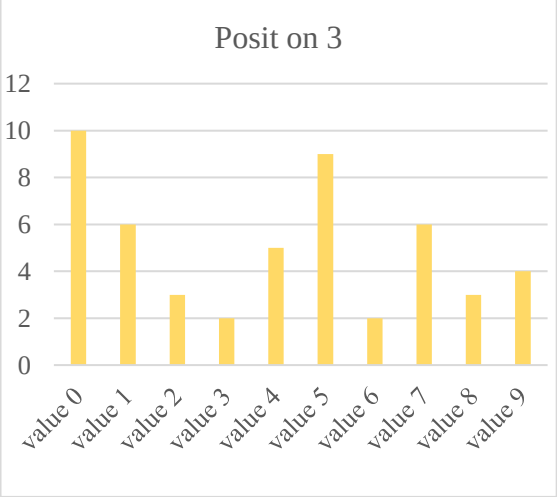


Fig. 9. The frequency on 3rd position

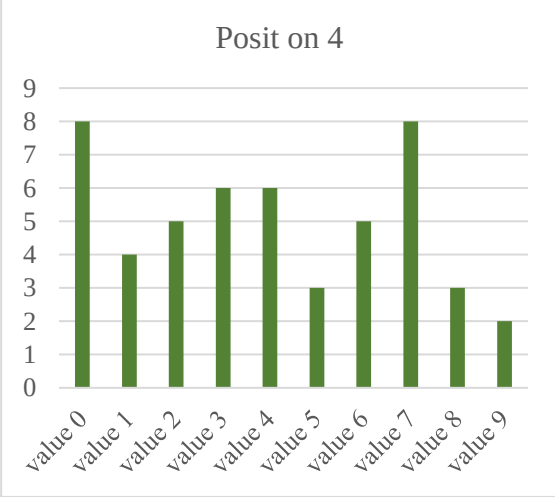


Fig. 10. The frequency on 4th position

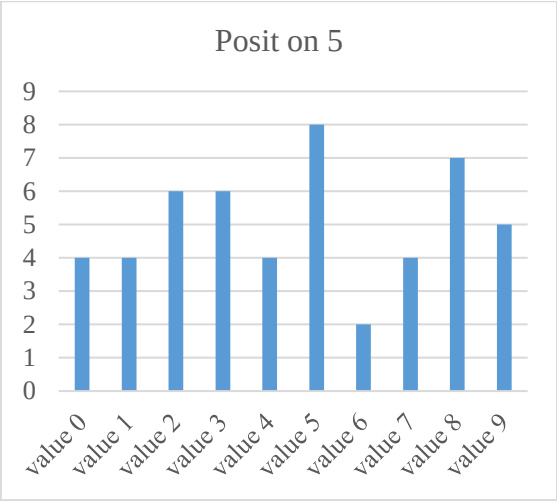


Fig. 11. The frequency on 5th position

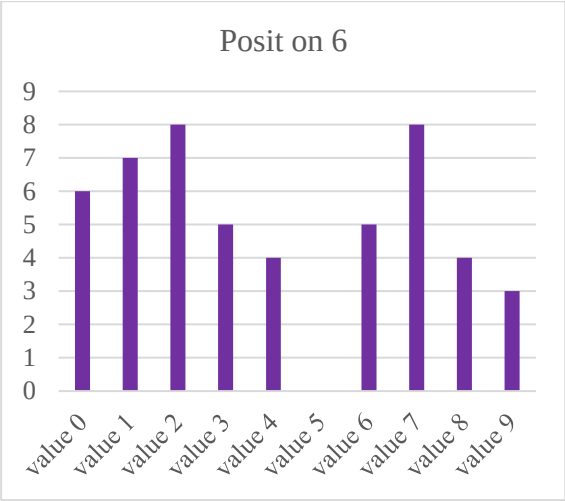


Fig. 12. The frequency on 6th position

6. Conclusions

The results of this evaluation showed that the overall algorithm performed best on the given dataset and could be used to generate repeating tokens and random numbers for the purpose of email authentication. Regardless of the type of pseudorandom number generator used, it is important to understand the strengths and weaknesses of each generator and the selection of the most appropriate one for the application. The advantage of using the LCG algorithm is that it generates pseudo-random numbers in a very accessible and fast way, but it comes with the disadvantage that its mass use can overload the program and generate the same corresponding token for two different email addresses.

In this paper, a modification was made to the classic LCG algorithm in order to generate a token. This algorithm takes an email address as input, checks that it is valid, and then generates 6 strings of numbers according to the algorithm shown. It takes the last digit of each number and creates a new number that will represent the authentication value.

Later, this application can be inserted into any other type of Web or mobile application to secure user authentication. Some examples of applications they could be embedded in are banking, medical results, and other applications that require increased attention to protecting customer or user data.

References

[1]. swarnavo09, “Elements of Cybersecurity,” 2022. Accessed: Mar. 15, 2023. [Online]. Available: https://www.geeksforgeeks.org/elements-of-cybersecurity/

[2]. H. Kavak, J.J. Padilla, D. Vernon-Bido, S.Y. Diallo, R. Gore, S. Shetty, “Simulation for cybersecurity: state of the art and future directions,” *Journal of Cybersecurity*, Vol. 7, Issue 1, 2021. Accessed Apr. 4, 2023. [Online]. Available: https://academic.oup.com/cybersecurity/article/7/1/tyab005/6170701.

[3]. J. V. Neumann, *Various techniques used in connection with random digits*, 1951.

[4]. B. Fathi-Vajargah, R. Asghari, "A Novel Pseudo-Random Number Generator for Cryptographic Applications," *Indian Journal of Science and Technology*, 9(6), 2016.

[5]. H. Tang, “Reverse multiple recursive random number,” *European Journal of Operational Research*, 164, 2005.

[6]. akshisaxena, Hash Map in Python, 2023. Accessed Mar. 22, 2023. [Online]. Available: https://www.geeksforgeeks.org/hash-map-in-python/

[7]. Python hash(). Accessed Mar. 20, 2023. [Online]. Available: https://www.programiz.com/python-programming/methods/built-in/hash

Open-Source Intelligence - Useful Tools in Data Analysis

Adelaida STĂNCIULESCU
Bucharest Court, Bucharest, Romania
adelaida.stanciulescu@gmail.com

Abstract

The paper aims to address how open sources, available in the public space, can provide relevant, high-quality information on which organizations (whether public or private) can strengthen their decision-making process. For example: the development of public policies, the development of security policies, law enforcement norms, the adaptation of tax systems to the digital age, the implementation of targeted marketing campaigns, the widespread access to continuing education, with the aim of creating an adapted workforce in the digital age, the business environment can support technology change through a more intense collaboration with authorities, local communities and society as a whole, etc.

Index terms: Open Source Data (OSD), Open Source Information (OSINF), Open Source Intelligence (OSINT)

1. Introduction

In this article I aimed to present the fundamentals of Open Source Intelligence (OSINT), how it is used, as well as the tools and techniques that can be used to collect and analyze information from open source (Open Source Data). In the era of globalization and digitization, information has become the resource without which progress, at this moment, seems impossible. In this context, the analysis of information from open source (Open Source Data), available on the web, has become a requirement, and even a necessity.

Starting from the premise that information is the first and most important element of the decision, we understand its importance and applicability in all fields, from political to military, from social to economic and even cultural [1]. Over the past two decades, the entire world has witnessed profound transformations as a result of globalization and technological change.

The galloping evolution of information and communication technology, recorded in the last two decades, has opened up new opportunities aimed at significantly improving the techniques and methods that can be used in data analysis.

Using Open Source Data (OSD) competitive intelligence analysis, organizational leaders can gain a valuable perspective and engage in debates to find beneficial solutions in a world of more and more virtual interactions. Before presenting the sources and use of Open Source Intelligence (OSINT) it is important to understand the terms used in this field.

2. Terms used in the field of Open Source Intelligence (OSINT)

2.1. Open Source Data (OSD)

According to US public law [1], Open Source Data (OSD) is publicly available information from open sources. According to the OSINT Guide developed by the Romanian Intelligence Service

[2], data from Open Source Data (OSD) are considered information communicated through radio/TV broadcasts, prints, unprocessed signals, photographs, tapes, satellite images and personal letters.

Although the definition published in the OSINT Guide seems more rigid, in my opinion it is much more clarifying, because it is accompanied by the source of this data, thus, as can be seen, the data from the open source Open Source Data (OSD), can be considered those data which are:

- Published or broadcast in the public space (for example, news media content);
- Available to the public on request (for example, census data);
- Available to the public by subscription or purchase (e.g. industry magazines);
- Could be seen or heard by any casual observer;
- Made available at a meeting open to the public;
- Earned by visiting a place or participating in any event that is open to the public.

From both perspectives on the definition of the notion of Open Source Data (OSD), the public nature of this data undoubtedly follows, this information is "publicly available". The term "open source" refers specifically to information that is available to the general public. If specialized skills, tools or techniques are required to access information, it cannot reasonably be considered open source. Paradoxically, open source information (OSD) is not limited to what we can find using the main search engines.

Web pages and other resources that can be found using Google are certainly massive sources of open source information (OSD), but they are far from the only sources. According to former Google CEO Eric Schmidt, a huge proportion of the information available on the Internet, more than 99%, cannot be found using the main search engines. This so-called "deep web" is a mass of websites, databases, files and more, which (for a variety of reasons, including the presence of login pages or barriers raised by payment mechanisms) they can be indexed by Google, Bing, Yahoo or any other search engine. Despite this, much of the content of the deep web can be considered open-source data (OSD) because it is readily available to the public.

Such legal, open sources (so-called "white sources") include, but are not limited to:

- National business registers;
- The official documentation that companies must present by law, including financial statements;
- Information from relevant state offices and units (Public Procurement Office, Chief Inspectorate for Environmental Protection, Consumer Protection Office, etc.);
- Bankruptcy notices in court and information from debt exchanges;
- Statements of spokespersons for companies and state persons;
- Press and mass media;
- Social networks;
- Social surveys;
- Public life.

2.2. Open-Source Information (OSINF)

Public legislation in the US does not make a strict delimitation of the notions: information produced from open source data (OSD), i.e. Open Source Information (OSINF) [2] and Open Source Intelligence (OSINT), letting them complement or substitute in places, while the OSINT Guide developed by the Romanian Intelligence Service [2], draws clear limits regarding this notion as follows, by Open Source Information (OSINF) - we mean: correlated and processed data to create information of general interest - articles from mass media, books, communiques.

2.3. Open Source Intelligence (OSINT)

As a complex, specialized and distinct process, Open Source Intelligence (OSINT) integrates human experience, with data obtained from open sources, in order to produce information and

informative documents relevant to the decisions of leaders, regardless of the type of organization they belong to.

According to US public law [1], Open Source Intelligence (OSINT) is:

"(1) ...information produced from open source data collected - Open Source Intelligence (OSINT) , exploited and disseminated in a timely manner to an appropriate public, in order to respond to a specific information requirement."

"..."

"(3) Open-source intelligence production is a valuable intelligence discipline that must be integrated into the tasks, collection, processing, exploitation, and dissemination of information to ensure that United States decision makers are fully and completely informed."

According to the OSINT Guide developed by the Romanian Intelligence Service [2], by Open Source Intelligence (OSINT) we mean the results of a complex OSD and OSINF processing process, which involves identification, validation of sources, collection, corroboration and analysis, in order to develop products with relevance in terms of national security, which correspond to specific intelligence requirements.

Open Source Intelligence (OSINT) uses advanced technology to discover and analyze massive amounts of data, obtained by scanning public networks, from publicly available sources such as social media and the deep web - content that is not crawled by engines but search, which is, however, publicly accessible.

OSINT tools can be open source or proprietary: a distinction must be made between open source code and open source content. Even though the tool itself is not open source, as an OSINT tool it provides access to openly available content known as open source intelligence [4].

OSINT is in many ways the mirror image of operational security, which is the security process by which organizations protect public data about themselves that could, if properly analyzed, reveal damaging truths. IT security departments are increasingly tasked with conducting OSINT operations within their own organizations to strengthen operational security.

2.4. Validated Open Source Intelligence (OSINT-V)

Another notion encountered in the field of Open Source Intelligence (OSINT) is: Validated Open Source Intelligence (OSINT-V). Thus, according to the same OSINT Guide, developed by the Romanian Intelligence Service [2] - by Validated Open Source Intelligence (OSINT-V) we mean those data which have a high degree of certainty, either because they are made by a professional analyst or because they come from reliable open sources.

As a partial conclusion, we can state that the four previously presented notions actually constitute the four stages completed in the Open source intelligence (OSINT) process.

So:

1. Open Source Data (OSD), represents the initial stage of collecting raw data from several different sources - without analyzing and processing this data;
2. Open Source Information (OSINF), the second stage, which consists of an initial grouping of the data collected during OSD and an initial general analysis of the information held;
3. Open Source Intelligence (OSINT), i.e. the transfer of processed data to the requester;
4. Validated Open Source Intelligence (OSINT-V), based on checking information already available in other open sources and sometimes comparing it with data obtained through other methods.

In the following we will focus on how Open Source intelligence (OSINT) can be used as best practices for cyber security. There are two common usage scenarios:

- Ethical hacking and penetration testing
- Identification of external threats

Ethical Hacking and Penetration Testing

Security professionals use open source intelligence to identify potential weaknesses in managed networks so they can be fixed before they are exploited by threat actors.

Common weaknesses include, but are not limited to:

- Accidental leaks of sensitive information, for example, through social networks;
- Open ports;
- Unsecured devices connected to the Internet;
- Software without updates, such as websites running old versions of common CMS products.

Identification of external threats

The Internet is a great source of information that can provide information about possible attacks on an organization, identifying new vulnerabilities and how they are being actively exploited, to intercepting threat actors' conversations about an upcoming attack.

Thus, Open Source Intelligence (OSINT) allows security professionals to prioritize their time and resources to address the most significant threats.

In most cases, this type of activity requires an analyst to identify and correlate data from multiple sources to validate a threat before taking action. For example, while a single threatening post on a social network may not be a cause for concern, the same post would be viewed in a different light if it were linked to a group of known threats that he is active in a certain community.

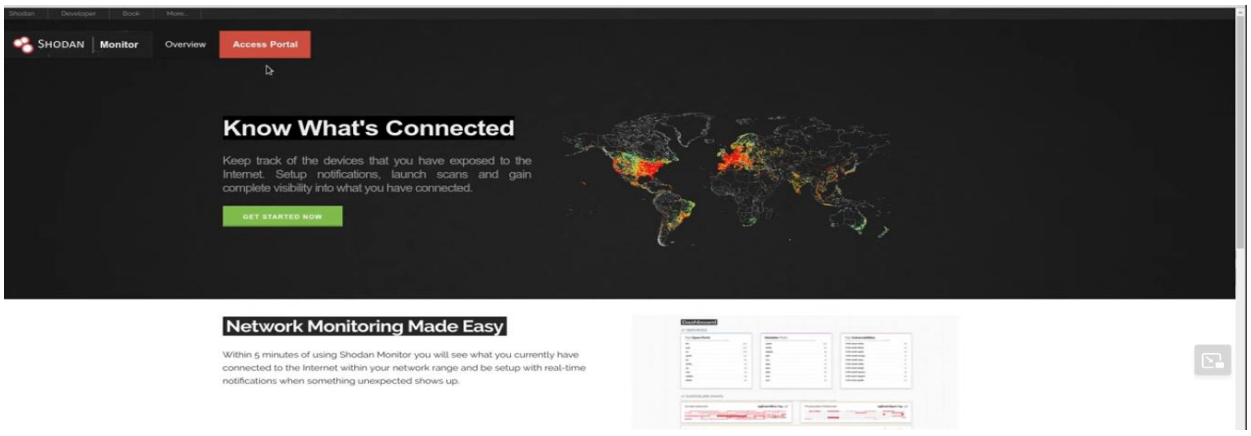
One of the most important things to understand about open-source intelligence is that it can often be used in combination with other subtypes of intelligence. Information from closed sources such as internal telemetry, dark closed communities and external information sharing communities are regularly used to filter and verify open-source information.

Case study on collecting data and turning it into useful information

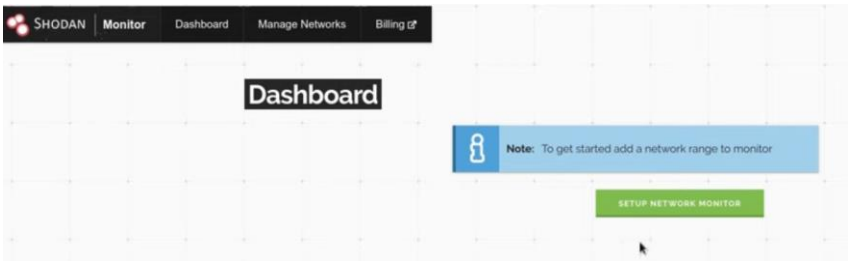
As an analyst, having a large amount of information at your disposal is both a blessing and a curse. On the one hand, you have access to almost anything you could possibly need, but on the other hand, you have to be able to find what you need by actually searching through a torrent of data.

In this example we will use online tools other than traditional search engines. As you know, Google is the most used search engine, but Shodan is a search engine that provides results for security professionals and more, being a goldmine for hackers to see exposed assets.

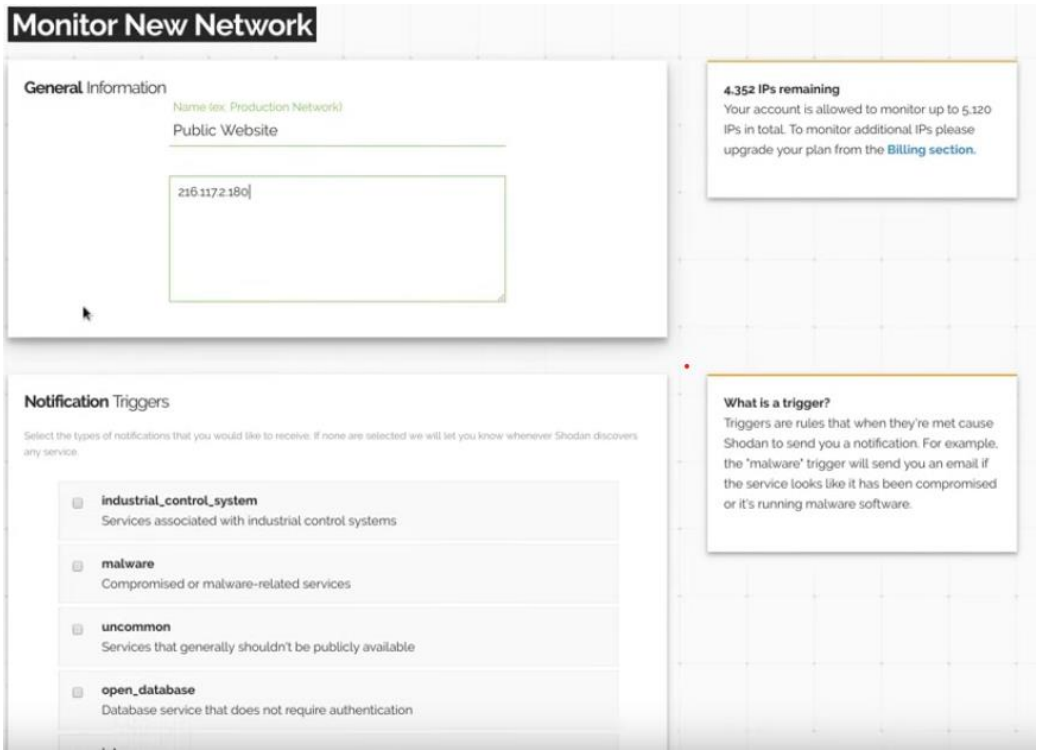
Shodan is a security monitoring solution that makes it possible to search deep web and IoT networks. It makes it possible to discover any type of device connected to a network, including servers, smart electronic devices and web cameras. It mainly includes information related to the assets that are connected to the network. Devices can range from laptops, traffic signals, computers and various other IoT devices. This open-source tool mainly helps the security analyst to identify the target and test it for various vulnerabilities, passwords, services, ports and so on.



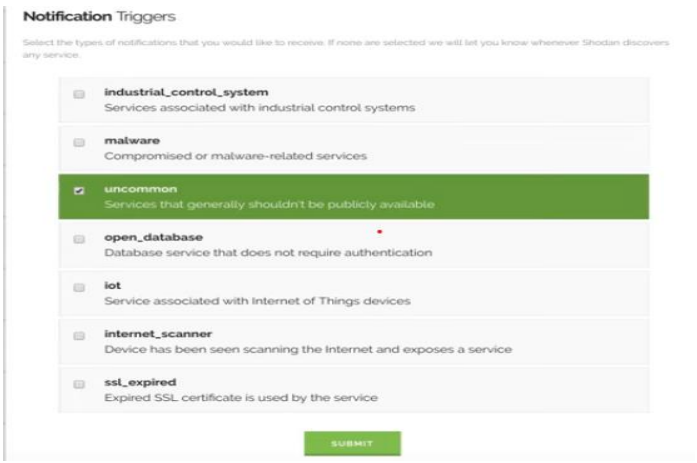
Shodan manages to identify and test "default passwords", devices with VNC viewer, use of open RDP port for testing available assets, etc. This tool is available at https://www.shodan.io and requires login to access the information. After authentication, the network to be monitored must be configured.

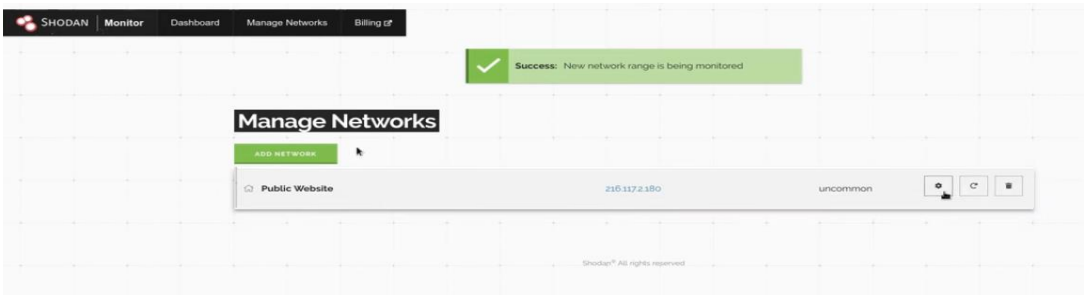


In this example, we configure for monitoring a site located at the IP address 216.117.2.180.

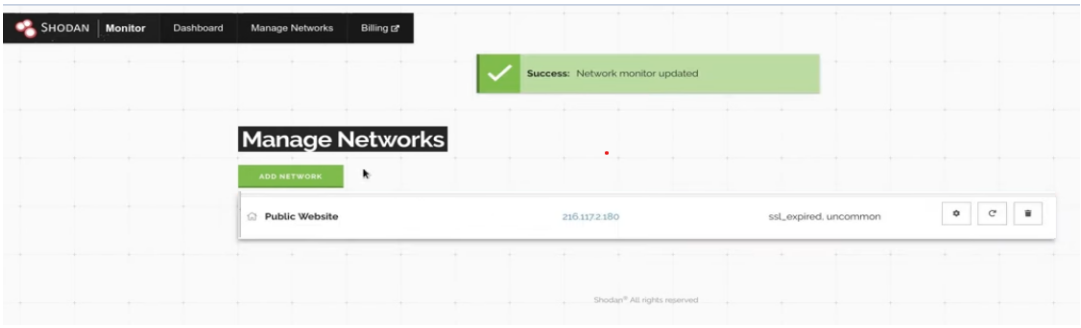


For this site, we configure receiving alerts for unusual threats to services that should not be available to the public.



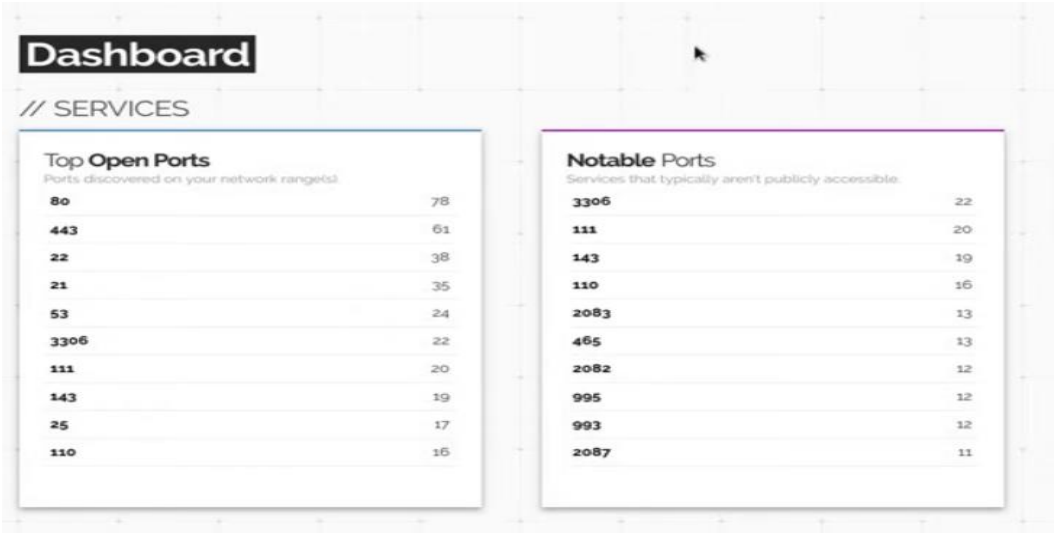


At the same time, I want there to be notifications about the expiration of the site's security certificate (SSL certificate) and therefore I will modify the previous configurations with this new alert.



Once these events are configured, for which the alert is to be made, the possible alerts can be consulted in the Dashboard section, from the same interface.

In this section the information is structured as can be seen in the following image, and at first glance we see information about the open ports on the server.



3. Conclusion

Open-Source Intelligence (OSINT) is the practice of collecting information from published or otherwise publicly available sources. OSINT operations, whether practiced by IT security professionals, malicious hackers, or government-authorized intelligence agents, use advanced techniques to search through the flood of data available on the Internet and find that data in order to achieve objectives.

According to the OSINT Guide developed by the Romanian Intelligence Service [2] it is estimated that OSINT provides between 80% and 95% of the total data used by the intelligence community, worldwide [5].

OSINT provides access to some of the best data available in the world, whether you're conducting a research project, looking to gain competitive intelligence, uncover vulnerabilities, or conduct an analysis of potential threats.

Even if you are simply a person concerned about your privacy and want to find out what personal information has been inadvertently leaked, OSINT can be useful.

Open-source Intelligence (OSINT) can help organizations gather high-quality, grassroots intelligence and make choices based on it.

Open source, in this context, does not refer to the open-source software movement, although many OSINT tools are open source, instead, it describes the public nature of the data analyzed [7].

Despite their great utility, open-source intelligence tools also have a dark side that hackers or people involved in illegal activities can exploit, so great care and caution is required when using these tools, so as not to exceed legal limits.

References

[1]. https://www.govinfo.gov/content/pkg/USCODE-2011-title50/html/USCODE-2011-title50-chap15-subchapI-sec403-5.htm
[2]. https://www.sri.ro/upload/Ghid_OSINT.pdf
[3]. https://opensource.com/
[4]. https://www.imperva.com/learn/application-security/open-source-intelligence-osint/
[5]. https://sri.ro/
[6]. Truyens, Johan, Developing Open Source Capabilities, EDA Bulletin, nr. 9, iulie 2008
[7]. https://www.csoonline.com/article/3445357/what-is-osint-top-open-source-intelligence-tools.html
[8]. Clark, Robert M. 2013. Intelligence Collection. Sage Publications.
[9]. https://data.europa.eu/en/publications/datastories/open-source-intelligence
[10]. https://eda.europa.eu/
[11]. http://www.risk-uk.com
[12]. https://i-intelligence.eu/

Unit Testing and Automate Security Testing

Roxana PRUTEANU
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
roxana.pruteanu@stud.etti.upb.ro

Abstract

In the current context, technology plays a crucial role in our lives, from the moment we wake up until the end of the day we interact directly or indirectly with this new world. Since it appeared, its purpose has been to come to the aid of humanity, to evolve in an efficient and effective way and with all that, it also represented an open door for people who used technology in an obscure way. The number of cases of cyber-attacks has increased exponentially, from data theft to the integrity of critical sectors (health, transport, energy, financial), every possibility was tried to be exploited, leading to serious consequences. Awareness is the first step towards safety, and further it is important how to use technology in order not to be the target of attacks, but also to stay informed and to become better persons in daily activities. This paper presents an analysis of automated testing for software applications, what it is, how automated testing is divided, the benefits brought by it, as well as unit testing details and some examples. Finally, automatic security testing is discussed, the most emerging web application security risks, suggestions about Android security testing tools and some automation frameworks. The focus is on creating an overview, differentiating between the terms used and exemplifying them.

Index terms: Automated Testing, Unit Testing, Development Testing, Dynamic Testing, Automated Security Testing

1. Introduction

To ensure the reliability of the software produced, a crucial term appears in software development: testing. Testing as part of Software Development Life Cycle is the phase where all activities that address the implications of producing quality products, work according to customer requirements and, although it does not add quality to the final product, testing contributes to a high-quality product. The smallest part that builds the software quality testing base is unit testing, and from this point a good quality program is built, saving time and costs. But unit tests can be written/executed manually as any regular program, however, living in the age of speed, unit tests acquired a new, more advantageous form, that of automated tests. In the following topic, automatic testing will be explained to understand its role at present, the advantages it comes with, but also the influence in the other test blocks. Then, unit testing will be analyzed, steps in a process of running unit tests, an example of a piece of the production code and an example of a false unit test.

Last but not least, a rather important topic in the current context, which grabs attention is security automation and testing. In this section, security testing is presented as a vital part of the quality assurance process, the terms are exemplified to create an improved overall image, the top ten web application security risks were presented, suggestion of Android security tools and the automation frameworks that may help in various kinds of security testing scenario. And finally, the paper ends with conclusions in Section 4.

2. Automated testing for software applications

Software testing is one of, if not the most critical phase of the Software Development Life Cycle and in this phase are measured the quality assurance criteria and the Key Performance Indicators already set for releasing software in the production environment. Software testing, as part of the Software Development Life Cycle, has seen exponential growth in both directions of types of testing and volumes of testing. Quality is often sacrificed in order to release software as quickly as possible. Testing has many benefits, including the confidence in the delivered code quality and increased application reliability [1]. All the activities of software testing can be conducted by two means: automated testing and manual testing. Manual testing is considered the fundamental software testing and is done without using automation tools or scripting [2].

One of the most dynamic parts of software testing is automated testing. Automation testing reduces both regression time and the overall costs. The most used automated testing types are Automated Unit Tests, Automated Functional and Non-Functional Tests, Automated Regression Tests, Automated Deployments. The automated testing will provide the following benefits: faster release cycles, extended test coverage, better code quality, improved reliability. Automated testing must be realized using proper tools that provide advanced capabilities for optimization, execution, and reporting [1]. The testing tool executes the test plan in order to assess both functional and non-functional attributes for the software iteration under test. The aim of automated testing is to reduce repetitive tests for a redundant action and to gain time, but is not a replacement for a manual testing.

Organization's success strongly rely on the quality of their products. To develop a product of good quality and without any critical/major defects within the cost and time constraints have become critical and implementing such products is a difficult task. Software testing is performed to evaluate correctness and functionality of software for assuring fulfillment of user requirements and the expected quality. "IEEE defines software testing as the process to evaluate the system or its components manually, or by automated means to determine whether it fulfills the user requirements, or to find the difference among actual result and expected result" [2]. Therefore, the purpose of software testing is to execute a software to identify defects or any missing features that were expected by the user requirements. If it is executed appropriately, software testing results in improved quality and effectiveness of the software system. The reduction of maintenance costs is done by detecting the defects in a software and removing those defects before the release of software.

Test cases describe the complete test scenario in terms of actions to be performed during testing, but when the automatic testing is performed, there is no manual navigation through the different parts of the application, the testing is conducted through some testing tool. Initially, only manual testing was performed, which did not ensured a good quality of software systems, and few defects may be ignored or unidentified through manual testing because of human errors. This situation led to the evolution of automatic testing, being useful in quicker testing process, recently many testers prefer to use automated testing for the variety of software systems. "The basic element behind automated testing is the automated testing tool that is used to conduct the tests" [2].

Due to the popularization of automated testing in software industries, the testing process become more effective, this one helps in easily executing various tests like performance testing and regression testing and many difficult testing activities got easier than before, as the automated testing evolved and improved. Automated software testing conducts the tests for various datasets and the tests can be executed repeatedly without human involvement, it can be performed in various phases (e.g. preparation of test plan or developing the test cases, selecting the testing tool, creation of the test script and finally executing the test by using the automated testing tool and the script). Testing automation results in improved efficiency, the main objective of automating software testing is to reduce the testing effort, costs and time, and to reduce human involvement in the testing process as

much as possible. "Automated testing supports the reusability of test scripts, using the testing tool, for different upgrades of the system under test. Automated testing has the following benefits:

- Tests are repeatable and reusable,
- Simplified regression testing,
- Reduces time and costs,
- Performance testing is possible due to simultaneous testing.

Automated testing has the following drawbacks:

- It is more expensive than manual testing,
- All areas cannot be automated,
- Manual testing cannot be fully discarded" [2].

2.1. Unit Testing

Viewed in the light of developer testing activities, testing is an activity performed to ensure correctness and quality of software (we can verify its functionality, measure progress while developing it) and this process start with unit testing. Unit testing is an integral part of development [3], and this type of testing is based on testing the smallest piece of code that can be logically isolated in a system, it is the easiest, fastest, and most consistent way to verify developers' assumptions about the code they produce [6]. According ISTQB (International Software Testing Qualification Board), unit tests are the first level of dynamic testing in the software testing process [3], is usually performed by developers, testers and QA engineers (they are written by developers and often involves a collaboration between these roles) and it helps uncover early bugs and flaws in application code [9]. Productivity is lower for software supported by tests, but it's kept constant over time, when they plot productivity versus time for software with and without tests. Initially, for software without unit tests, productivity is higher, but it plummets after a while and becomes negative [6]. According to the test pyramid, the higher the level, the higher the involved costs. This effect is multiplied because the involved cost increases involve:

- Setting up the tests,
- Maintenance of existing tests,
- The cost of test execution,
- Test result waiting time,
- Analysis and resolution time when a bug is detected [3].

The process of running unit tests consists of four steps:

1. Creating test cases: this stage requires writing multiple test cases for web applications,
2. Review and re-write: written test cases are reviewed and re-write if there are any mistakes,
3. Baseline: it is checked if each line of code is in a manner or another,
4. Execution: it is performed test execution using an online Selenium Grid [9].

To create a clearer picture, we have the following example which shows a unit test for a piece of the production code [10]:

```
import pytest
#Production Code
def str_len( theStr ):
   return len(theStr)
#A Unit Test
def test_string_length():
   testStr = "1"                    // Setup
   result = str_len(testStr)        // Action
   assert result == 1               // Assert
```

"The production code is the function that returns the length of past in string and the unit test is a single positive test case that verifies the length of one has returned for a string with one character in it. The test string length call is the unit test for the string length production code." This structure represents a common structure that all unit tests should follow being composed of three steps: **a setup step** where it creates a test string, **an action step** where calls a production code to perform the action that is under test, and the last one **an assertion step**, where the test validates results of the action [10].

After the unit tests are performed, the "higher-level tests" follow (integration tests, system tests, system integration tests, solution tests, acceptance tests) [6], but there is a family of tests that shares some characteristics with unit tests and some with higher-level tests, which tends to cause confusion. A common trait of such tests is that they're not unit tests although they seem, due to the fast execution time - in the range of 1 to 2 seconds - these tests they look deceivingly simply and are fast, but they're often integration tests, or even system tests. How do they make it into the unit test suite? The most plausible reasons are: they are not classified (if no attention is paid to how and when different tests are done, these tests end up in unexpected places), the test suite is small (if the test suite consists of a small number of unit tests it doesn't matter if some of them take a few extra seconds to run), hurry (for example, in the beginning of a project it is wanted to prove that the product has the potential to be commercially viable and in this stage, fast medium tests may live in the unit test suite). The easiest way to recognize these tests is to see a concrete example that have been appearing rather consistently in developers' projects over the years, named tests using In-memory databases.

Various SQL-compliant in-memory database implementations exist that are very fast, which not just only do they perform reads/writes much faster than databases that make use of disk storage, but they're also easier to set up, because they require virtually no installation and provide programmatic APIs for configuration. Databases are an important component for tests that require a data source and that validate vendor-agnostic functionalities. Let's take the case of the authentication because it is complicated and this is quite a good case, AuthenticationManager class uses the database correctly and that password hashing seems to work as expected, but nevertheless it is not a unit test. It loads classes, starts a database, and establishes a connection to it, at the time of writing, it ran in less than a second [13].

```
private Connection conn
def setupSpec() {
    Class.forName("org.hsqldb.jdbc.JDBCDriver")
    conn = DriverManager.getConnection("jdbc:hsqldb:mem:db", "SA", "")
    Sql.newInstance(conn).execute(
          "CREATE TABLE users(id BIGINT IDENTITY, " +
                "name VARCHAR(45), "+
                "password_hash VARCHAR(45))")
}
def "Authenticate user"() {
    given:
    Sql.newInstance(conn).execute("INSERT INTO users " +
        "(id, name, password_hash) VALUES (NULL, 'regular_user', '%ReG^7@R')")
    expect:
    new AuthenticationManager(conn).authenticate("regular_user", "secret")
}
```

Any application can host a multitude of opportunities to create tests similar to unit tests by running just as fast, but which in essence are not, they are part of the category of false unit tests. The least common denominator of these tests is that they all start a server somehow, but server took relatively little time to start, so waiting has been acceptable. One in-memory database test takes one second, ten of the same tests 2 seconds and combining this with other almost unit tests, the unit test suite will start getting sluggish. If the latency will pass a certain threshold, the tests will no longer be

executed, but leaving that aside, running these tests as unit tests is not a good idea, they make test suite more brittle and sensitive to environment settings [13][14].

Unit tests, even in the absence of test-driven development (TDD), are vital when writing new code (because they contain the highest amount of detail), their presence ensures that the code is testable, and they serve as specifications [6]. In those analyzed previously, the first safety net for catching bugs are unit tests, tests made in the production code, they build and run in the development environment [10].

3. Security automation and testing

Security testing is an integral part of the testing process, which verifies the software's vulnerability to cyber-attacks and tests the impact of malicious or unexpected inputs on its operations. These tests are evidence that systems and information are safe and reliable, and unauthorized entries are not accepted [4]. Returning to what was previously presented, about the dynamic part of software testing and its most used parts, automated functional and non-functional testing [1] represent subdivisions that cover different requirements regarding software testing. Dynamic testing includes functional testing, which focuses on what the software does, if software's functions are working properly, and non-functional testing which focuses on the design and correct configuration of the application (targets a solution's quality attributes, these being usability, reliability, maintainability, security, portability, etc. [6] [14]), security testing being part of non-functional testing [4].

For example, to differentiate functional from non-functional testing:

- A functional unit test: is considered a unit test of a sorting algorithm who verify that the input is indeed sorted,
- A non-functional unit test: the previous sorting algorithm it is subject to a unit test that times it to make sure that it runs within a specified time constraint.

Probably, this example seems not to be part of the present chapter, but this is aimed at differentiating terms according to context, security testing was described as part of non-functional testing, but the example can be extrapolated to security too. For example, a functional security test may be about a user who logging in nonprivileged, and attempting to do something in the system that only users with administrative privileges are allowed to do [14]. This example was created for a better understanding, security testing will be used further as part of non-functional testing.

The purpose of security automation is to reduce repeated manual testing and increase testing coverage in an efficient manner. According to OWASP, they presented top 10 web application security risks. "Using the OWASP Top 10 is perhaps the most effective first step towards changing the software development culture within organizations into one that produces more secure code" [8].

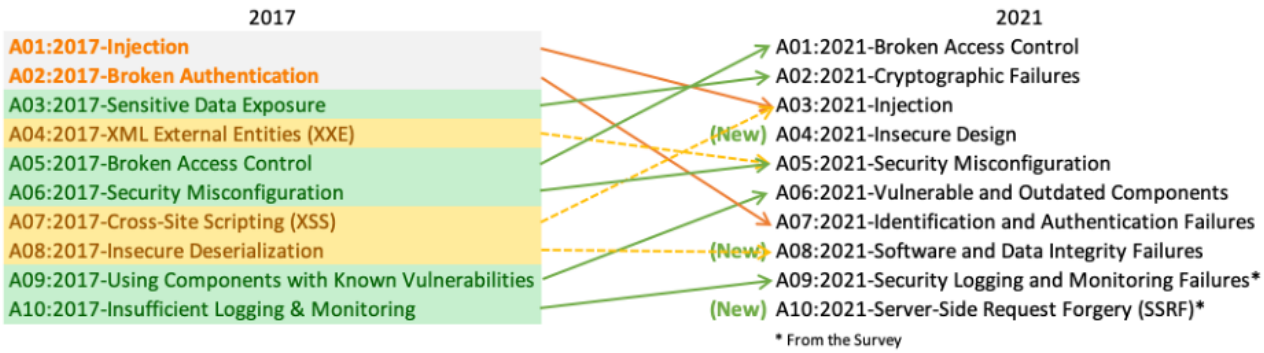


Fig. 1. Top 10 web application security risks [8]

Some categories have changed from the previous installment of the OWASP Top Ten, but it takes time to integrate these tests into tools and processes, and in the figure presented previously is a

high-level summary of the category changes. Top ten is an awareness document, it is the bare minimum and just a starting point for a coding or testing standard [8].

Next is presented a suggestion of Android security testing tools:

Table 1. Suggested Android security testing tools and approach [5]

Scanning approach	Automated tools	Description
Secure code scanning	Fireline	Static Java source code scanning. It's a light-weight secure code scanning tools, but it may require the Java source and the reverse of APK.
Privacy and sensitive information scan	Androwarm	The focus is on privacy and sensitive information scanning of any given APK. Static analysis of the application's Dalvik bytecode, represented as Smali for PII and sensitive information leakage or exfiltration such as telephony identifiers, geolocation information leakage, audio/video flow interception, and so on.
Light-weight all in one APK security scanning	Quick Android Review Kit (QARK)	It's a Python program that can automatically do security scanning of any given APK.
All in one security scanning	Mobile Security Framework (MobSF)	The MobSF is similar to QARK. In addition, MobSF supports Android, Windows, and iOS applications. It not only does the static security analysis, but also does dynamic runtime behavior analysis.

Android application security testing techniques include: the source code scan, privacy information inspection, reverse engineering of an APK, the adoption of automated security testing frameworks (for example QARK or MobSF) [5]. There are some key considerations when is required selecting security automation tools and these selected may depend on the integration of the existing automation testing framework [7]. In general, a security automation framework includes security testing tools, a web service, testing results, an automation framework (for example Robot Framework), automation scripts, and security payloads. "Security testing tools are in charge of testing for specific security vulnerabilities, such as cross-site scripting (XSS) and SQL injection, and also analyze HTTP responses for security issues, security testing tools may provide initial testing reports and testing results can be further integrated by either a testing framework" [12].

Automated UI testing may only cover the scenarios from a user perspective, while the system testing may cover more business logic and user expectations. The following table shows the automation frameworks that may help in various kinds of security testing scenario:

Table 2. Automation frameworks [11]

Types of automation frameworks	Usages
Web UI automation (Selenium or Robot Framework)	*User registration* flow *Authentication/authorization* flow Shopping cart and *Payment* flow *Forget password* flow PII (Personally identifiable information) – sensitive operations, such as *Profile update*
API testing (JMeter, Postman)	RESTful API testing with injection payloads
Behavior-driven development (BDD) testing (Robot Framework or Gauntlt)	When a BDD framework is applied to security testing, the purpose is to enhance cross-team communication and enable a non-security team to understand how security is covered
Fuzz Testing	Security payload testing with various injection and buffer-overflow testing
Data-driven testing (DDT)	DDT testing is normally done with fuzzy testing DDT is normally included in the unit testing framework of the programming language

"The Selenium Web UI framework is used to walk through the UI flow for security tools to inspect security issues, JMeter can be used with security payloads to do RESTful API security testing. Robot Framework can be integrated with ZAP to introduce BDD testing into the security testing cycle. Robot Framework is a common keyword driven testing acceptance automation framework, and Gauntlt is a purpose-built for security BDD framework in Ruby" [11].

To produce secure software systems, all aspects of the development environment must be in symbiosis, including the organizational culture, the tools used, the use of libraries and code repositories, the standards and processes used by the organization. The development environment can be considered a major source of potential vulnerabilities, so it is necessary that the development environment to be well-managed and secure in order to be able to produce secure, quality software systems, the end product being the image of the environment in which it was produced. The tools used to develop secure software systems need to be reviewed carefully to ensure that using the tool does not insert new vulnerabilities into the development process. The final results are directly affected by the quality of the tools used in the development process, and the process of choosing the necessary tools must be carefully approached [7]. Manual security testing are still important in creating a completely secure software (such as full penetration tests or security audits), but organizations must automate security testing and perform these for the most part, preferably with every change to applications or computing infrastructure. Security must be incorporated into every part of the development process to ensure a continuous integration (CI) and to reduce compliance costs [4]. Security requires a large part of our attention in the development process, being a key aspect in the quality assurance process.

4. Conclusions

In this paper were introduced notions such as automated testing, the most used types of automated testing, the benefits of automated testing, included in section 2 of this study. The automation process is in full development, it monopolizes most fields and especially in the creation of high-quality software is present. It was wanted to explain these notions as clearly as possible, and after that the basis of software development will be discussed in detail: unit testing. The notion of unit testing represents the key to starting software testing with purpose to ensure correctness and quality of software. In this section the purpose was to exemplify why it is important this type of test, how such a test is composed and how it differs from other higher-level tests. And the same with unit tests, which ensure quality in the testing environment, security testing aims with maintaining quality high. This process evidence that system and information are safe and reliable, and unauthorized entries are not accepted. In section 3 the subject is discussed at length, differentiates from the previously presented notions and some suggestions of Android security testing tools, a suite of automation frameworks and top ten web application security risks. The research will be continued by constructing more advanced cases scenario and an implementation of the notions discussed.

References

[1]. Flaviu Fuior, "An overview of some tools for automated testing of software applications," in Romanian Journal of Information Technology and Automatic Control, Vol. 29, No. 3, 97-106, 2019.

[2]. Haneen A., Maham K., Zainab S., Muhammad I.B., Saima C., Furkh Z., Muhammad J., Summiyah S., Shahid N.B., "A Comparative Analysis of Quality Assurance of Mobile Applications using Automated Testing Tools" in International Journal of Advanced Computer Science and Applications,Vol. 8, No.7, 2017.

[3]. Alexander Aubert. (2020, June 25th). "Why invest in unit testing?". Available: https://blog.atinternet.com/en/why-invest-in-unit-testing/.

[4]. Oliver Moradov. (2022, May 29th). "Security Testing: Types, Tools, and Best Practices". Available: https://brightsec.com/blog/security-testing/.

[5]. Tony Hsiang-Chih Hsu, "Android Security Testing" in Practical Security Automation and Testing, 2019, ch.7, Available: https://learning.oreilly.com/library/view/practical-security-automation/9781789802023/1f29016b-1353-4061-81c2-1a82a45d1ea6.xhtml.

[6]. Alexander Tarlinder, "Developer Testing Activities" in Developer Testing: Building Quality into Software, 2016, ch. 1, pp. 1-8.

[7]. Erik Fretheim, Marie Deschene, "The Development Environment", in Secure Software System, ch. 13, pp. 221-227.

[8]. Andrew van der Stock, Brian Glas, Neil Smithline, Torsten Gigler. (2021 September 24). "OWASP Top 10:2021", Available: https://owasp.org/Top10/.

[9]. LAMBDATEST, "Unit Testing Tutorial: A comprehensive Guide With Examples and Best Practices". Available: https://www.lambdatest.com/learning-hub/unit-testing#n.

[10]. Richard Wells. (2018, June 6th). "What is unit testing?", Available: https://www.linkedin.com/learning/unit-testing-and-test-driven-development-in-python/what-is-unit-testing?autoplay=true&u=2037052.

[11]. Tony Hsiang-Chih Hsu, "Automating existing security testing" in Practical Security Automation and Testing, 2019, ch. 2, Available: https://learning.oreilly.com/library/view/practical-security-automation/9781789802023/f3a726d8-f7a4-4cd0-a0c4-fe694e30bb69.xhtml.

[12]. Tony Hsiang-Chih Hsu, "Project Background And Automation Approach" in Practical Security Automation and Testing, 2019, ch. 10, Available: https://learning.oreilly.com/library/view/practical-security-automation/9781789802023/34cbdd6a-f26f-4ca0-8c83-a2f125615967.xhtml.

[13]. Alexander Tarlinder, "Almost Unit Tests" in Developer Testing: Building Quality into Software, 2016, ch. 11, pp. 151-157.

[14]. Alexander Tarlinder, "The Testing Vocabulary" in Developer Testing: Building Quality into Software, 2016, ch. 3, pp. 21-36.

An Efficient Security System That Uses Artificial Intelligence to Detect and Identify Objects

Grigor PARANGONI, Dumitru-Iulian NĂSTAC
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
geriparangoni@gmail.com, iulian.nastac@upb.ro

Abstract

Object identification is a significant task in computer vision due to the complexity and diversity of the things that must be detected. Rapid response time and precision are critical, particularly in security applications. We investigate YOLOv5, one of the most efficient object identification algorithms on the market, in this study. Our goal is to show how successful this algorithm is in a security system when compared to other existing alternatives. We also created a web interface that allows visitors to view the live camera feed and track the object detection process in real time. We provide our action plan, as well as the technology and knowledge required to complete this project. The suggested security system consists of a high-resolution surveillance camera and the YOLOv5 object detection algorithm. We created and implemented this system using computer programming and image processing technologies. Our findings reveal that the YOLOv5 algorithm outperforms alternative solutions in terms of speed and accuracy.

Index terms: Object detection, YOLOv5, computer vision, security system, image processing

1. Introduction

Object detection is a well-studied subject in computer vision, with a variety of strategies available to meet the growing demand for effective object recognition models. The complexity and diversity of the items that must be detected, particularly in security applications, present considerable obstacles to the development of effective object detection algorithms. In such applications, quick response time and excellent accuracy are critical. In this paper we investigate YOLOv5, one of the most efficient object detection algorithms on the market and show how it works in a security system.

We begin by reviewing the technologies and knowledge required to complete this project. Then we show our action plan, the security system we intend to build, and the results gained. The suggested security system consists of a high-resolution surveillance camera and the YOLOv5 object detection algorithm. We created and implemented this system using computer programming and image processing technologies. We also created a web interface that allows visitors to view the live camera feed and track the object detection process in real time.

The efficacy of the YOLOv5 algorithm in detecting and recognizing objects in security applications is the topic of this paper. Our goals are to demonstrate the algorithm's effectiveness in contrast to other existing solutions and to create a user-friendly interface for monitoring the system's performance.

2. Related work

Object identification has been the subject of countless research in recent years, with various algorithms being developed to tackle this difficult issue. Faster R-CNN, YOLO, SSD, and RetinaNet are among the most popular algorithms. Each algorithm has advantages and disadvantages, and researchers are always exploring for new ways to improve object identification accuracy and speed. Several studies have examined the performance of these algorithms, with YOLOv5 outperforming them in many circumstances in terms of speed and accuracy. Moreover, some research has concentrated on combining object detection with security systems such as surveillance cameras. Many of these investigations, however, have used standard object identification methods, and there has been little study on the usefulness of YOLOv5 in such systems. Our research intends to fill this need by investigating the usage of YOLOv5 in a security system and comparing its performance to existing alternatives.

3. The implementation of YOLO

In this section, we demonstrate the YOLOv5 algorithm implementation in our suggested security system. The system is made up of a high-resolution security camera and a computer that is linked to it and runs the YOLOv5 algorithm. We created and implemented this system using computer programming and image processing technologies [1].

We begin by preprocessing the photos captured by the security camera in order to increase their quality and remove noise. The preprocessed photos are then run through the YOLOv5 algorithm to recognize and classify objects in the image. The program can accurately detect people, vehicles, and other things in images.

A tracking technique is then used to maintain track of the observed items' motions across frames. This enables the system to identify potential security threats and detect questionable behavior. The use of YOLOv5 in our security system has yielded encouraging results in terms of detecting and preventing security risks. In real-world circumstances, the system successfully detected and stopped several security breaches [2].

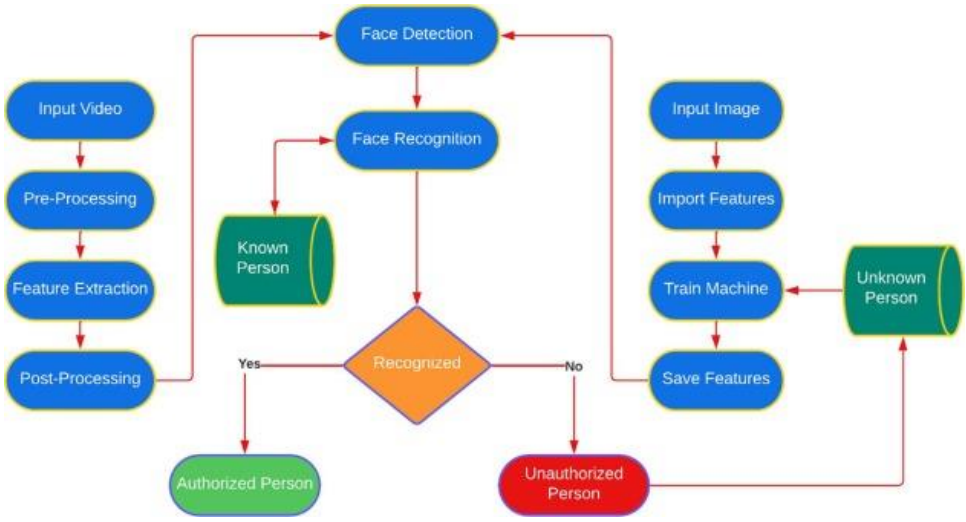


Fig. 1. Proposed Model for Security

4. Training over the YOLOv5 algorithm

A vast quantity of data and processing resources are required to train an object detection model like YOLOv5. To increase its capacity to recognize items in real-world circumstances, the YOLOv5

algorithm is trained on a dataset including a varied range of objects and backdrop sceneries. The YOLOv5 algorithm examines and learns from the dataset during training to enhance its accuracy and speed [3].

The initial stage in training the YOLOv5 algorithm is to acquire a large collection of photos containing items of interest. After that, the dataset is tagged with bounding boxes around the items of interest. The labeled data is used to train the YOLOv5 algorithm using supervised learning, in which the system learns to anticipate the position and class of objects in a picture using the labeled data.

The YOLOv5 method employs a deep learning technique known as a convolutional neural network (CNN). CNNs can learn complicated patterns and characteristics from pictures, making them ideal for object recognition applications. YOLOv5 employs a modified version of the well-known EfficientNet architecture, allowing for quicker and more accurate training [4].

The YOLOv5 algorithm learns to distinguish distinct objects based on their visual properties, such as color, texture, and form, during training. It also learns to distinguish between objects and background scenes, which aids in the reduction of false positive detections.

To summarize, training the YOLOv5 method entails gathering and classifying a large collection of photos, which is then used to train the system using supervised learning [5]. The YOLOv5 algorithm use a CNN architecture to learn complicated patterns and characteristics from photos. The YOLOv5 algorithm learns to recognize and discriminate between objects and background scenes through training, increasing its accuracy and speed in object identification tasks.

5. YOLOv5 compared with other algorithms

Several object detection algorithms have been investigated from novel perspectives in recent years. YOLO (You Only Look Once) is a prominent neural network-based algorithm capable of extracting complicated information from photos and applying them to identify objects. YOLOv5, the most recent version of this algorithm, has demonstrated substantial gains in terms of speed and accuracy. We compare YOLOv5 to various existing solutions in this section [6].

The speed and accuracy of algorithms is one of the key issues in object detection. Objects can be detected in real time by faster algorithms, which is critical for security applications. YOLOv5 is one of the quickest object identification algorithms, outperforming popular solutions such as Faster R-CNN, RetinaNet, and EfficientDet [7]. YOLOv5 has also achieved cutting-edge accuracy on various object detection benchmarks, including COCO, PASCAL VOC, and Open Images. YOLOv5 outperforms other algorithms in detecting small objects, which is critical in security applications [8].

Fig. 2. YOLO V5 Object detection

6. Conclusion

Implementing the YOLOv5 algorithm in our suggested security system has proven to be effective in real-time object detection and recognition. The device can be deployed in a variety of locations, including airports, public spaces, and other venues where security is a top priority.

Furthermore, this implementation demonstrates the capability of AI-based solutions in tackling security concerns. Traditional techniques of monitoring and analyzing data can be time-consuming and inefficient with the rising volume of data collected by surveillance systems. However, by implementing AI-based solutions such as YOLOv5, real-time object detection and recognition can be accomplished with improved accuracy and speed.

Despite the encouraging results, the YOLOv5 algorithm has certain limitations, such as its inability to recognize objects in adverse weather conditions or when the object is partially concealed. Furthermore, there are still worries about the ethical and privacy consequences of deploying AI-based security solutions.

We foresee more enhancements to the YOLOv5 algorithm in the future, as well as the development of new AI-based security solutions. These improvements may result in the development of more efficient and precise surveillance systems capable of preventing security threats and ensuring public safety.

References

[1]. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.

[2]. Liew, J. W. M., & Law, N. F. (2018). A review on video-based human activity recognition. arXiv preprint arXiv:1807.06306.

[3]. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[4]. Saha, S., & Chowdhury, A. S. (2020). Artificial intelligence-based smart security and surveillance system: a review. Multimedia Systems, 26(2), 149-164.

[5]. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

[6]. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[7]. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). ImageNet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211-252.

[8]. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (pp. 6105-6114).

Cybercrimes in the Metaverse: Challenges and Solutions

Alexandru-Valentin TEODOROV
Faculty of Business Administration, Bucharest University of Economic Studies, Romania
alexandruvalentin.teodorov@stud.ase.ro

Abstract

The emergence of the metaverse has brought about novel opportunities for user interaction and commerce. However, with these new technologies also comes the rise of cybercrime as well as new types of cybercrime. The current article aims to delve into the manifold forms of cybercrime that loom large in the metaverse - from virtual theft and identity theft, to cyberbullying. At the same time, the paper explores the multiple challenges that come with preventing and addressing such crimes, such as the arduous task of identifying perpetrators and the inefficacy of law enforcement as well as the necessity for new laws created for the metaverse. In conclusion, the study will explore viable solutions for preventing and mitigating cybercrimes in the metaverse. The article aims to do exploratory research of cybercrimes and technological solutions such as blockchain and AI, as well as policy and legal changes, so that the metaverse can be a safe and secure haven for all users.

Index terms: AI policy, blockchain, cybercrime, metaverse, prevention, response

1. Introduction

The metaverse is a digital realm where users can interact with each other through virtual avatars and has become a popular platform for socializing, gaming, and commerce. However, as the number of users in the metaverse increases, so does the risks of these worlds increase. Cybercrimes on the metaverse refer to illegal activities committed in virtual worlds, social games, and other digital spaces where users can interact with each other through avatars or other digital representations.

Studies have shown that cybercrime in the metaverse can have serious consequences for users of these platforms. These consequences include but are not limited to, financial losses, emotional distress, and damage to reputation and privacy. In a survey of 3,000 users of virtual worlds and social games, “26% reported experiencing cybercrime or knowing someone who has been a victim of cybercrime in the metaverse” [1]. According to Time Magazine, in 2021 alone, “14 billion worth of cryptocurrencies was sent to “illicit” wallet addresses”. [2] What is more, phishing websites have become more and more genuine looking, as cybercriminals become more skillful, making it difficult, to discern between scam and reality.

To combat cybercrime in the metaverse, it is important to understand how it works. This includes a study of threats, the technological fortes and liabilities, as well as, social, and psychological factors that contribute to its occurrence. Legal frameworks, specific cyber policies, as well as user education and awareness campaigns, can also play a role in preventing cybercrime in the metaverse.

2. Background

The metaverse, as defined by the Oxford Dictionary, "a virtual reality space in which users can interact with an environment generated by computer and with other users". The term 'metaverse' appeared in 1992, when it was first introduced by Neal Stephenson in the science-fiction novel Snow Crash [3]. The metaverse has since been developed by researchers, game developers, and technology companies to create various virtual worlds and online games that provide immersive experiences and social interactions. However, with its increasing popularity and complexity of virtual platforms, so has the occurrence of cybercrime in the metaverse grown to a concerning amount [4].

Cybercrime in the metaverse can take many forms, such as hacking, phishing, identity theft, virtual property theft, cyberbullying, and sexual harassment. Cybercriminals can exploit vulnerabilities in the virtual environment, the user devices, or even the user itself, in order to steal sensitive information or control users' accounts. These crimes can have serious consequences, such as financial losses, emotional distress, and damage to reputation and privacy. It is of upmost importance to understand, the different types of cybercrime, and the factors that contribute to their occurrence so that one can prevent these types of attacks.

Hacking is one type of cybercrime in the metaverse, where hackers exploit vulnerabilities in virtual environments or user devices, in order to gain unauthorized access to user accounts or steal data, property or sensitive information. Phishing attacks can occur through a variety of means, including email, instant messaging, and social media. In the metaverse, phishing attacks may be disguised as messages from trusted sources, such as virtual world administrators or other users. Identity theft is another common cybercrime in the metaverse, where cybercriminals use stolen login credentials or personal information to impersonate users or create fake accounts to commit fraud or other crimes.

Virtual property theft involves stealing virtual goods, virtual currency, or other assets that users have acquired through gameplay or purchases. Cyberbullying and sexual harassment are also concerns, where users can use the anonymity of the virtual environment to harass or intimidate others.

To combat cybercrime in the metaverse, users, developers, and law enforcement agencies must work together to implement effective prevention and response strategies. This can involve using technological solutions such as encryption, firewalls, and other security measures to protect user data and prevent cybercrime. Developers can also implement reporting and blocking features to enable users to report cybercrime and prevent further victimization. Legislators may need to cooperate with virtual world operators to investigate and prosecute cybercriminals who use the metaverse to commit crimes. Legal frameworks can also help combat cybercrime in the metaverse. Some countries have enacted laws that address cybercrime, including cyberbullying and virtual property theft [5]. However, the legality of virtual crimes is still being debated, and the lack of a clear legal framework can make it difficult to hold cybercriminals responsible for their actions in the metaverse.

To combat cybercrimes in the metaverse, users should take steps to protect their devices and personal information, such as using strong passwords and avoiding sharing personal information with strangers. User education and awareness campaigns can also be effective in preventing cybercrime in the metaverse. Users should be advised to use strong passwords and avoid sharing personal information with strangers. They should also be made aware of the risks of cybercrime and how to report it to developers or legal entities.

3. Cybercrimes in the Metaverse

As the metaverse continues to evolve and become more immersive, it also becomes more vulnerable to cybercrimes. In this chapter, we will explore various types of cybercrimes that may occur in the metaverse, including virtual property theft, identity theft, cyberbullying, harassment, and

phishing. We will also discuss the potential economic impact of these crimes on both individuals and businesses in the metaverse. Finally, we will examine some current and future measures that can be taken to prevent and mitigate these cybercrimes.

3.1. Virtual Property Theft

Virtual property theft is one of the most common types of cybercrime in the metaverse. In virtual worlds, users can acquire and accumulate virtual assets such as virtual currency, virtual real estate, and virtual items. These assets can have real-world value, and as a result, they can be targeted by cybercriminals.

This phenomenon can occur through various means, including hacking, phishing, and social engineering. Cybercriminals may steal login credentials and gain access to a user's virtual property. They may also use phishing tactics to trick users into giving up their virtual assets. Additionally, virtual property theft can occur through social engineering tactics, where a cybercriminal gains the trust of a user and then steals their virtual assets. Other tactics applied by cybercriminals recreate an app or a website in order to trick users. The theft of virtual goods such as cryptocurrencies, NFTs or other assets, results not only of the financial deficit, but also in the loss of trust of users towards the companies managing such services [6].

3.2. Identity Theft

Identity theft is another common cybercrime in the metaverse. In virtual worlds, users can create and customize their avatars to represent themselves. These avatars can be highly personalized and can even have real-world characteristics, such as a user's name or likeness. As a result, they can be targeted by cybercriminals looking to steal a user's identity [7].

Identity theft in the metaverse can occur through various means, including hacking, phishing, and social engineering. Cybercriminals may steal login credentials and gain access to a user's avatar (online identity), or they may use phishing tactics to trick users into giving up their avatar information. Additionally, cybercriminals may use social engineering tactics to gain a user's trust and then use that trust to steal their avatar information.

3.3. Cyberbullying

Cyberbullying is another growing concern in the metaverse. In virtual worlds, users can interact with each other in a variety of ways, including text chat, voice chat, and even physical interactions between avatars. However, these interactions can also be used to harass and bully other users.

This phenomenon can take many forms in the metaverse, including verbal abuse, harassment, and even physical assault between avatars. Cyberbullies may use text chat or voice chat to harass and intimidate other users, or they may use physical interactions between avatars to physically harm other users [8].

3.4. Phishing

Phishing is a common cybercrime in the metaverse, and it involves the use of deceptive tactics to trick users into revealing their login credentials or other sensitive information. Phishing attacks can occur through various means, including email, instant messaging, and social media.

In the metaverse, phishing attacks may be disguised as messages from trusted sources, such as virtual world administrators or other users. Cybercriminals may use different techniques to trick users into revealing their login credentials or other sensitive data, which can then be used to steal virtual property or even real-world financial information [9].

3.5. Economic Impact

The economic impact of cybercrimes in the metaverse can be significant, both for individuals and for businesses. Virtual property theft can result in the loss of virtual assets that have real-world value, such as virtual currency or virtual real estate [6]. Additionally, identity theft can result in the loss of personal information that can be used to steal real-world financial information [7].

Cyberbullying can also have a significant economic impact, particularly for businesses that operate in the metaverse. Businesses may suffer reputational damage if cyberbullying occurs within their virtual world. In such cases, businesses they may also face legal liabilities should they not take adequate measures to prevent cyberbullying [4].

4. Prevention and Response in the Metaverse

As discussed in the previous chapter, cybercrimes in the metaverse can have severe impacts on the economy, society and on the end consumer of digital products. In this chapter, we will explore various prevention and response measures that can be taken to mitigate the risks enumerated previously.

4.1. Prevention Measures

Prevention measures are proactive measures that can be taken to reduce the risk of cybercrimes in the metaverse. These measures include education, technology, and policy [10].

4.1.1. Education

Education is one of the most important prevention measures that can be taken to reduce the risk of cybercrimes in the metaverse. Users should be educated on the risks associated with using virtual worlds and how to protect themselves. Education can be delivered in various ways, including through online tutorials, workshops, and training programs. Cyber or prevention education in the online environment should include information on creating strong passwords, recognizing phishing attacks, and reporting cyberbullying and other types of cybercrimes [11].

Virtual world administrators can also incorporate education into their user agreements and terms of service. What is more, there are some platforms that have created massive educational campaigns in order to inform and educate users on the dangers that the online medium might pose. A good example in this case is the ING Bank of Romania, that not only wrote the information on their website [12], but went as far as to create a campaign with influencers, in order to educate the user on how not to act online [13].

4.1.2. Technology

Technology can also play a significant role in preventing cybercrimes in the metaverse. There are various technological solutions that can be implemented to reduce the risk of cybercrimes, including authentication, encryption, and monitoring [14]. Another way in which a user can ensure that he/she is safe, is to always keep the software up to date.

4.1.3. Policy

Policy is another important prevention measure that can be used to reduce the risk of cybercrimes in the metaverse. Policies can be implemented at the organizational level or at the virtual world level.

Virtual world policies can be implemented to establish rules and guidelines for users of the virtual world. These policies can address issues such as acceptable behavior, virtual asset ownership, and reporting procedures for cybercrimes.

Organizational policies can be implemented to establish rules and guidelines for employees who use virtual worlds for work purposes. These policies can address issues such as acceptable use, data security, and reporting procedures.

4.2. Authentication

Authentication is the process of verifying the identity of a user. In the metaverse, authentication can be used to prevent unauthorized access to user accounts and virtual assets. One effective method of authentication is two-factor authentication (2FA), which requires users to provide two forms of identification, such as a password (something you know) and a code sent to their mobile device [15], or a token (something you own), or it can require a scan of the user – either a fingerprint or face scan, in the case of mobile devices (something you are).

4.3. Encryption

Encryption is the process of converting data into a code to prevent unauthorized access. Encryption can be used to protect user data, virtual asset transactions, and other sensitive information in the metaverse. Virtual world administrators can implement encryption by using secure communication protocols, such as Transport Layer Security (TLS) and Secure Sockets Layer (SSL) [16].

4.4. Monitoring

Monitoring is the process of observing user activity for suspicious behavior. In the metaverse, monitoring can be used to detect cyberbullying, phishing attacks, and other types of cybercrimes. Virtual world administrators can monitor user activity by analyzing user data and implementing automated monitoring tools.

4.5. Response Measures

Response measures are reactive measures that can be taken in response to cybercrimes in the metaverse. These measures include investigation, prosecution, and remediation.

4.5.1. Investigation

Investigation is the process of gathering evidence and information to identify the perpetrator of a cybercrime. In the metaverse, investigation can be challenging due to the anonymous nature of virtual worlds. However, virtual world administrators can implement tools to track user activity and investigate cybercrimes [17].

4.5.2. Prosecution

Prosecution is the legal process of pursuing criminal charges against a perpetrator of a cybercrime. In the metaverse, prosecution can be challenging due to jurisdictional issues and the difficulty of identifying perpetrators. However, virtual world administrators can work with law enforcement agencies to pursue legal action against cybercriminals [18].

4.5.3. Remediation

Remediation is the process of restoring the victim of a cybercrime to their pre-incident state. In the metaverse, remediation can involve restoring virtual assets that were stolen or destroyed, as well as providing counseling services for victims of cyberbullying, and implementing stronger security measures to prevent future incidents.

5. Future directions and challenges

The metaverse is a constantly evolving concept, making predicting future cybercrime trends a challenge. Therefore, analyzing current developments and trends reveals some potential future directions and challenges.

One possible direction is the increased utilization of virtual and augmented reality technologies, enabling cybercriminals to conduct more convincing attacks. For instance, virtual phishing attacks could be carried out by creating fake physical objects through augmented reality.

Another possible direction is the growing incorporation of artificial intelligence and machine learning, facilitating automation of various cybercrimes such as phishing, fraud, and spamming. Furthermore, this technology could enable the creation of more sophisticated attacks that are harder to detect and thwart.

As the metaverse gains popularity, more people unfamiliar with its risks will begin using it, leading to more people falling prey to cybercrimes, which could result in new challenges for prevention and response. Preventing cybercrimes in the metaverse is fraught with challenges. For example, the lack of regulation makes it challenging for legislators to prosecute cybercriminals. Moreover, the decentralized nature of the metaverse makes it tough to gather and analyze data on cybercrimes, creating obstacles for policymakers and researchers.

Despite these challenges, there are some possible solutions. One possible solution is the development of blockchain technology to create more secure and transparent virtual environments. Additionally, educating metaverse users on the risks and threats of the virtual world could help them protect themselves from cybercrimes, through teaching them how to identify phishing scams, secure virtual assets, and report crimes to law enforcement agencies.

Overall, the future of cybercrimes in the metaverse, evolves along with the evolution of security systems. It is for this reason that permanent betterment should be at the basis of cybersecurity.

6. Conclusion

The potential risks and threats of cybercrimes in the metaverse have been explored in this article, along with the preventive measures and response strategies that can be implemented to mitigate them. With the metaverse constantly evolving and becoming more integrated into our daily lives, the risks and threats of cybercrimes are likely to increase. This is especially true given the lack of regulation and oversight in the current metaverse landscape.

To address these risks, it is essential to establish clear organizational policies and guidelines, use encryption technologies, and implement incident response plans. However, there are still significant challenges in implementing these strategies effectively. As the technology continues to evolve, new methods and techniques used by cybercriminals to exploit vulnerabilities will also emerge. Therefore, it is crucial that policymakers, researchers, and industry stakeholders work together to develop new approaches and technologies that can effectively address these challenges.

In summary, cybercrimes in the metaverse are a growing concern that requires serious attention from all stakeholders. By establishing clear policies, implementing effective preventive measures, and developing new technologies and strategies, we can ensure that the metaverse remains a safe and secure environment for all users.

References

[1]. J. van der Meer, J. S. Doorn, and S. R. de Groot, "Cybercrime in the Metaverse: An Empirical Analysis of Cybercrime in Virtual Worlds and Social Games," J. Cybersecurity, vol. 5, no. 2, pp. 41-58, 2019. doi: 10.1093/cybsec/tyz010.

[2]. I. Dodds, “Why Crypto Scams Are Driving an Online Crime Boom — And How to Outsmart Them,” Time, Mar. 29, 2022. https://time.com/6162350/crypto-scams-online-crime-boom/.

[3]. J. S. Brown, "Snow Crash and the Metaverse: A Critical Analysis of Neal Stephenson's Vision," in Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 2005, pp. 1-9.

[4]. Ioannis Hatzilygeroudis, “Metaverse,” Encyclopedia, vol. 2, no. 1, pp. 486–497, Feb. 2022, doi: https://doi.org/10.3390/encyclopedia2010031.

[5]. M. J. H. Overmars and M. A. de Vries, "Virtual World Crime: The Emergence of Cybercrime in the Metaverse," J. Criminol., vol. 4, no. 3, pp. 67-83.

[6]. A. B. Ahmed, "Virtual Property Theft in the Metaverse: Types, Prevention, and Economic Impact," in Proceedings of the 2020 IEEE International Conference on Cybersecurity and Privacy (ICCP), San Francisco, CA, USA, 2020, pp. 1-7.

[7]. T. J. Smith and K. L. Wang, "Identity Theft in the Metaverse: Risks, Implications, and Mitigation Strategies," IEEE Security & Privacy, vol. 18, no. 2, pp. 43-51, Apr. 2020.

[8]. H. Lee and M. Lee, "Cyberbullying in the Metaverse: Characteristics, Consequences, and Countermeasures," in Proceedings of the 2021 IEEE International Conference on Cybersecurity and Cyberforensics (ICCCF), Rome, Italy, 2021, pp. 1-8.

[9]. S. Kim, S. Choi, and K. Lee, "Phishing in the Metaverse: Tactics, Trends, and Detection Techniques," IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 4, pp. 917-930, Jul./Aug. 2021.

[10]. A. Abbasi, M. Naveed and R. Nazir, "A survey on cybercrime in the metaverse," 2019 International Conference on Innovative Computing (ICIC), Leshan, China, 2019, pp. 262-266, doi: 10.1109/ICIC48177.2019.00068.

[11]. S. K. Ghosh, R. Pandey and D. K. Bhattacharyya, "Cybercrime in virtual worlds: A survey," 2015 International Conference on Computing and Network Communications, Trivandrum, India, 2015, pp. 545-548, doi: 10.1109/CoCoNet.2015.7411241.

[12]. “ING Bank Masuri Antiphishing,” Ing.ro, 2023. https://ing.ro/lp/masuri-antiphishing (accessed Apr. 28, 2023).

[13]. I. Romania, “Epic Show Romania Hacker School ENG,” *YouTube*. Nov. 11, 2021. Accessed: Apr. 28, 2023. [YouTube Video]. Available: https://www.youtube.com/watch?v=xxwhEnxd4NQ.

[14]. R. K. Sharma and S. S. Tyagi, "Security issues in virtual worlds," 2017 International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2017, pp. 1212-1215, doi: 10.1109/ICCMC.2017.8074659.

[15]. N. G. Carr and M. F. M. Yassin, "Two-factor authentication in virtual worlds," 2014 International Conference on Information Science, Electronics and Electrical Engineering (ISEEE), Sapporo, Japan, 2014, pp. 1843-1846, doi: 10.1109/InfoSEEE.2014.6947985.

[16]. Y. Sun, L. Zhao, J. Zhao and Y. Shen, "Design and implementation of security encryption in virtual world," 2011 International Conference on Electric Information and Control Engineering, Wuhan, China, 2011, pp. 2126-2129, doi: 10.1109/ICEICE.2011.5777899.

[17]. S. R. Patil and S. R. Suralkar, "Detection of cyberbullying in virtual world," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 2017, pp. 1-5, doi: 10.1109/ICCUBEA.2017.8329707.

[18]. J. C. Yang, J. C. Chen, and C. H. Wu, "Detection of unauthorized virtual money transactions in online games," Journal of Network and Computer Applications, vol. 33, no. 1, pp. 1-9, 2010.

Financing Terrorism: Economy's Dark Side

Andreea-Mădălina VÂRTEI
Faculty of Economics Sciences, Lucian Blaga University of Sibiu, Romania
andreea.vartei@ulbsibiu.ro

Abstract

In the hidden depths of covert operations and clandestine dealings, the sinew of currency entwines in a sinister ballet, fueling the malevolent fires of terrorism, bestowing upon it the means to unleash havoc and anguish upon unsuspecting souls. Within the intricate web of the global economy, the haunting specter of terrorist financing looms large, its tendrils entangling a labyrinthine network of cartels funding jihadist endeavors, all entwined within the ideological struggle between material wealth and religious fervor. The aim of this study is to delve into the intricate layers of terrorist financing. The first layer involves conducting a literature review focusing on the 2000s, providing insights into the subject. Moving on to the second layer, a behavioral analysis of terrorist financiers is presented, highlighting the formation of alliances between terrorists and financiers. The third layer examines the utilization of advanced technology and intelligent materials in combating the issue of terrorist financing. Finally, the research concludes with an overview of the challenges posed by the influence of the black economy in a globalized world.

Index terms: ATF, Black Economy, Financing Terrorism, Hidden transaction, White-Collar Crime

1. Introduction

Terrorism, a persistent menace that knows no borders, continues to pose a significant threat to global peace and security. Behind every act of terror lies a complex web of funding that fuels these operations, allowing extremist groups to carry out their deadly agendas. Unraveling the intricate mechanisms and sources of terrorist financing is crucial in our collective efforts to combat this grave threat.

Terrorism can be seen as a rational call from the jihadists, who focus on power and intimidation to finance their activity and to fulfill their terrorist objectives which cannot be obtained by legal and democratic methods.

Furthermore, we can say that the terrorists weigh the costs and benefits they would have before acting to raise funds simultaneously the costs and risks they assume to create new strategic gates. This happens as a result of achieving the objectives of attracting a large number of financial sources, new members in jihadist organizations or establishing new relationships with other extremist groups, all merging in the hope of achieving political and ideological goals.

Terrorist financing emphasizes the idea of a phenomenon that involves the solicitation, collection or provision of funds with the intention that they can be used to support terrorist acts or organizations, where the bases of financing can be both licit and illicit.

Nonetheless, the term and phenomenon of terrorist financing does not have a generally accepted interpretation, but depends on the circumstances of the terrorist Islamists.

Gaining a comprehensive grasp of the fundamental elements of terrorist financing is of utmost importance. To achieve this, we delve into the illicit channels, money laundering techniques, and the

exploitation of legitimate sectors that form the bedrock of the financial machinery empowering extremists in funding their malevolent endeavors.

To achieve meaningful progress in combating terrorist financing, a multi-faceted approach is required. This encompasses robust legislative measures, effective intelligence gathering and sharing, capacity building, financial institution vigilance, public-private partnerships, and public awareness campaigns to promote financial transparency and integrity.

Besides, the entire process of terrorist financing follows a circuit that includes the request, collection or supply of funds with the intention that they enter the area of support of jihadist groups. The focus of the Mujahideen's power games in financing their activity often competes in hiding the financing and the nature of the financed activity, not so much the sources of the dirty money as in the case of money laundering [11].

Financing terrorism is said to be deeply embedded in and represent a substantial portion of global capitalism, where this association gotten in touch with capitalism originates through the prism of the connection with the idea of competition, economic freedom and yield maximization, which define the capitalist system [6].

This paper examines the field of financing jihadist activities through a three-pillar structure. The first pillar involves a literature review on current approaches to terrorist financing. The second pillar explores the connection between terrorist groups and funding entities, providing a behavioral analysis of terrorist financiers. The third pillar responds to the second pillar and highlights the methods employed by international organizations to combat this phenomenon, with a particular focus on the contribution of advanced technology. The article concludes by summarizing key points regarding the global economy and the challenges posed by terrorist financing.

2. Literature review

Terrorism remains a significant global threat, as terrorist groups continually explore new avenues to finance their operations. In recent years, the utilization of tax havens as a method for terrorist financing has garnered considerable concern. Tax havens, characterized by low tax rates or tax exemptions, stringent bank secrecy laws, and limited financial regulations, have become conducive environments for terrorist financing due to the anonymity and lack of transparency they provide.

In this literature review, we analyze the scope of the issue of financing terrorism, the strategies employed for laundering money and transferring funds to terrorist groups and the efficacy of measures implemented by governments and international organizations to combat these concerns.

Black-collar criminals in high finance are the representatives of illegal activities with direct involvement in terrorism. Brokers, lawyers, financial advisers, bankers or directors of fictitious or underground companies who manage other people's money are some of these organized crime actors [9].

Terrorist groups have strengthened/coagulated the financing of illegal activities by disguising them as goods and cash, creating Islamic charities not controlled by the government, and involving small transfers of money by the capital market [5].

As per the information provided by Stephen, terrorist financing can involve the collaboration and merger between terrorist organizations and non-state actors. Notably, close cooperation has been observed between terrorist organizations and smuggling groups, as well as the utilization of cartels to facilitate their activities [12].

According to William, the measures implemented to combat terrorist financing heavily rely on surveillance and criminal investigation. However, these measures, employed in the context of the `war on terror,` can have unintended repercussions on the financial transactions of individuals, regardless of their citizenship status [16].

Jeffrey Simser underlined that the system aimed at countering the financing of terrorism can be enhanced to reduce costs and risks for financial institutions while increasing actionable intelligence. Striking a balance is crucial between the defined objective, actionable intelligence, and the mechanism employed to achieve that objective [15]. In the perception of Katarzyna Bilicka and Clemens Fuest, countries categorized as tax havens have consistently entered into a greater number of TIEAs (Tax Information Exchange Agreements) with nations that they share stronger economic connections with, such as through foreign direct investment. In this way, TIEAs illustrate tax information exchange agreements that each tax haven is obliged to sign with at least two other countries, in order to position itself in accordance with OECD standards [4].

Frederic Compin has analyzed in his paper called `Terrorism financing and money laundering: Two sides of the same coin?`, the connection between a terrorist and a financial criminal. The reaction for this highlighted that financial criminals focus to maximize investments and minimize risks, while terrorists seek to maximize risks in order to achieve their objectives. Nevertheless, both rely on current data, financial instability, and market capitalism to carry out their actions [7].

From the point of view of Gheorghe Cosmin Manea and Cristian Valeriu Păun, terrorist networks continue to benefit from the flow of illegal resources that circulate through tax evaded regimes that do not cooperate for rules, agreements, standards, treaties, conventions or regulations against the terrorist apparatus [13].

According to Noura Ahmed Al-Suwaidi and Haitham Nobanee approach the directions of anti-terrorist financing (ATF) and the main perspective, the terrorist groups often operate in secrecy through closed networks and industries with low levels of transparency. Despite this, terrorists may offer lucrative economic incentives, such as smuggled cigarettes, fake products, illegal drugs, and supposed charitable donations [1].

Michele Sabatino described the terrorism financing process as an illustration of globalization, where globalism was a bridge between organized crimes and the global economy. Essentially, money from illegal activities is laundered by criminals or terrorists and then used to fund new illicit ventures. Money laundering is the preliminary stage of financing terrorism, which is why the so-called tax havens appeared, where everything appears legal in the legislative loopholes so that the chain of dirty money cannot be broken [14].

Carolyn Alfieri highlighted the idea that terrorist groups are increasingly employing new methods, such as cryptocurrency, to generate funds. Cryptocurrencies facilitate seamless global transactions for these groups, enabling easy sending and receiving of funds. Examples involving Hamas, an al-Qaeda affiliate, and ISIS highlight how terrorists are utilizing cryptocurrencies to generate revenue through donations. In these instances, Hamas's military branch, AQB, solicited Bitcoin donations through their website, while ISIS and an al-Qaeda affiliate developed a scheme involving the purchase of cryptocurrency coupons. These cases underscore the diverse strategies employed by terrorist groups to acquire financial resources [2].

The literature review on terrorist financing provides a comprehensive analysis of the existing research and scholarly works on the subject. It sheds light on the various aspects of terrorist financing, including its underlying causes, funding sources, and methods employed by terrorist groups. The review also highlights the challenges faced in combating terrorist financing and explores the effectiveness of different preventive measures and international initiatives. Overall, the literature review serves as a valuable resource for understanding the complexities of terrorist financing and informing future efforts to address this global security threat.

3. Behavioral patterns exhibited by the terrorist financier

Terrorism funding originates from multiple sources, facilitated by a well-connected network that utilizes various means such as cryptocurrencies, charities, looting, theft, money laundering, shell

companies, tax havens, drugs, arms trade, narcotics, drug trafficking, organ trafficking, and more. Considering these implications for terrorist financing, here is a proposed series of key elements characterizing the financier of jihadist activity [3].

- ***Use of networks and connections***: terrorist financiers frequently depend on vast networks and personal connections to acquire and move funds, involving both internal collaborators and external partners.
- ***Anonymity***: terrorist financiers employ various methods, including money laundering, cryptocurrencies, or intermediaries, in an attempt to hide their identities.
- ***Use of informal financial networks***: terrorist financiers can exploit informal financial channels like hawala, enabling them to transfer money covertly while minimizing traceable evidence.
- ***Unusual financial transactions***: terrorist financiers possess the ability to utilize informal financial channels, such as hawala, enabling them to clandestinely transfer funds while leaving behind limited discernible traces.
- ***Motivation:*** terrorist financiers can be motivated by various goals, such as extremist ideology, support for a radical political or religious cause or a desire to destabilize society.
- ***Connections to terrorist groups***: terrorist financiers can be affiliated with established terrorist organizations or operate independently while holding ideological sympathies and affinities.
- ***Funding scheme***: specific funding models utilized by terrorist financiers may include cash donations, international bank transfers, or the exploitation of informal financial systems.
- ***Use of illicit and alternative financial resources***: terrorist financiers depend on illicit or alternative funding sources, including drug trafficking, smuggling, robbery, extortion, or financial support from state or terrorist-supporting organizations.
- ***Use of advanced technology and encrypted communication channels***: terrorist financiers employ advanced technologies, such as cryptography and encrypted communication channels, to conceal and safeguard fund transfers as well as the exchange of sensitive information.
- ***Financial expertise and capabilities***: there might be indications that terrorist financiers possess advanced understanding of the financial system and techniques related to money laundering.
- ***Engagement with other individuals under suspicion***: terrorist financiers may have connections with other individuals or groups suspected of engaging in terrorism or illegal activities.
- ***Flexibility and adaptability***: terrorist financiers adapt quickly to changes in the regulatory environment and detection methods, adjusting their tactics and using new techniques and funding channels to avoid sanctions and achieve terrorist objectives.
- ***Use of non-profit and charitable organizations***: terrorist financiers can employ non-profit and charitable organizations as conduits to conceal money flows and provide a façade of legitimacy for seemingly lawful financial transactions.
- ***Use of the global financial system***: terrorist financiers endeavor to exploit vulnerabilities and loopholes in the international financial system, enabling them to move funds across borders and evade detection.
- ***Use of corruption and infiltration of financial institutions***: terrorist financiers may employ corruption and infiltration of financial institutions to acquire confidential information, evade investigation, or facilitate illicit transfers of funds.

A crucial aspect is to acknowledge that these characteristics are general and can vary based on terrorist groups and specific contexts. Authorities and law enforcement agencies are continuously striving to identify and counter these attributes and behaviors, aiming to combat terrorist financing and prevent acts of terrorism.

4. Technological weapons

The institutional brigade, composed of various specialized institutions dedicated to detecting and halting terrorist financing, employs cutting-edge weaponry in its battle against the phenomenon. At the forefront stands the Financial Action Task Force (FATF), the main governing body, armed with the most advanced tools and strategies. These professional bodies propose so-called anti-terrorist financing (ATF) measures, which come as an escape hatch from jihadist problems [10].

To combat terrorist financing, a range of counter-offensive tactics can be employed, leveraging advanced technology and intelligent materials capable of infiltrating the realm of jihadists and dismantling their activities [8].

- ***Natural language processing (NLP) techniques***: using NLP can help identify key words and linguistic signals that indicate terrorist financing in written or verbal communications. This approach can help detect threats and radical speech early.
- ***Facial recognition and fingerprint identification technology***: cutting-edge technologies can be employed to compare facial images and fingerprints with existing databases, facilitating the identification and tracking of individuals affiliated with terrorist organizations. This enables enhanced surveillance and monitoring capabilities to mitigate potential threats.
- ***Social network analysis***: by monitoring and analyzing activities on social media platforms and other social networks, valuable insights can be gained regarding the connections and interactions among individuals and groups suspected of engaging in terrorist financing. Utilizing algorithms for sentiment analysis and identification of communication patterns enables the detection of potential threats.
- ***Big Data analysis***: the application of big data analytics technologies plays a crucial role in processing and identifying intricate patterns and interconnections across various data sources. By harnessing these advanced tools, it becomes possible to uncover elusive terrorist financing schemes that may elude traditional detection methods. Through comprehensive analysis of extensive data sets, hidden relationships and anomalies can be revealed, enabling proactive measures to combat and disrupt terrorist financing activities.
- ***Geospatial data analysis***: by utilizing geospatial information and mapping technologies, it becomes possible to detect geographical areas or regions where suspicious activities or unconventional financial transactions associated with terrorist financing may be occurring. These technologies enable the identification, visualization, and analysis of such activities, facilitating targeted monitoring and intervention efforts. Integrating geospatial data into the detection process enhances the ability to identify and disrupt terrorist financing networks effectively.
- ***Online behavior analysis***: by closely monitoring the online behavior of individuals suspected of terrorist financing, authorities can gather crucial clues regarding their intentions and activities. Engaging in activities such as accessing radical websites, searching for bomb-making information, or participating in suspicious transactions on online platforms can serve as red flags for potential involvement in terrorist financing. Through vigilant scrutiny and analysis of these online behaviors, law enforcement can

identify and investigate individuals involved in illicit financial activities associated with terrorism.

- ***Using automatic image recognition and text analysis***: technology can analyze images and texts to uncover possible connections to terrorism and its funding. This involves examining social media posts, threatening letters, and other related materials associated with terrorist activities.
- ***Alliance with the private sector***: collaboration with banks, financial service providers, and other private sector entities is instrumental in the detection of terrorist financing. These organizations play a crucial role by providing valuable data and information that can contribute to the identification of suspicious financial activities. By fostering knowledge sharing and exchanging experiences, the effectiveness of counter-terrorist financing efforts can be significantly enhanced. The partnership between public and private sectors establishes a robust framework for proactive measures and swift response in combating terrorist financing networks.
- ***Advanced financial transaction monitoring technologies***: implementing financial transaction monitoring and analysis systems is a valuable approach to identify money laundering schemes and suspicious transactions linked to terrorist financing. These advanced technologies enable the detection of irregular patterns and activities within financial transactions. By utilizing algorithms and rules-based systems, these tools can generate.
- ***Use of artificial intelligence and predictive analytics***: can be harnessed to create predictive models aimed at detecting potential risks related to terrorist financing. These models enable the real-time identification of suspicious transactions and behaviors, empowering swift and well-informed decision-making processes.
- ***Communications monitoring technologies***: communications interception and analysis technologies play a crucial role in detecting the exchange of information and financial transactions associated with terrorist financing. These technologies encompass the monitoring of various communication channels such as phone calls, text messages, and encrypted communication to identify potential indicators of terrorist financing activities.
- ***Open-source intelligence analysis (OSINT):*** by leveraging open intelligence sources such as websites, blogs, and public reports, valuable insights and details regarding terrorist financing activities and schemes can be obtained. Through careful monitoring and analysis of these sources, significant connections and information can be uncovered, aiding in the understanding and detection of terrorist financing networks.
- ***Using satellite monitoring technology***: satellite monitoring technology is instrumental in identifying suspicious activities, including the movement of vehicles or construction in sensitive areas. By utilizing satellite imagery, valuable data and contextual information can be obtained, shedding light on potential sources of terrorist funding.
- ***Use of voice recognition technology:*** voice recognition technology enables the identification and analysis of communications and telephone calls made by individuals suspected of engaging in terrorist financing. This technology can help detect distinctive speech patterns and language cues, contributing to the identification and investigation of potential instances of terrorist financing.

Weapons of the technological age encompass sophisticated systems and technologies that heavily rely on electronic, computer, and communications components. Their purpose is to enhance the effectiveness and precision of military endeavors, enabling enhanced surveillance, command, and destructive capabilities.

5. Conclusions

The article emphasizes the importance of enhancing international cooperation in the fight against terrorist financing. Governments should increase the sharing of information and collaborate in implementing both preventive and punitive measures. It is crucial to effectively implement financial security policies and adhere to pertinent international regulations.

The dynamics of the global economy are shaped by the proportion of legal and illegal financial activities. In this regard, the funding of terrorist groups, driven by a desire for destruction, highlights how the economy inadvertently facilitates the growth of illicit activities operating covertly.

Certainty in the world of financing terrorism is far given the fact that it is a clandestine and secret activity, thus: it is not known how big this industry really is, the shrouding in mystery of other sources through which terrorists finance their activity, how much they put the basis of their terrorist activity in tax havens, what is the volume of prohibited transactions that pass through banking institutions, the main channels through which the subsidization of jihadist activity passes, but also what is the number of those who know that they are involved in cases of supporting Islamic terrorists, secret cases in charitable causes. Everything seems to be a nebula if we dig deeper.

The utilization of advanced technology in the fight against terrorist financing can yield substantial advantages, improving the efficiency and precision of financial monitoring. It enables swift identification of suspicious transactions and the prevention of funds reaching terrorist organizations, thus diminishing their operational capacity.

However, it is crucial to acknowledge the dual nature of technology in this context. While it aids in advancing and refining methods for financing from the perspective of terrorist financiers, it also presents challenges that must be addressed. Safeguarding data, ensuring privacy, and fostering collaboration between authorities, the financial sector, and technology providers are vital considerations.

To close, While the fear of death may influence terrorist financing, it is important to note that funding sources for terrorism are diverse, including donations, drug trafficking, ransom payments, and other illicit activities. To address this issue, a comprehensive strategy is necessary. Promoting security, justice, and sustainable development is vital in reducing vulnerabilities to terrorist financing and undermining the appeal of extremism. By addressing underlying causes such as political instability, regional conflicts, and economic disparities, we can create an environment less conducive to terrorist activities.

Essentially, the economic theory of terrorism allows us to draw a prototype of the behavior exhibited by jihadist Islamic organizations. These organizations are driven by two fundamental aspects: money and religion, which both symbolize power. It is this convergence of money and religious ideology that serves as the catalyst for the catastrophic events humanity has witnessed.

References

[1]. Al-Suwaidi, A.N., Nobanee, H., 2020, Anti-money laundering and anti-terrorism financing: a survey of the existing literature and a future research agenda. Available: https://www.emerald.com/insight/content/doi/10.1108/JMLC-03-2020-0029/full/html?utm_source=TrendMD&utm_medium=cpc&utm_campaign=Journal_of_Money_Laundering_Control_TrendMD_0&WT.mc_id=Emerald_TrendMD_0, accessed on April 24, 2023.

[2]. Alfieri, C., 2022, Cryptocurrency and National Security. Available: https://www.criminologyjournal.org/uploads/1/3/6/5/136597491/cryptocurrency_and_national_security.pdf, accessed on April 23, 2023.

[3]. Bantekas, I., 2017, The International Law of Terrorist Financing. Available: https://www.cambridge.org/core/journals/american-journal-of-international-law/article/abs/international-law-of-terrorist-financing/92839FB46DFEFC8ACBAAEB1C2A09F7CF, accessed on April 22, 2023.

[4]. Basile, M., 2010, Going to the Source: Why Al Qaeda's Financial Network Is Likely to Withstand the Current War on Terrorist Financing. Available: https://www.tandfonline.com/doi/abs/10.1080/10576100490438237, accessed on April 24, 2023.

[5]. Bilicka, K., Fuest, C., 2013, With which countries do tax havens share information? Available: https://link.springer.com/article/10.1007/s10797-013-9267-y, accessed on April 24, 2023.

[6]. Campbell, A.M., 2021, Money Laundering, Terrorist Financing, and Tax Evasion. Available: https://link.springer.com/book/10.1007/978-3-030-68876-9, accessed on April 24, 2023.

[7]. Compin, F., 2018, Terrorism financing and money laundering: two sides of the same coin? Available: https://www.emerald.com/insight/content/doi/10.1108/JFC-03-2017-0021/full/html, accessed on April 24, 2023.

[8]. Davis, J., 2020, New Technologies but Old Methods in Terrorism Financing. Available: https://static1.squarespace.com/static/5e399e8c6e9872149fc4a041/t/5f18579880818c6feb0e6a49/1595430886467/CRAAFT+Jessica+Davis.pdf, accessed on April 24, 2023.

[9]. Engdahl, O., 2001, Offshore Financial Centres, Tax Havens and White-Collar Crime: Historical Developments and Contemporary Usage. Available: https://bra.se/download/18.12305534131e173a7f180001915/2001_white-collar_crime_research.pdf#page=154, accessed on April 2024, 2023.

[10]. Financial Action Task Force. Available: https://www.fatf-gafi.org/en/home.html, accessed on April 26, 2023.

[11]. International Monetary Fund, 2004, Anti-Money Laundering/Combating the Financing of Terrorism. Available: https://www.imf.org/external/np/leg/amlcft/eng/aml1.htm, accessed on April 25, 2023.

[12]. Kiser, S.D., 2004, Financing terror: An analysis and simulation for affecting Al Qaeda's financial infrastructure. Available: https://www.proquest.com/openview/e6c25ef3f7feb253ed9af8b53e6d2a70/1?cbl=18750&diss=y&parentSessionId=21ENa1uQm5Rv4KmKl2QqKkNUiAPQE1FQ35%2BbH9HvwGQ%3D&pq-origsite=gscholar&parentSessionId=CTTfWYqkcoKaHWFdM7Rt6cSlBMqRxVtJVavsuYP5SDQ%3D, accessed on April 24, 2023.

[13]. Manea, G.C., Paun, C.V., 2019, Terrorist Threats on the Economic System and Combating Financing Terrorist Organizations. Available: https://sciendo.com/article/10.2478/picbe-2019-0081, accessed on April 24, 2023.

[14]. Sabatino, M., 2020, Crime Treasure Islands: Tax Havens, Tax Evasion and Money Laundering. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3530218, accessed on April 25, 2023.

[15]. Simser, J., 2011, Terrorism financing and the threat to financial institutions. Available: https://www.emerald.com/insight/content/doi/10.1108/13685201111173811/full/html?fullSc=1&fullSc=1&mbSc=1&fullSc=1, accessed on April 24, 2023.

[16]. Vlcek, W., 2007, Surveillance to Combat Terrorist Financing in Europe: Whose Liberty, Whose Security?. Available: https://www.tandfonline.com/doi/abs/10.1080/09662830701442436, accessed on April 25, 2023.

Security Testing for E-Commerce Applications

Alexandru-Petrișor LAZĂRA
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
alexandru.lazara@stud.etti.upb.ro

Abstract

Over the past decade, as the e-Commerce market has evolved into a shopping ecosystem involving multiple devices and store concepts, retailers have been continuously innovating the online shopping experience introducing convenient features like multi-device optimizations, product customization, quick and secure checkout processes, or recurrent payments to attract more customers and influence purchase decisions. The main guidelines that are followed in this paper are revolving around security testing and how it can be performed in the form of manual and automated testing, with aid from automated security tools. This paper looks at the threats e-Commerce Applications are facing in regards with cybersecurity and intends to assist preventing vulnerabilities being exploited by malicious intended users by showing the importance of performing security testing to identify weaknesses, mitigate risks and to raise awareness of the importance of strong security measures and procedures.

Index terms: e-commerce security, security testing, software vulnerability, software risk mitigation, automated security tools

1. Introduction

As e-Commerce continues to evolve, the increased use of digital technology has made it vulnerable to cyber-attacks and fraudulent activities. The implementation of online shopping was accelerated by the pandemic, which registered a major worldwide impact. As reported by the International Trade Administration, global e-Commerce revenue has increased by 19% between the period preceding and following the COVID-19 outbreak in 2020 and reported an additional 19% sales growth for 2020 to the existing 9% regular forecast sales growth rate [1].

Although the market already reached its maturity phase during pandemic and the competition is extremely high and cost intensive in most developed countries, the global e-Commerce market is anticipating to continue its growth trend, with an expected Compound Annual Growth Rate of 11.51% from 2023 to 2027, resulting in a projected market volume of $6.35 trillion and an anticipated user base that amounts to 5,288.5 million worldwide, as Statista is highlighting [2].

Following the above-mentioned trends in the e-Commerce market and given the integration of digital information and technology into the basic operations of the retail activity, the risk of cyber-attacks and fraudulent activities related to online shopping also increased. The repercussions of cybersecurity breaches can lead to operational, financial, reputational, and strategic ramifications for an organization, all of which may result in significant costs [3], yet their effects are not limited to the organization itself, but also extends to its customers which may suffer from the loss of personal information and financial resources. Consequently, ensuring the security of e-Commerce Applications has become a major concern for both businesses and customers.

2. Cybersecurity measures against cyber-attacks

Gartner defines cybersecurity as "the practice of deploying people, policies, processes, and technologies to protect organizations, their critical systems and sensitive information from digital attacks" and enlists some of the most common and notable types of cybersecurity attacks – that are also impacting the e-Commerce market – as follows: phishing and social-engineering-based attacks, internet-facing service risks (including cloud services), password-related account compromises, misuse of information, network-related and Man-in-the-Middle attacks, supply chain attacks, ransomware, and Denial-of-Service attacks (DoS) [3].

In addition, Gartner also provides a list of technical defense measures against cyber-attacks, which includes performing network and perimeter security, endpoint security, application security, data security, Identity and Access Management, and implementing a Zero-Trust Architecture [3]. One critical technical measure that is not mentioned in the list, but is equally important, is software security testing. **Security testing** is a type of software testing – an important phase in the Software Development Lifecycle (SDLC) – that asserts "whether software is vulnerable to cyber-attacks and tests the impact of malicious or unexpected inputs on its operations", providing confirmation that "systems and information are safe and reliable, and that they do not accept unauthorized inputs" [4]. By continuously performing security testing, the software is scanned for weaknesses in its implementation, design, and coding. This help organizations ensuring that software meets the necessary security requirements, by fixing any identified vulnerabilities before cybercriminals take advantage of them, thus avoiding serious security breaches. However, security testing is a specialized testing process and should be conducted by a team of certified testers that have a certain level of expertise in this field.

Furthermore, technology control isn't the only line of defense against cyber-attacks. Building employee awareness and establishing cybersecurity procedures at the organizational level are also essential measures for ensuring an effective cybersecurity strategy. The **ISO/IEC 27000** family of standards define requirements, provide direct support guidance and interpretation for the process to establish, implement, maintain, and improve an Information Security Management System, including risk management, information security controls, and incident management [5].

3. Security Testing

When an application is developed, numerous levels of software testing should be carried out by the development team to detect any potential defects – **bugs** – that may occur and result in an observed failure in the application execution [6]. Left unchecked, many of these bugs can be exploited by attackers, but as already stated, security testing can be used to verify whether Software Applications carry such vulnerabilities and to minimize their associated risks.

Security testing for e-Commerce Applications refer to the process of rigorously searching and identifying vulnerabilities in the shopping application, starting from the fundamental interactions present in all different types of implementations, such as user authentication and authorization, continuing with more specialized elements like promotions, product configurations, shopping carts, payment gateways, or the order management system, which are usually targeted by malicious users. Performing security testing can help highlight weaknesses in the implementations of such elements, as those that permit unauthorized users to force and bypass certain authorization restrictions to modify relevant product information, including descriptions, prices, or quantities, through injection methods as SQLi (also known as SQL injection) – an attack that "consists of insertion or injection of a SQL query via the input data from the client to the application" to gain access to read or modify sensitive data from a database and even execute administration operations on it [7].

Session hijacking is another form of attack that grants the attacker unauthorized access, which consists in "the exploitation of the Web session control mechanism, which is normally managed for a session token". Because TCP (Transmission Control Protocol) connections are used intensively in HTTP (Hypertext Transfer Protocol) communication, the Web Server needs a method to recognize every user's connection. The most advantageous method "depends on a token that the Web Server sends to the client browser after a successful client authentication". There are different ways to compromise the session token, however the most common ones are predictable session token, session Sniffing, Man-in-the-Middle and Man-in-the-Browser attacks, and client-side attacks (such as XSS – Cross Site Scripting, or the use of malicious JavaScript Codes) [8].

To detect and address vulnerabilities in this kind of software, both manual and automated security testing needs to be used. In **manual testing**, the software engineer mainly "assumes the role of a user executing the System Under Test (SUT) to verify its behavior and find any observable defects", while **automated testing** is using "scripts that execute without human intervention to test the SUT's behavior". However, security engineers can fully benefit from automated security testing only if they acquire expertise in using testing techniques together with available testing tools such as **vulnerability scanners and analyzers** [9]. Brian Hambling's ISTQB Foundation Guide describes security testing tools as instruments used "to test the functions that detect security threats and to evaluate the security characteristics of software" and to assess the capacity of the SUT to handle computer viruses, protect data confidentiality and integrity, prevent unauthorized access, carry out authentication checks of users, remain available under a DoS attack, or check non-repudiation attributes of digital signatures [6]. Even if vulnerability scanners and analyzers alone are powerful tools, an even more accurate risk rating can be obtained by combining the mentioned testing types with the services provided by a **penetration tester**, that plays the role of an attacker to find and exploit vulnerabilities. While "penetration testing occurs at the end of the SDLC, the results of the penetration test can provide feedback for tests even in its earlier phases". The penetration test report should provide clear details about the discovered vulnerabilities, including how to fix them, and how a specific vulnerability can be exploited by an attacker. A security team needs to help the development interpret the penetration test report and provide guidance in addressing the found vulnerabilities [10].

It is important to note that testing benefits most from implementing **CI/CD (Continuous Integration and Continuous Delivery)** pipelines. This ensures that testing is an ongoing process and that any issues are identified and resolved quickly, reducing the risk of security breaches and other defects that can negatively impact the software development process.

Regardless of the type of testing, it is crucial to avoid any negative impact on the production environment. Working on a **test environment** that is an exact replica of the production environment, with personal identifiable information anonymized, is an endorsed practice to avoid this risk.

3.1. Manual Security Testing

Staying at the forefront of any application, manual security testing must be carried out for every new iteration of the SDLC, for every new added feature, as well as for every change that is impacting the existing features, as part of the regression testing. To improve this process and increase vulnerability discovery, when a new weakness is discovered during testing, it is essential to create a new test case in the designated testing suite. This will help ensuring that the bug can be addressed and resolved in future iterations of the SDLC. It is also important to document all actions taken during the security testing process, including any changes made to the system, and to ensure that the system is returned to its initial state after the testing is completed.

The payment gateway is one of the most critical and complex element of an e-Commerce application. In this case, manual security testing should be performed to ensure that users cannot make unauthorized electronic payments or that their sensitive payment information is being handled securely to prevent it from being intercepted and used by attackers. As a vendor, it is usually a good

practice to use a payment gateway provider instead of developing an internal one, since they can afford investing more time and resources in security measures like tokenization, "the process of protecting sensitive data by replacing it with an algorithmically generated number called a token"; or by using a payment processing encryption method that uses "keys which are encrypted and decrypted to keep sensitive customer information safe" [11].

Handling information securely is mandatory not only in the interactions with the payment gateway, but in all the aspects that are implying sensitive data. The level of access to data and deciding which data can be provided to users or third-party agents should be highly dependent of context. For instance, to maintain their privacy, when sending notification alerts to users, only essential information should be included. Another example consists in sharing client information to the company responsible with delivering customer's order. In this case, the e-Commerce platform must confirm that only the necessary information like delivery address and contact information is shared, and not something sensible like user's payment details or order content.

Other important aspects to be taken into consideration, as security reports are revealing, are the vulnerabilities within APIs (Application Programming Interfaces) implementation and request responses. Thus, during security testing it is considered good practice to assert whether the best practices in API authentication and authorization are being followed. Between these can be enlisted the usage of transport layer security, usage of access control measures by enabling configuration of different permissions for different API keys, as well as constructing a good management system to protect them [12]. Another good practice that needs to be followed in the security testing process is verifying that the responses received from the API requests (such as error messages, internal application errors or configurations) are not disclosing any confidential and system information.

3.2. Automated Security Testing

In comparison with manual testing, automated security testing is a faster, more effective, and dependable approach to assert the vulnerabilities and risks of a SUT, since it can perform tests more efficiently and generate reports containing information about the identified security weaknesses, as well as recommendations for corrective or preventive measures. Moreover, while this method doesn't require a certified security engineer to conduct the testing, the automated security tests that will be performed still needs to be developed by a security team of software developers (in form of unit and integration tests) and test engineers (in form of system tests) that has the required expertise.

Traditionally, security testing "have been conducted after code has been completed", however with aid from security tools, tests can also be performed "in earlier phases of the SDLC, such as develop and commit, and can run as the code is written, with feedback delivered directly in the Integrated Development Environment (IDE)". This behavior implies that security tools integrated within a technology stack are more accessible to software developers rather than to security testing engineers. As a result, the primary purpose of this kind of tools is to provide a good indication of what the found issues are and how to fix them. In this way, security tool findings can be provided directly to the developers, to enable rapid feedback on any issues, even if in some cases the security team must conduct an additional security analysis [10].

Static Application Security Testing (SAST), or Static Analysis of applications, is a security testing technique that examines an application's source code without executing it, and is used to analyze lexical, grammar, and semantic features, as well as data flow and model checking to detect hidden bugs. The main advantage of this technique consists in its high detection speed, as a Static Analysis tool can quickly check the target code. However, such tools carry "a high false rate in practice, due to the lack of an easy-to-use vulnerability detection model", making it challenging to identify satisfactory results [13] [14]. Between the tools available for performing SAST, we can enlist names as SonarQube or Roslynator. **SonarQube** is an open-source platform developed by SonarSource. It offers the capability to assess the health of an application and to identify newly

introduced issues, provides reports on duplicated code, coding standards, unit tests, code coverage and code complexity, bugs, and security recommendations. It helps analyze the quality on more than 30 programming languages, frameworks, and Infrastructure as Code (IaC) Platforms [15]. **Roslynator** is another powerful open-source SAST tool which consists in a collection of more than 500 analyzers, refactorings and fixes for C#, and is powered by Roslyn, an open-source implementation of both the C# and Visual Basic compilers with an API surface. It provides a wide range of features for improving code quality, including code style analysis, code formatting, and code refactoring [16] [17].

In contrast, **Dynamic Application Security Testing (DAST)**, or Dynamic Analysis of applications, requires code execution to observe its behavior. "By monitoring the running states and analyzing the runtime knowledge, Dynamic Analysis tools can detect program bugs precisely". The advantage of this method is its high accuracy, but there still exist the shortcomings of slow speed, low efficiency, high requirements on the technical level of testers and application architecture, poor scalability, and the difficulty to carry out large-scale testing [13] [14]. In industry, there is a series of DAST analyzers as the ones provided by **GitLab**, for scanning websites: DAST proxy-based analyzer and DAST browser-based analyzer; and for scanning APIs: DAST API analyzer. Based on differences between different scan results on the source and target branches, these analyzers can produce a DAST report that GitLab utilizes to identify vulnerabilities in applications [18].

At commit time, automated testing known as SAST, is checking for "known insecure patterns in the source code before being added to the code repository or merged to the main branch". Security tests run at build time are known as **Software Composition Analysis (SCA)**, and are checking for vulnerabilities in libraries, or in used container images. Security tests that run at deploy time, which "allows automated testing on a running application", are the ones referenced as DAST. For development on a new codebase, it can be useful to incorporate security testing from the start, as automated tests can be configured to fail a code build if the tests do not pass [10].

The most popular automated vulnerability discovery technique for an application is currently represented by **Fuzzing, or Fuzz testing**. This entails in generating enormous amounts of random inputs, both expected and unexpected, and tries "to detect exceptions by feeding the generated inputs to the target applications and monitoring the execution states". The application may subsequently execute successfully, raise an exception, or crash. Compared with SAST or DAST, Fuzzing is easier to deploy, has higher extensibility, applicability, and accuracy in actual testing execution, larger scalability and can be carried out with or without the source code. Thanks to this, despite having many disadvantages including the low efficiency and code coverage, Fuzzing has become the most effective and efficient state-of-the-art vulnerability discovery technique [13] [14]. One testing tools that can be used for Fuzzing is **APIFuzzer**, a public HTTP API Testing Framework that reads the API description and, fuzzes the fields to validate if the application can cope with the fuzzed parameters [19].

4. Conclusions

In conclusion, with the increase in cyberattacks, it became critical to implement multiple security measures at both organizational and technology control levels, to protect sensitive data from breaches. Adding multiple layers of security through Defense in Depth is offering stronger protection against potential threats.

However, even with such measures in place, performing security testing remains crucial to find and fix weaknesses as early as possible in the SDLC process, since early testing can prevent costly defects and improve the overall quality of the product. Moreover, supplementing manual testing with automated testing, using different testing tools, and combining internal testing with testing from

external sources, ensure that most of the possible testing scenarios are covered, mitigating risk in software by decreasing the number of potential undetected weaknesses and vulnerabilities.

References

[1]. "Impact of COVID Pandemic on eCommerce", International Trade Administration 10 2021. Available: https://www.trade.gov/impact-covid-pandemic-ecommerce.

[2]. "eCommerce - Worldwide", Statista – The Statistics Portal for Market Data, Market Research and Market Studies, 02 2023. Available: https://www.statista.com/outlook/dmo/ecommerce/worldwide.

[3]. "What Is Cybersecurity?" Gartner, 18 11 2021. Available: https://www.gartner.com/en/topics/cybersecurity.

[4]. "Security Testing: Types, Tools, and Best Practices" Bright Security, 22 05 2022. Available: https://brightsec.com/blog/security-testing/.

[5]. "ISO/IEC 27000:2018(en) Information technology – Security techniques – Information security management systems" International Organization for Standardization, 2018. Available: https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-5:v1:en.

[6]. B. Hambling, P. Morgan, A. Samaroo, G. Thompson and P. Williams, "Software Testing, An ISTQB–ISEB Foundation Guide, Second Edition", British Informatics Society Limited, 2010.

[7]. "SQL Injection", The Open Worldwide Application Security Project. Available: https://owasp.org/www-community/attacks/SQL_Injection.

[8]. "Session hijacking attack", The Open Worldwide Application Security Project. Available: https://owasp.org/www-community/attacks/Session_hijacking_attack.

[9]. V. Garousi and F. Elberzhager, "Test Automation: Not Just for Test Execution" in IEEE Software, vol. 34, no. 2, pp. 90-96, 28 03 2017, doi: 10.1109/MS.2017.34.

[10]. "OWASP Security Culture", The Open Worldwide Application Security Project. Available: https://owasp.org/www-project-security-culture/v10/7-Security_Testing/.

[11]. "Tokenization, Encryption, and Secure Payment Processing", TrueMerchant. Available: https://truemerchant.com/tokenization-encryption-and-secure-payment-processing/.

[12]. S. Scott and G. Neray, "Best practices for REST API security: Authentication and authorization", The Overflow, 06 10 2021. Available: https://stackoverflow.blog/2021/10/06/best-practices-for-authentication-and-authorization-for-rest-apis/.

[13]. J. Li, B. Zhao, and C. Zhang, "Fuzzing: a survey" in Cybersecurity 1, 05 06 2018. Available: https://doi.org/10.1186/s42400-018-0002-y.

[14]. P. Raghu and J. Agrah, "Practical Security Testing of Electronic Commerce Web Applications" in International Journal of Advanced Networking and Applications, 01 08 2021, doi: 10.35444/IJANA.2021.13109.

[15]. Github, SonarSource / sonarqube. Available: https://github.com/SonarSource/sonarqube.

[16]. Github, JosefPihrt / Roslynator. Available: https://github.com/JosefPihrt/Roslynator.

[17]. Github, dotnet / roslyn. Available: https://github.com/dotnet/roslyn.

[18]. Gitlab Docs, "Dynamic Application Security Testing (DAST)". Available: https://docs.gitlab.com/ee/user/application_security/dast/.

[19]. Github, KissPeter / APIFuzzer. Available: https://github.com/KissPeter/APIFuzzer.

Prevention of Widespread Ransomware Cyber-Attacks through the SEAP Platform

Eduard-Ștefan SANDU
Faculty of Applied Sciences, University POLITEHNICA of Bucharest, Romania
edy.eminem@yahoo.com

Abstract

This scientific study aims to explore the potential for launching a cyber-attack through SEAP platform, particularly in light of the increasing use of ransomware as a tool to cause widespread damage to critical infrastructure. The study focuses on the methodology of a ransomware attack on a critical infrastructure, with a specific emphasis on the analysis of the infection process, persistence mechanism, encryption process, recovery prevention, and propagation mechanisms, as well as the communication with command and control servers.

Index terms: critical infrastructures, cyber-attack, cybersecurity, ransomware, SEAP

1. Introduction

The Electronic Public Procurement System, commonly referred to as S.E.A.P., is a web-based platform designed to facilitate electronic public procurement in Romania [1]. As the official website operated by the Authority for the Digitization of Romania (A.D.R.), S.E.A.P. serves as a public utility IT system accessible through the internet, which enables public procurement by electronic means. As the operator of the electronic system, A.D.R. is a legal entity under public law responsible for providing technical support and establishing the specific operating framework required to award public procurement contracts to contracting authorities through the electronic procedure, in accordance with relevant legislation.

S.E.A.P. was developed to increase the efficiency and transparency of public procurement processes in Romania, and to enable wider participation in the procurement process. By using electronic means to carry out procurement, S.E.A.P. aims to reduce the administrative burden associated with traditional paper-based processes, and to promote fair competition among suppliers. Plus, the platform aims to enhance the quality and reliability of procurement data, thereby supporting the development of evidence-based policies and practices.

The use of S.E.A.P. is mandatory for all public procurement procedures in Romania, regardless of the value of the contract being awarded. The platform provides a comprehensive set of tools and features for managing the entire procurement process, from publishing tender notices and receiving bids, to evaluating proposals and awarding contracts. Contracting authorities can use S.E.A.P. to create, publish, and manage procurement procedures, and to communicate with potential bidders throughout the process.

All in all, S.E.A.P. represents a significant step forward in the digitization of public procurement processes in Romania. By providing a secure and efficient means of conducting procurement procedures online, S.E.A.P. has the potential to increase the speed and effectiveness of public procurement, while also improving transparency and accountability.

The use of computers and electronic devices has become an integral part of daily life and their widespread use has resulted in a growing concern for cybersecurity. Malware is a type of malicious software that can cause significant harm to computer systems and electronic devices. Malware can take on various forms, ranging from a simple pop-up window to a sophisticated application that tricks the user into giving away sensitive information. This article will provide an overview of malware, its different types, and the impact it can have on computer systems and electronic devices. In particular, the focus will be on how malware can affect the security and functionality of these systems.

The term "ransomware" is used to describe a type of malicious software that demands a ransom from the victim in exchange for the release of compromised data. The name is derived from the combination of "ransom" and "software". Payment of the ransom is often requested in cryptocurrencies like Bitcoin, but there is no guarantee that the data will be restored even after payment. Regrettably, as ransomware continues to evolve and the anonymity of the internet provides ample opportunities for exploitation, cyber attackers are able to take advantage of legal loopholes and evade detection and retribution, transforming ransomware into a highly attractive and low-risk criminal enterprise.

A significant challenge in combating ransomware attacks is the intricacy involved in identifying the specific variant of ransomware utilized. The attack method utilized by ransomware entails the deletion of the original files during a data-level attack, and custom ransomware systems have been developed to incorporate an executable component. The said executable component comprises the malware's payload, expressed as lines of code that execute on the victim's virtual machine to gain control of the victim's files. The malware's behavior is regulated by a remote server that issues instructions and/or controls to the executable, facilitating the execution of its primary function.

Ransomware, a type of malicious software, infiltrates computer systems and restricts users' ability to access internal data stored on virtual machines. Data recovery in the aftermath of a ransomware attack is a precarious process, often involving the targeted individuals paying a ransom imposed by the attacker. However, the efficacy of such payments in ensuring file recovery remains uncertain.

The diversity of ransomware cyberattacks are manifested in several aspects, including but not limited to the encryption methodology used to restrict access to victim files, the complexity and intricacy of the ransomware's architecture, the dissemination channels employed to propagate the malicious code, and the strategies utilized for data recovery.

Ransomware, as a type of malware, exhibits diverse potentialities for propagation through various vectors, including but not limited to:

- one of the prevalent methods for distributing ransomware is through traffic redirection, where victims are lured into accessing specific websites that present themselves as exploit kits. When a user downloads the program, the malware payload exploits vulnerabilities in the virtual machine, leading to the encryption or locking of systems and files.
- a prevalent vector for ransomware propagation is through e-mail attachments or links that entice the target to access web portals containing malware. This tactic often employs the use of social engineering techniques to entice the recipient to access attachments or follow links that lead to websites containing malware.
- botnets are a pervasive method for ransomware propagation, which involves the distribution of malware through infected programs and legitimate programs with infected code. The botnet infects a large number of devices, which can then be used to distribute the ransomware to other devices, causing widespread damage.

2. Crypto-ransomware

To classify the encryption methodology of victim files, crypto-ransomware can be classified into three distinct types:

- which overwriting the data of the encrypted file as metadata of the original file and this method can be considered a type of file obfuscation, as the encrypted data is hidden within the original file, making it difficult to detect.
- which changing the original file's location from the source directory to the directory where the encryption takes place, along with renaming the encrypted file and aims to further conceal the encrypted data and make the detection process more challenging.
- which creating a new file and entirely replacing the original file and is the most challenging to detect using a data or system-centric detection model that typically looks for bulk deletion behavior of running processes.

There exist diverse techniques for detecting and mitigating ransomware attacks, which possess strengths and limitations that may be leveraged by cybercriminals to render the malware more elusive. Given that the code underlying ransomware is perpetually evolving in terms of complexity, inventiveness, and targeted objectives, effective detection and mitigation of such attacks necessitate continued research and development of novel countermeasures.

In the event of a ransomware attack targeting multiple critical infrastructures that lack a robust cyber defense plan, the probability of successful data recovery is substantially reduced. In such cases, relinquishing the stolen data may be the only viable option. However, if the attacker has made an error in the development of the ransomware source code, another approach could be to attempt to crack the decryption key, which involves solving a mathematical puzzle.

Often, the targeted infrastructures opt to pay the ransom in order to retrieve the decryption key and regain access to their data. However, this course of action does not always guarantee the successful recovery of the stolen data or the attacker's compliance with the agreement. The use of cryptocurrency as the preferred method of payment presents a further layer of complexity for decision-makers who may be unfamiliar with the technology.

I shall present a tabulated account of the most massive ransomware attacks to date, along with their respective dates of detection and the classification of the encryption algorithms employed.

Table 1. Encryption algorithms for ransomware attacks [2]

#	**Malware**	**First known appearance**	**Encryption algorithms**
1	GPcoder	2004	AES - ECB
2	Crypto Locker (Gameover ZeuS)	2013	AES
3	Crypto Wall	2014	AES - CBC
4	CTB Locker		AES - ECB
5	Torrent Locker		AES - CTR | CBC
6	Tesla Crypt	2015	AES - ECB | CBC
7	Crypt Vault		RSA - OAEP
8	Locky	2016	AES - CTR | EBC / AES RSA + ECB
9	Petya		Salsa20
10	Not Petya		MFT - Salsa20 - AES
11	WannaCry		AES - RSA
12	SamSam		RSA
13	Hermes	2017	RSA - AES - CBC
14	Ryuk	2018	RSA - AES - CBC

The data presented in Table 1 suggests that ransomware attackers frequently exploit the encryption algorithms that are trusted by critical infrastructures to safeguard their important data. It is additionally noted that ransomware developers tend to replicate established encryption processes,

with the majority of their efforts directed towards the development of new methods for infiltration and infection.

2.1. Analyzing the classic methodology employed by ransomware

The initial step in the ransomware attack process involves the downloading of the ransomware onto the targeted virtual system. However, this download alone does not necessarily trigger the destructive consequences associated with ransomware. Instead, the download serves as the first phase of a complex process informally referred to as "infection". The downloaded payload, known as a binary, contains the code that directs the actions of the ransomware and is typically transmitted through a variety of infection methods that are tailored to the specificities of the attack.

The second step comprises the code that governs the actions of the ransomware according to the instructions provided by the attacker. The execution instructions differ widely based on the specifics of the attack, making it especially challenging to develop an accurate ransomware execution model. Still, the execution of ransomware can be separated into three general steps, as outlined below [3]:

- **stealth operations**. Before initiating an attack, ransomware must become acquainted with the victim's system and cybercriminals aim to keep their ransomware undetected during this preliminary step to avoid prematurely aborting the attack.
- **suspicious activities**. Ransomware launches the malevolent component of the attack covertly without disclosing its presence and has the ability to evade detection. In the case of locker-ransomware, this phase involves impeding the user interface, while in crypto-ransomware, it entails encrypting the targeted data.
- **obvious actions**. Upon completion of the encryption process, ransomware typically presents a ransom note, which serves as a form of communication to the victim regarding the ransom demand and the method of payment. The ransom note can be displayed as a pop-up window, a file or a wallpaper, and usually includes instructions on how to access the decryption key.

In the absence of a successful defense against the ransomware attack, the critical infrastructure is left with the decision of whether or not to acquiesce to the attacker's demands. If the organization chooses to comply, payment is typically made in the form of digital currency, as these transactions are unregulated and allow for a degree of anonymity. The ransom note will often provide instructions for purchasing digital coins from online exchanges and transferring them to the attacker's designated wallet address.

In the next-to-last stage, pertaining to the specific scenario of a crypto-ransomware attack, it is noteworthy that subsequent to the payment, the ransomware may selectively undertake a course of action, which could involve automatic decryption of the files or provision of a decryption binary to the victim, or it may alternatively fail to render the encrypted data intelligible.

In conclusion, the final step in a cyber-attack involves converting the digital currency reward into national currency. This process poses significant risks as it can compromise the anonymity of the attacker and link them to the crime. To mitigate this risk, attackers may use the shuffling technique to cover their tracks. By obfuscating redemption payments, the origin and destination of funds are concealed, preserving the anonymity of the attacker. Despite its effectiveness, the exchange of digital currencies remains a precarious step, requiring careful consideration and strategic planning to avoid legal and criminal consequences.

2.2. Analyzing Ransomware Attacks on Virtual Machines: Understanding the Effects

Ransomware attacks are executed in five stages, which are as follows:

- during the deployment phase, ransomware undergoes a process whereby its various components are installed in order to effectively infect, encrypt, or crash the targeted

virtual machine. This is achieved through the utilization of drive-by downloads, phishing e-mails, and the exploitation of system vulnerabilities that are readily accessible to the malware.

- within the installation phase of a ransomware attack, a payload is delivered to the targeted system, thereby triggering the start of the infection process. The code responsible for the attack is carefully crafted to avoid detection and subsequently establish communication with command and control (C&C) servers. This is achieved through the manipulation of keys within the operating system registry, which are specifically configured to activate starting code. As the malware propagates itself throughout the network, it utilizes standard processes such as "explorer.exe" or "svchost.exe".
- after the previous steps have been successfully completed, the code starts reaching the command and control server for instructions, which include commands about the types of files to be encrypted, how long to wait before starting the process, and similar commands based on the specifics of the attack. The code will also extract system information such as the IP address, domain name, operating system, and anti-malware products to the C&C channels which can be unencrypted HTTP, encrypted HTTP or anonymous via TOR.
- the fourth stage, which involves destruction, will mark the beginning of an encryption procedure for all targeted files, which involves all types of documents .JPEG, GIF, .CAD, or others without being limited to these circumstances only.
- in the ultimate stage, the coercive strategy of blackmail is employed, whereby a pop-up interface materializes subsequent to the complete encryption of files. This interface contains explicit instructions regarding the remittance of ransom and the consequences that will ensue if the payment is not made within the designated timeframe.

The use of programming languages like C and C++ in the development of ransomware attacks provides attackers with a high degree of flexibility and control. These languages are well-suited to the development of malware due to their low-level access to computer hardware and operating system resources. The development of ransomware attacks is typically characterized by the use of programming languages such as C or C++, and the evolution of these attacks results in increasingly complex and dangerous forms of malware. Understanding the programming languages and evolutionary patterns of ransomware attacks is crucial for developing effective mitigation strategies and protecting against these threats.

The latest generation of ransomware has made it increasingly difficult to gain access to files without complying with the attacker's demands. This is achieved through the use of a combination of symmetric and asymmetric encryption techniques in the encryption process. Symmetric encryption, where the same key is used for both encryption and decryption, is employed to encrypt the bulk of the data quickly. Asymmetric encryption, on the other hand, where a public key is used for encryption and a private key is used for decryption, is used to encrypt the symmetric key used for encryption.

In the following section, we present a breakdown of a particular case of ransomware:

- when the link that contains the malicious code is accessed for the first time, the ransomware will be downloaded on the targeted system.
- after the ransomware is accessed and executed, the execution stage is initiated.
- as per reference [4], the ransomware gathers hashed details about the targeted virtual machine, which may include but are not restricted to CPU, hostname, and RAM. This data is used to recognize the device architecture and generate a type key that is stored in the registry at the following paths: *HKCU\Software\<uniquecomputer id>\<random id>* and *HKCU\Software\[random]*.

- the program has a built-in mechanism to detect the presence of any previous ransomware that may have infected the system. It performs a thorough scan of the system's files and folders to search for any indicators of ransomware activity, such as encrypted files, ransom notes, or suspicious processes running in the background.
- the ransomware is designed to inject itself into the "explorer.exe" extension, which is a legitimate executable file used by Windows to manage the desktop, taskbar, and file explorer. To achieve this, the ransomware creates a new instance of "explorer.exe" and modifies its code to include the malicious payload.
- the ransomware establishes a connection with a command and control server to remain undetected by the target's internal security protocols. This connection allows the ransomware to receive commands from the attacker and send back information about the infected system and makes it more difficult for the target's security team to detect and respond to the ransomware attack in a timely manner.
- reference [5] suggests that ransomware can behave differently depending on the specifics of the attack. Upon execution, the ransomware may start running and performing its malicious activities in one or more of the following ways: % Localappdata %, % ProgramData%, % UserProfile%, % Temp%.
- alternatively, it is also possible to create the extension "svchost.exe" for the injection of the ransomware. This extension is typically located in the "C:\Windows\System32" directory and is a legitimate process used by Windows to host multiple services.
- once the ransomware is injected into the target system, the attackers may use various services to carry out their malicious activities. For instance, the "vssadmin.exe" service can be used to delete or encrypt system copies. In some cases, the "wmic.exe" service may also be employed to delete all files on the infected system.
- the attackers may attempt to prevent backup applications from altering the startup options of the infected system. This is typically achieved by disabling startup recovery using the "bcdedit.exe" extension.
- after disabling backup options and other security measures, the attackers may use various Windows services to encrypt the victim's files. The attackers may use "syskey.exe" to encrypt the victim's files using a randomly generated key and, once the files are encrypted, the attackers may use another Windows tool called "cipher.exe" to manage the encrypted data.
- once the attackers have successfully encrypted the victim's files using tools like "syskey.exe" and "cipher.exe", they notify the command and control (C&C) server to provide the encryption keys along with the target ID and password.
- in order to keep the communication between the attackers' command and control (C&C) server and the infected system secure and private, data traffic is generally executed using an encrypted protocol such as HTTPs or TOR where the attackers can hide the data traffic between the server and the infected system from detection by security tools and network administrators.
- after the encryption process is completed, the files are transformed into a format that is not readily accessible to the victim. Once the encryption process is finished, a notification message is displayed on the user's device. This message typically contains a ransom demand from the attacker, which outlines the conditions for the release of the encrypted files.
- Ultimately, as part of the encryption process, the ransomware will often change the file extensions of the encrypted files to specific ones like ".crypt", ".cryptolocker", ".cryptowall", and ".encrypted" and to make it easier for the attacker to identify them for future reference.

To summarize, ransomware is a type of malicious software that can infect a computer system and encrypt files using complex mathematical algorithms. The ransomware may migrate and hide in the "explorer.exe" extension, changing it to a new, infected extension that it copies to paths like "% Appdata%" and "% Programdata%", and modify the registry value in "HKEY_LOCAL_MACHINE\Software\Microsoft\Windows\Currentversion\Run". The ransomware can also disable useful services using extensions like "vssadmin.exe" and "bcdedit.exe" to prevent the system from reacting and delete sensitive data to force the target to pay the ransom. After encryption, a public key is applied using one of the asymmetric encryption algorithms. If the system is rebooted, the ransomware will run continuously and keep the connection between the command and control server at all times.

3. Case study - WannaCry

This scientific research describes a method for carrying out a cyber-attack and assesses the risks that cyber systems face through the SEAP platform. The SEAP platform is used to publish procedures related to awarding public procurement contracts/framework agreements or advertising notices. As a result, it represents a potential target for cyber-attack and, in this case study also highlights the vulnerabilities of cyber systems that can be exploited by attackers.

To access the procedure, users can navigate to the related initiation notice and select the "View procedure" button. This will redirect the user to a list of initiation notices of the corresponding type, from which the user can choose the desired initiation notice. Once registered in the procedure, the economic operator can proceed to submit their offer by accessing the "My offer" section of the procedure viewing screen. The application process involves uploading the necessary files in the "My Offer" section, specifically in the Qualification Documents subsection.

The contracting authority will download and review the documents that have been submitted during the award procedure/advertisement in order to assess the potential bidders and ultimately determine the winning bid.

Our study aims to examine the process of transmitting the offer, which is submitted in a PDF format. We will focus on a particular type of worm that is injected into the offer, and which is designed to execute a cyber-attack of the ransomware variety. The worm's primary objective is to encrypt the contracting authority's data, rendering it inaccessible and held for ransom. Through our investigation, we hope to gain a better understanding of the potential vulnerabilities within the offer submission process and to develop effective countermeasures against such attacks in the future.

The contracting authority is likely to face significant challenges in protecting against a ransomware attack similar to WannaCry. Such attacks have been observed to spread rapidly through the use of a worm component, making it difficult to contain and mitigate the damage caused. Furthermore, the encryption component of WannaCry employs public key cryptography, which is a robust and widely-used encryption method that presents significant challenges for decryption without the proper key, so it is crucial for the contracting authority to develop and implement comprehensive cybersecurity measures to minimize the risk of such attacks and ensure the integrity of their systems and data.

Moving forward, we will analyze two distinct executable components: the worm and encryption components. Table 2, 3 and 4 [6] provides a detailed breakdown of these components, including their respective functions and properties.

Table 2. Executable components: the worm and encryption components

	Worm component
MD5	db349b97c37d22f5ea1d1841e3c89eb4
SHA 1	e889544aff85ffaf8b0d0da705105dee7c97fe26

	Worm component
SHA 256	24d004a104d4d54034dbcffc2a4b19a11f39008a575aa614ea04703480b1022c
File type	PE32 executable (GUI) Intel 80386, for MS Windows
	Encryption component
MD5	84c82835a5d21bbcf75a61706d8ab549
SHA 1	5ff465afaabcbf0150d1a3ab2c2e74f3a4426467
SHA 256	ed01ebfbc9eb5bbea545af4d01bf5f1071661840480439c6e5babe8e080e41aa
File type	PE32 executable (GUI) Intel 80386, for MS Windows

The Pestudio analysis conducted on WannaCry's executable components yielded the following results regarding its worm and encryption functionalities:

Table 3. The Dynamic Link Libraries (DLLs) that comprise the structure of a worm

Library	**Imports**	**Description**
ws2_32.dll	13	Windows Socket 2.0 32-bit DLL
iphlpapi.dll	2	IP Helper API
wininet.dll	3	Internet Extensions for Win32
kernel32.dll	32	Windows NT Base API Client DLL
advapi32.dll	11	Advance Windows 32 Base API
msvcp60.dll	2	Windows NT C++ Runtime Library DLL
msvcrt.dll	28	Windows NT CRT DLL

Table 4. The Dynamic Link Libraries (DLLs) that constitute the encryption component

Library	**Imports**	**Description**
kernel32.dll	54	Windows NT Base API Client DLL
advapi32.dll	10	Advance Windows 32 Base API
user32.dll	1	Multi-User Windows User API Client DLL
msvcrt.dll	49	Windows NT CRT DLL

During its execution, the worm framework DLLs call upon the iphlpapi.dll extension to obtain the network configuration settings of the infected host. On the other hand, the encryption component heavily relies on the kernel32.dll and msvcrt.dll libraries, which are among the most frequently invoked DLLs. This suggests that these two malicious libraries are responsible for the primary encryption functionality implemented in the component. To verify this assertion, a closer examination of the imported functions of these libraries will be carried out.

Table 5. The functions that facilitate the encryption format [6]

Function	**Location**
GetCurrentThread	0x53a
GetStartupInfoA	0xa97a
StrartServiceCtrDispatcherA	0xa6f6
RegisterServiceCtrDispatcherA	0xa6d8
CreateServiceA	0xa688
StartServiceA	0xa662
CryptGenRandom	0xa650
CryptAcquireContextA	0xa638
OpenServiceA	0xa714
GetAdaptersInfo	0xa792
InternetOpenUrlA	0xa7c8
OpenMutexA	0xda84
GetComputerNameW	0xd8b2
CreateServiceA	0xdc2a
OpenServiceA	0xdc62
StartServiceA	0xdc52
CryptReleaseContext	0xdc14
RegCreateKeyW	0xdc04

Function	Location
fopen	0xdcd4
fread	0xdccc
fwrite	0xdcc2
fclose	0xdcb8
CreateFileA	0xd922
ReadFile	0xd964

Table 5 displays the list of the most suspicious functions identified among them. Overall, the analysis indicates that WannaCry predominantly employs Microsoft's Crypto, File Management, and C Runtime APIs for its operations. The crypto API library is specifically used to generate and manage both symmetric and asymmetric cryptographic keys, which play a vital role in ensuring secure transmission and storage of sensitive information.

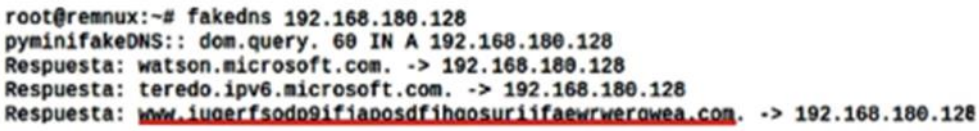


Fig. 1. The process of capturing malicious DNS requests through FakeDNS [6]

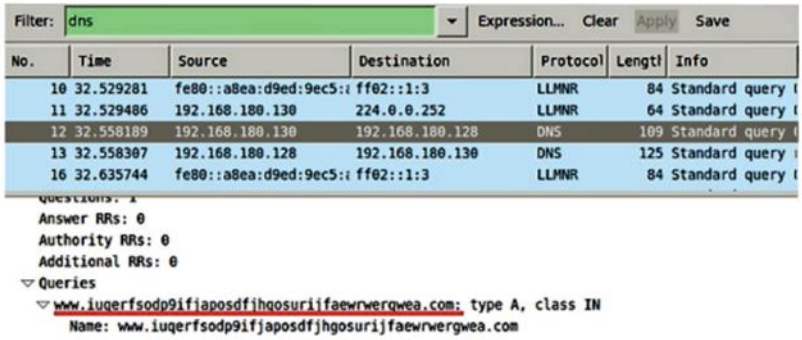


Fig. 2. The malicious DNS request was captured using Wireshark [6]

Throughout the live analysis, it was observed that upon initialization, the worm component makes an attempt to establish a connection with a specific domain. This connection is initiated using the InternetOpenUrl function, which is commonly used for accessing resources over the internet. The domain in question is identified as “www.iuqerfsodp9ifjaposdfjhgosurijfaewrwergwea.com” [6], and is suspected to be associated with the spread of the malware.

The examination carried out the worm component via runtime execution revealed that upon initialization, the component attempts to establish communication with a specific domain utilizing the InternetOpenUrl function. The mentioned domain serves as a kill-switch domain, meaning that if the domain is active, the worm component terminates its execution. Conversely, if the worm component fails to establish a connection with this domain, it continues to run and installs itself as the "Microsoft Security Center Service (2.0)" process mssecsvs2.0 on the compromised system [6]. Therefore, the existence of this kill-switch domain can be employed as a detection technique when developing a defense mechanism against WannaCry.

The malicious DNS request on port 80 was captured by the FakeDNS utility of REMnux, as illustrated in Fig. 1. In addition, Fig. 2 depicts the query field of DNS packets sent from the infected machine (IP 192.168.180.130) to the DNS server on REMnux (IP 192.168.180.128), as observed through Wireshark [6].

Upon failure to connect to the kill-switch domain, the WannaCry worm component initiates the creation of an mssecsvs2.0 process with DisplayName of "Microsoft Security Center (2.0) Service".

This event is visible in the Process Hacker tool, which displays the process with a PID of 4016 (Fig. 3). Furthermore, the malware extracts the hardcoded R resource binary, which represents the WannaCry encryption component binary, and copies it to the "C:\Windows\taskche.exe" directory path. The worm subsequently launches the executable with the command line parameters "C:\Windows\taskche.exe/i". Additionally, it attempts to move the file "C:\Windows\taskche.exe" from "C:\Windows\qeriuwjhrf" to replace the original file, if it exists. This is executed to enable multiple infections and to circumvent any issues related to tasksche.exe process creation [6].

Fig. 3. Microsoft Security Center (2.0) Service [6]

At the end of its execution, WannaCry creates a persistent mechanism in the Windows registry to ensure that the malware runs automatically after every system restart. This involves generating a unique string, such as "midtxzggq900", using the computer name and storing it as a new entry in the registry. After this step, the malware copies itself to a randomly named folder within the Common Appdata directory on the infected computer. Finally, WannaCry attempts to establish memory persistence by adding itself to the AutoRun feature, thus ensuring that it executes every time the system boots up [6].

```
Created         C:\ProgramData\midtxzggq900\b.wnry
Modifed 15F936  C:\ProgramData\midtxzggq900\b.wnry
Created         C:\ProgramData\midtxzggq900\c.wnry
Modifed 30C     C:\ProgramData\midtxzggq900\c.wnry
Created         C:\ProgramData\midtxzggq900\msg
Created         C:\ProgramData\midtxzggq900\msg\m_bulgarian.wnry
Modifed         C:\ProgramData\midtxzggq900\msg
Modifed BB07    C:\ProgramData\midtxzggq900\msg\m_bulgarian.wnry
Created         C:\ProgramData\midtxzggq900\msg\m_chinese (simplified).wnry
Modifed D457    C:\ProgramData\midtxzggq900\msg\m_chinese (simplified).wnry
Created         C:\ProgramData\midtxzggq900\msg\m_chinese (traditional).wnry
Modifed 135F2   C:\ProgramData\midtxzggq900\msg\m_chinese (traditional).wnry
```

Fig. 4. WannaCry is known to launch files in the working directory [6]

Fig. 5. Ransom [6]

The comprehensive dynamic analysis of the WannaCry ransomware was performed in a specially designed virtual testbed [6].

4. Conclusion

The study revealed that WannaCry is composed of two distinct components that work together to enable the worm-like self-propagation mechanism and the combined encryption process. The

worm component is responsible for spreading the ransomware to other vulnerable systems on the network, while the encryption component is responsible for encrypting the files on the infected machine.

The evaluation implemented on the WannaCry ransomware provided an in-depth understanding of the technical aspects of the malware, exposing its intricate structure and intricate functionality. This analysis demonstrated the ability of the malware to propagate quickly and effectively, infecting multiple systems and causing significant damage. Moreover, it highlighted the importance of developing and implementing robust defense mechanisms to prevent the spread of similar threats in the future. Overall, the results of this analysis provided valuable insights into the workings of WannaCry, emphasizing the need for continued research and development in the field of cybersecurity to mitigate the impact of such malicious attacks.

The SEAP platform was chosen as the focus of this study due to its potential vulnerability to a WannaCry attack. The analysis aimed to investigate the specific aspects of the attack, including the process of infection, its mechanism for persistence, encryption, and prevention of recovery. Additionally, the study examined the methods by which the ransomware propagated and communicated with its command and control servers. The findings of the analysis were essential in identifying the characteristics and behaviors of the WannaCry ransomware, which have implications for the development of effective defense mechanisms and incident response strategies.

References

[1]. www.e-licitatie.ro.

[2]. A. Palisse et al., “Ransomware and the Legacy Crypto API”, The 11th International Conference on Risks and Security of Internet and Systems. 5th-7th September 2016 (Roscoff, France: Springer).

[3]. G. Hull, H. John, and B. Arief, “Ransomware deployment methods and analysis: views from a predictive model and human responses”, Crime Science, vol. 8, no. 1, 2019; K. Savage, P. Coogan, and H. Lau, The evolution of ransomware. Symantec, 2015.

[4]. BleepingComputer (2016). Locky Ransomware Information, Help Guide, and FAQ, https://www.bleepingcomputer.com/virus-removal/lockyransomware-information-help.

[5]. L. Abrams (2014). CryptoDefense and How_Decrypt Ransomware Information Guide and FAQ, https://www.bleepingcomputer.com/virusremoval/cryptodefense-ransomware -information; Webroot (2017). MSP Guide: Stopping Crypto Ransomware Infections in SMBs, 16 Easy Actions for MSPs, White Paper.

[6]. WannaCry Ransomware: Analysis of Infection, Persistence, Recovery Prevention and Propagation Mechanisms Maxat Akbanov, Vassilios G. Vassilakis, and Michael D. Logothetis.

[7]. D. O’Brien, “Ransomware 2017”, Internet Security Threat Report, Symantec, July 2017 [Online]. Available: https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/istr-ransomware-2017-en.pdf.

[8]. K. Savage, P. Coogan, and H. Lau, “The evolution of ransomware”, Security Response, Symantec, June 2015 [Online]. Available: http://www.symantec.com/content/en/us/ enterprise/media/securityresponse/whitepapers/the-evolution-of-ransomware.pdf.

[9]. Pestudio, Malware Assessment Tool [Online]. Available: https://www.winitor.com.

[10]. K. Cabaj and W. Mazurczyk, “Using software-defined networking for ransomware mitigation: The case of CryptoWall”, IEEE Network, vol. 30, no. 6, pp. 14–20, 2016.

A Method of Warning About Unauthorized Access to a Room

Cristian-Ovidiu OPRIȘ
Faculty of Electronics, Telecommunications and Information Technology,
University POLITEHNICA of Bucharest, Romania
cristian.opris@upb.ro

Abstract

This paper is based on the study of cybercrime in the context of a world based on technology. Whether it is financial losses, data leaks or mental trauma resulting from harassment in the online environment, cybercrime is part of the reality of the modern world, where the multiple advantages of using the most advanced technologies bring with them disadvantages that cannot be ignored. We will treat the types of cyberattacks, but also the methods by which we can protect ourselves as much as possible. An example of increasing the degree of security in terms of physical access to a room containing sensitive information, achieved at low cost, is also provided. A "smart" entrance mat is used to provide access, a coconut fiber mat into which Linqstat (Velostat) tactile force sensors and the data processing electronics provided by them have been inserted.

Index terms: access security, cyberattacks, tactile sensor, Velostat

1. Introduction

Cybercrime is increasingly common, and current technical methods of combating cybercrime are often ineffective. Therefore, preventive strategies become necessary to reduce cybercrime. The following two aspects are important: the characteristics of criminals to understand the motivations behind the crime, but also the characteristics of users' computer systems to better understand how they are victims of cybercrime.

Cybercriminals are developing increasingly intelligent techniques for their victims in the online environment: individuals, companies or organizations. Cybercrimes are on the rise due to lack of cyber security. All types of computer crimes consist of both the computer and the person behind it as victims. A generalized definition of cybercrime may be “Unlawful acts wherein the computer is both a tool and target”. Cybercriminal is a person who commits an illegal act with a guilty intention or commits a crime in context of cybercrime [1]. Cybercriminal can also be the person who illegally enters a secure room containing classified data in digital form to steal it. In this context, a method for detecting the number and even the approximate weight of people who can illegally access a secure space is presented.

Cyber security is a term of security which is implicated through diversified disciplines, most of them focusing on technical or psychological problems such as computer science, criminology, economics, engineering, information systems, management, medicare, neurophysiology, psychology, sociology, etc. It affords the people with discussions about behaviors and motivations, benefits and consequences about cybercrime and security [2]. The scope of cyber security is not just limited to securing the information in IT industry but also to various other fields like cyber space etc. The latest

technologies like cloud computing, green computing, mobile computing, e-commerce, net banking are required high level of information security [2].

Measures aimed at mitigating cyber-attacks should be implemented through the replication of already existing approaches and methods established to curb conventional crimes in society. These measures should be in the form of government policies, enactment of legislative laws, education and awareness and cooperation between government agencies, private sectors, the public and relevant international bodies. Special committee of experts and stakeholders should be instituted to oversee the modification of conventional strategies and approaches in order to fine tune implemented policies to be cybercrime specific [3]. Some measures can be:

- *Reduce Opportunities*: elaborate system design so that hackers do not hack the computer.
- *Use Authentication Technology*: use password bio-metric devices, fingerprint or voice recognition technology and retinal imaging, greatly immense the difficulty of obtaining unauthorized access to information systems.
- Data Recovery: develop tools for data recovery and analysis.
- Reporting: always report the crime to cyber fraud complaint center in one's country as they maintain huge data and have better tools for controlling cybercrime.
- Install firewalls: as they block particular network traffic according to security policy.
- Attachments: Avoid opening attachments or e-mails which were not expected and have come from an unknown source or person [4].

2. Types of cybercrimes

Among the reasons behind the actions of cybercrime are fighting for a cause in which the criminal believes, financial gain, public recognition. Types of cyberattacks will be classified as follows (Table 1).

Table 1. Types of cybercrimes [1]

Cybercrime against individuals	**Cybercrime on property**	**Cybercrime within organisations**
E-mail Spoofing	Intellectual Property crime	Hacking
Phishing	Cyber squatting	Password
Spamming	Cyber vandalism	Denial attack
Cyber defamation	Hacking system	Virus attack
Cyber stalking	Alerting way of unauthorized	Mail bomb
Salami attack	Logic bomb	
Computer sabotage	Trojan horse	
Malware		

2.1. Cybercrime against individuals

E-mail is one of the most used applications for communication. Security from this point of view is defined as the ability to provide confidentiality. E-mail date and E-mail address spoofing are the two important forms of E-mail spoofing. E-mail date spoofing occurs when someone changes the sending date and e-mail address spoofing refers to sending mail which pretends to come from someone else [5]. The date and time of the email are particularly important in such cases bank statements, communications from the chief / chief in terms of terms, etc. Simple Mail Transfer Protocol (SMTP) is the largest and most important protocol for e-mail. Unfortunately, this does not include security policies.

Phishing uses a technique in which the attackers are trying to trick victims. These attacks start with an email from an attacker pretending be someone or something you know or trust (bank,

knowledge, shop). In the content of the email appears a redirecting to a link or a file attached, the purpose is to show everything as convincing as possible to get money from the victims.

With spam emails being around for more than 40 years, cybercriminals have always found new ways to us them to catch victims out, having gauged "click rates rising from 13.4% in the second half of 2017 to 14.2% in 2018," according to Adam Sheehan from MWR InfoSecurity [6].

Defamation is a false statement that harms the reputation of individual person, business, product, group, government, religion of nation. For a statement to constitute defamation a claim must generally be false and made to someone other than the person defamed, and result injures the reputation of a person who is defamed [7].

Malware, short for malicious software, consists of programming (code, scripts, active content, and other software) designed to disrupt or deny operation, gather information that leads to loss of privacy or exploitation, gain unauthorized access to system resources, and other abusive behavior. The expression is a general term used by computer professionals to mean a variety of forms of hostile, intrusive, or annoying software or program code. Malware is not the same as defective software - software that has a legitimate purpose but contains harmful bugs (programming errors) [8].

2.2. Cybercrime within organisations

In 1997, the Group of Eight (G8) established a "Subcommittee986 on High-tech Crimes", dealing with the fight against cybercrime. During their meeting in Washington DC, United States, the G8 Justice and Home Affairs Ministers adopted ten Principles and a Ten-Point Action Plan to fight high-tech crimes.988 The Heads of the G8 subsequently endorsed these principles, which include [9]:

- There must be no safe havens for those who abuse information technologies.
- Investigation and prosecution of international high-tech crimes must be coordinated among all concerned states, regardless of where harm has occurred.
- Law-enforcement personnel must be trained and equipped to address high-tech crimes.

The risks associated with weak protection measures could in fact affect developing countries more intensely, due to their less strict safeguards and protection. The ability to protect customers, as well as firms, is a fundamental requirement not only for regular businesses, but also for online or Internet-based businesses. In the absence of Internet security, developing countries could encounter significant difficulties promoting e-business and participating in online service industries [9].

Several EU legislative actions contribute to the fight against cybercrime. These include:

- 2013 - A Directive on attacks against information systems which aims to tackle large-scale cyber-attacks by requiring Member States to strengthen national cyber-crime laws and introduce tougher criminal sanctions. In 2017, the Commission has published a Report assessing the extent to which Member States have taken the necessary measures in order to comply with the Directive.
- 2011 - A Directive on combating the sexual exploitation of children online and child pornography, which better addresses new developments in the online environment, such as grooming (offenders posing as children to lure minors for the purpose of sexual abuse).
- 2002 - ePrivacy Directive whereby providers of electronic communications services must ensure the security of their services and maintain the confidentiality of client information. In 2017, the Commission proposed to repeal the Directive and replace it with a Regulation concerning the respect for private life and the protection of personal data in electronic communications.
- 2001 - Framework Decision on combating fraud and counterfeiting of non-cash means of payment, which defines the fraudulent behaviors that EU States need to consider as punishable criminal offences. On 13 September 2017, the Commission has proposed a new Directive aiming at updating the current legal framework, removing obstacles to

operational cooperation and enhancing prevention and victims' assistance, to make law enforcement action against fraud and counterfeiting of non-cash means of payment more effective [10].

According to Check Point (Figure 1), in 2021 education and research were the sectors with the highest volume of attacks, up 75% compared to 2020. On the next positions, "the government/military sector had 1,136 attacks per week (47% increase), and the communications industry had 1,079 attacks weekly per organization (51% increase)" [11].

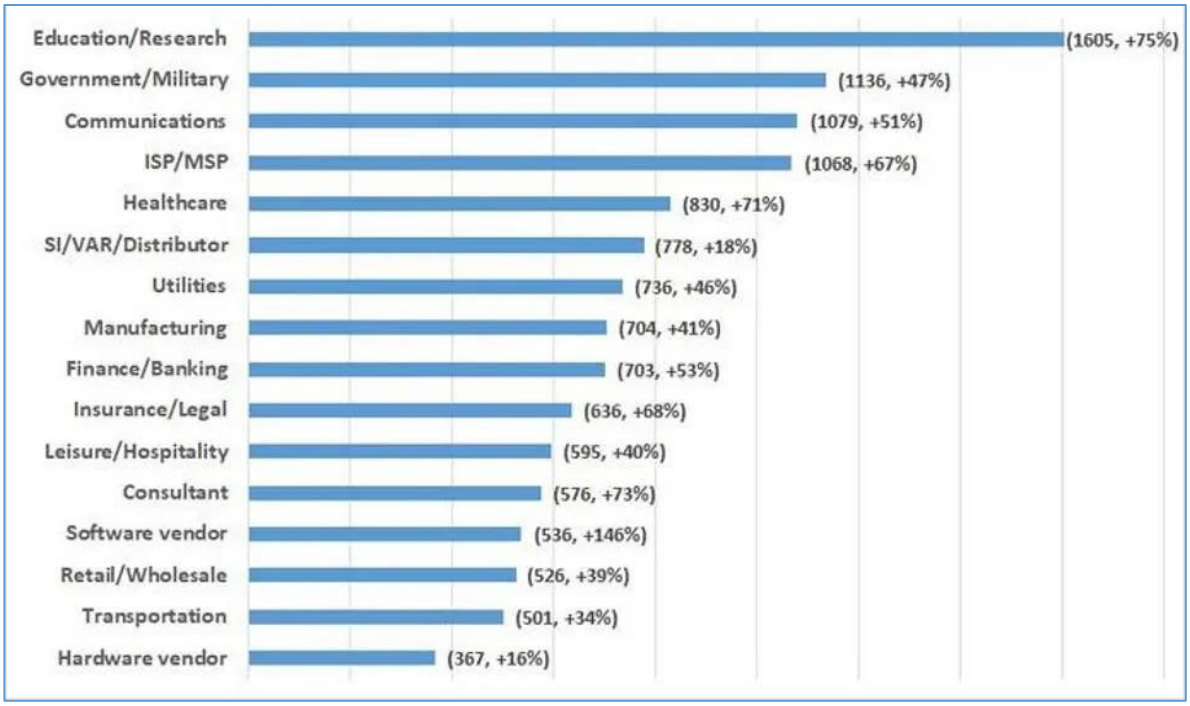


Fig. 1. Average weekly attacks per organisation by industry [11]

Possible reasons for these differences may be the types and frequencies of attacks experienced as well as the importance that each company places on the theft of information assets versus other consequences of the incident [12]. The following figure shows the types of cyberattacks for 60 companies.

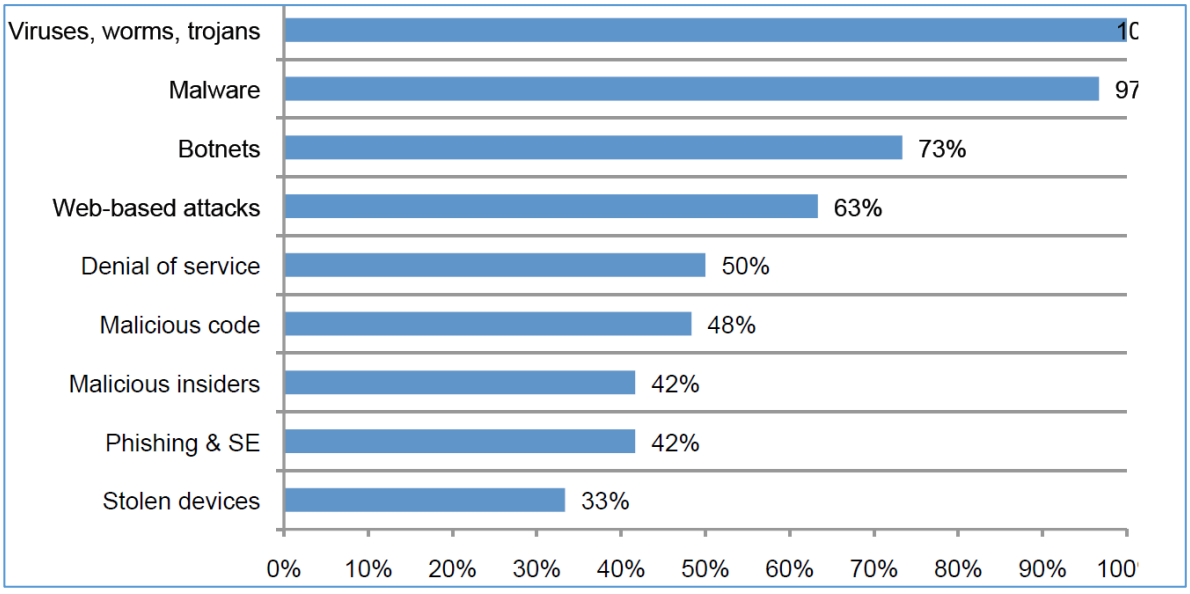


Fig. 2. Types of cyberattacks for 60 different companies [12]

3. The method of warning of unauthorized access to a room containing sensitive data

To obtain data about the approximate weight and number of people accessing a secure room, it was decided to use eight tactile force sensors [13] mounted in an entrance mat. The sensors will be placed so that the surface of the mat is distributed equally for each sensor (Figure 3).

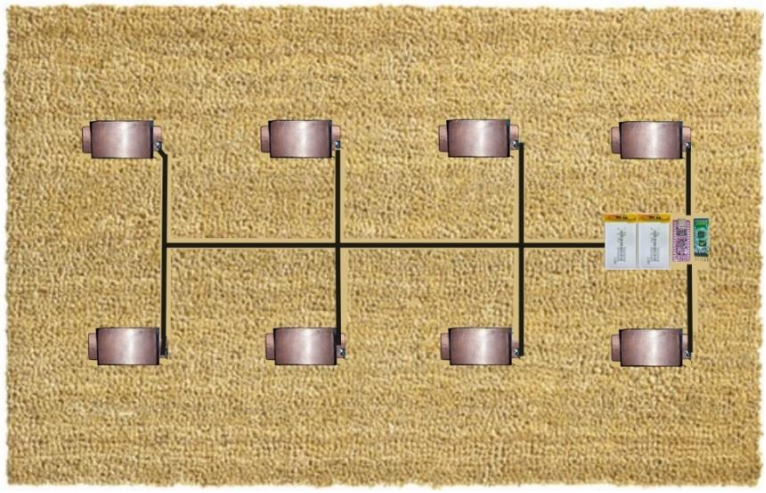


Fig. 3. "Smart" entrance mat

Two 40x60 cm entrance mats were used to make the "smart" entrance mat. In the first mat, eight cutouts were made in which eight tactile force sensors were placed, and on one side of the mat a cutout was made in which the electronic circuits were mounted. After commissioning, the entire surface is covered with a durable cloth for protection. The second mat is glued back-to-back with the first. This is the one that will be positioned normally, face up at the entrance to the secure room.

The block diagram of the electronic assembly is shown in Figure 4. Sensors are indicated by their resistance values (Rs1 to Rs8). The Velostat material is characterized by the phenomenon of piezo resistivity, i.e., it changes its electrical resistance because of its deformation [13].

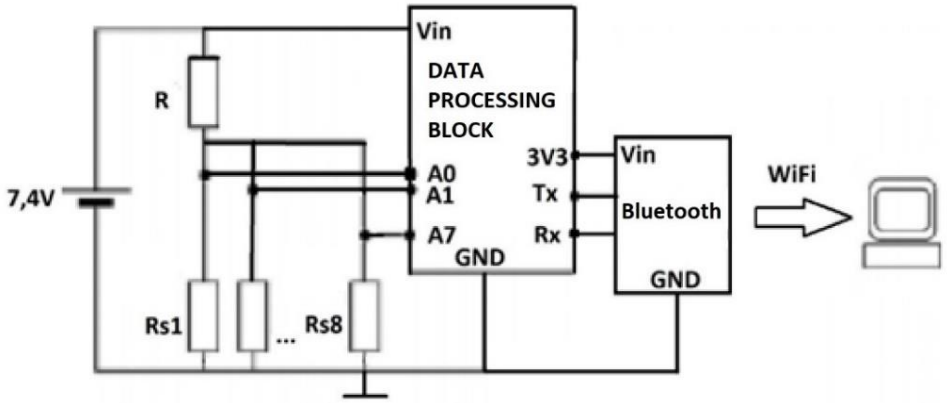


Fig. 4. The block diagram of the electronic assembly

The value of the resistor R is chosen so that we have optimal voltages at the input of the data processing block. The data processing block contains signal amplifiers and the Arduino Nano Board which contains 8 analog inputs. Warning decisions made by programming the Arduino board are transmitted via the Bluetooth Module to a computer or mobile phone via the Bluetooth Terminal application. The Arduino Nano Board and Bluetooth Module were chosen due to their small size and low power consumption. The entire electronic assembly is powered by 7.4 V lithium-polymer batteries.

4. Conclusion

In conclusion, the foundation of criminal justice system is to keep order and secure justice in the society. Safety is one of the basic needs of people according to Maslow's pyramid of hierarchical needs and crime prevention [14], in this respect, emerges as a necessity for development of a healthy and safe society. Therefore, the state has to be one step ahead the criminals in terms of the use of technology or the control/surveillance of the technology in order to prevent unlawful actions in cyberspace. As the uncontrolled power is not (a true and just) power, similarly, the uncontrolled technology is not (a helpful and good) technology for citizens. Without safety, which is one of their basic needs, the citizens will not feel like they are living in a free and happy world [15].

This paper presented a way to warn of a break-in (by forcibly removing a locking system) in a room containing sensitive data. The mode is also valuable if the classic alarm device is disabled. The Internet is a powerful tool and effective means of communication, but it is vulnerable just like anything else. To defend against cybercrimes, intrusion detection techniques should be designed, implemented, and administrated.

References

[1]. E. N. Kaur, "Introduction of Cyber Crime and its Types," International Research Journal of computer Science (IRJCS), 2018.
[2]. V. Kavitha, "Cyber Security issues and challenges - a review," International Journal of Computer Science and Mobile Computing, vol. 8, no. 11, 2019.
[3]. M.B. Owiso, Cyber Crime, https://www.academia.edu/12741661/Cyber_Crime.
[4]. M. Lakshmi Prasanthi, Tata A.S.K. Ishwarya, "Cyber Crime: Prevention & Detection," International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 3, March 2015.
[5]. E.S. Pilli, "Forensic analysis of e-mail address spoofing," in 2012 Third International Conference on Computer and Communication Technology, 2014.
[6]. Spam still first choice for cyber crime, according to study. https://www.information-age.com/spam-still-first-choice-cyber-crime-according-study-123473840/#.
[7]. A.L. Ishabakaki, "Defamation in social media (cyber defamation) legal perspective in Tanzania," Victory Attorneys & Consultants.
[8]. http://cs.sru.edu/~mullins/cpsc100book/module05_SoftwareAndAdmin/module05-04_softwareAndAdmin.html. [Online].
[9]. I.T.D. Sector, "Understanding cybercrime: phenomena, challenges and legal response," Sept. 2012.
[10]. Cybercrime. European Commission. https://ec.europa.eu/home-affairs/what-we-do/policies/cybercrime_en.
[11]. Check Point, Check Point Research: Cyber Attacks Increased 50% Year over Year, 2022. https://blog.checkpoint.com/security/check-point-research-cyber-attacks-increased-50-year-over-year/.
[12]. P. Institute, "Cost of Cyber Crime Study: United States," HP Enterprise Security, 2013.
[13]. C.O. Opris, I.B. Bacis, L. Milea, A. Vasile, "Implementation of a Resistive Pressure Sensor Made With "Linqstat" for Automotive". ISSE 2023, Timișoara, Romania.
[14]. S. Mcleod, Maslow's Hierarchy of Needs, Simply Psychology, http://www.simply psychology.org/maslow.html. Accessed Apr. 20, 2023.
[15]. Z. Gul, R. Terkesli, "Crime of the Millennium: Cyber Crime," Humanity & Social Sciences Journal 7, 2012.

Guarding the Nation: A Comprehensive Look at State Cybersecurity Measure

Marian-Emilian SPĂTARU, Alexandru BARCAN
"Alexandru Ioan Cuza" Police Academy, Bucharest, Romania
marian.spataru@outlook.com, alexbarcan23@gmail.com

Abstract

In a continuously evolving world, technology has not been left out of the process which consists of studies and research done by specialists in the field of cyber technology. Although the latter has brought along benignant effects in society, it can be considered a controversial domain due to those effects that can be used against the public safety and national security. Cyber-attacks & Cyber terrorism are just two of them, usually countered by Cyber intelligence, OSINT security, Cyber risk management. These actions are coordinated by different intelligence services such as: Federal Bureau of Investigation – FBI, Romanian Intelligence Service – SRI, Federal Security Service – FSB, while they have to cooperate with civilians, due to a shortage of employees. The lack of qualified staff on the following domain: awareness of the different types of cyber-attack, such as malware, web-based attacks, phishing, web application attacks, spam, distributed denial of service (DDoS), identity theft, data breach, insider threat, botnets, physical manipulation, damage, theft and loss, information leakage, ransomware, cyber-espionage, industrial espionage and crypto jacking, reaches an amount of 7.659 officials that are needed in this area.

Index terms: cyber intelligence, OSINT security, cyber-attacks, cyber terrorism, cyber risk management

1. Introduction

State cybersecurity weaknesses refer to the vulnerabilities and the gaps that exist in the cybersecurity defenses of a state. These vulnerabilities and gaps can arise due to several reasons, including inadequate funding, lack of skilled personnel, outdated technology and also poor security protocols.

The major challenge in state cybersecurity is the lack of adequate funding. Many states have limited budgets for cybersecurity, which makes it challenging to implement and maintain robust security measures. As a result, states may have outdated software and hardware systems that are vulnerable to cyber-attacks. The demand of cybersecurity experts has increased significantly in recent years, but the supply has not kept up with the demand. This shortage of skilled personnel can lead to inadequate cybersecurity measures and poor incident response.

So, in order to prevent state cybersecurity weaknesses, governments must allocate sufficient resources to fund and maintain robust cybersecurity measures. All these measures consist of investing in technology, hiring skilled personnel, and implementing best practices for security protocols.

This research aims to provide an overview on cybersecurity of states, how it works, the stages of cyber protection against cyber-attacks, cyber terrorism, and also cyber risk management which also **categorize the levels of risk** [2].

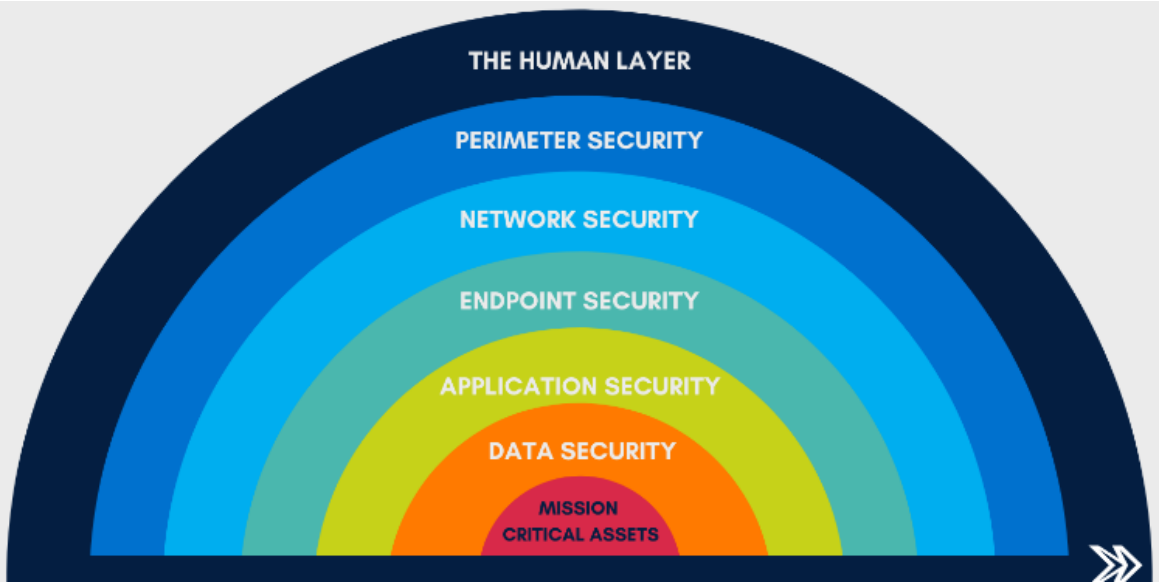


Fig. 1. The 7 layers of cybersecurity (Source: www.diamondit.pro/7-layers-of-cybersecurity/)

2. The Invisible Threat: APT28's Sophisticated Tactics for Hacking and Undermining State Cybersecurity

The hacking groups are divided into two categories:

- **"White Hat"** - In general, "white hat" hackers are individuals who use their technical skills to identify vulnerabilities in computer systems, networks, or applications, with the goal of improving their security.
- **"Black Hat"** - hackers are individuals who use their technical skills to exploit vulnerabilities in computer systems, networks, or applications, for personal gain or to cause harm to others. They are often involved in illegal activities such as stealing data, installing malware, or conducting distributed denial-of-service (DDoS) attacks. Their actions are considered illegal and unethical because they do not have the permission or authorization to perform their activities.

APT28, also known as **Fancy Bear**, is a cyber espionage group that is believed to be based in Russia. The group has been active since at least 2007 and has been responsible for a number of high-profile cyber-attacks around the world. This group is known for using sophisticated and highly targeted techniques to gain access to sensitive information. They have been linked to several attacks on government agencies, military organizations, and political groups in the United States and Europe [11], [14], [15].

One of the group's most well-known attacks was the hack of the DEMOCRATIC NATIONAL COMMITTEE (DNC) during the 2016 US presidential election. The group is believed to have stolen sensitive emails and other data from the DNC and released it to the public in an effort to influence the election. APT28 has also been linked to a number of other attacks, including the hack of the German parliament, the French presidential campaign, and the World Anti-Doping Agency [4].

The group has been known to use a variety of techniques, including spear-phishing malware, and social engineering, to gain access to its targets. The group's motivations are believed to be primarily political in nature, and it is believed to be sponsored by the Russian government. APT28 is considered to be one of the most sophisticated and dangerous cyber espionage groups in the world, and its activities have raised serious concerns about the sensitive information and infrastructure around the world.

The variety of techniques that they use to attack political groups in the United States includes spear-phishing, malware, and social engineering. The techniques are a variety of malware strains,

including X-AGENT and SEDNIT, to gain access to systems and steal data. Once installed on a system, the malware can be used to remotely control the system, exfiltrate data or even carry out other malicious activities. [11], [14], [15].

In the case of the hack of the DEMOCRATIC NATIONAL COMMITTEE (DNC) during the 2016 US presidential election, APT28 used a combination of these techniques to gain access to the organization's systems. The group sent spear-phishing emails to DNC staff members, which contained a link to a fake login page. When the staff members entered their login credentials, APT28 was able to steal their usernames and passwords, giving the group access to the DNC's system.

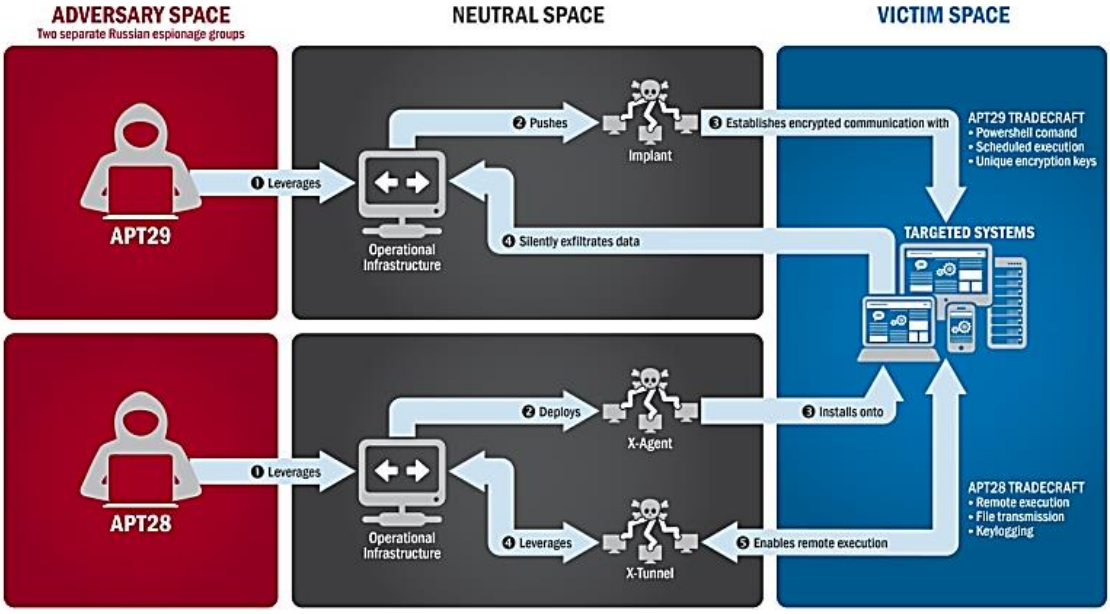


Fig. 2. The tactics and techniques used by APT29 and APT28 to conduct cyber intrusions against target systems

X-Agent and SEDNIT are two malware tools believed to be developed and used by the Russian hacking group known as APT28 or Fancy Bear (diagram taken from [6]).

X-Agent is a remote access Trojan (RAT) designed to give hackers remote access and control over a targeted system. It is capable of performing various malicious activities such as stealing sensitive data, capturing screenshots, recording keystrokes, and executing commands. X-Agent has been used in several high-profile cyber espionage campaigns, including the 2016 Democratic National Committee (DNC) hack [4], [11], [14], [15].

SEDNIT is another malware tool used by APT28. It is a modular malware that consists of multiple components, including a downloader, a backdoor, and a rootkit. SEDNIT is known for its stealthy and sophisticated techniques, such as using anti-debugging and anti-VM techniques to evade detection. Like X-Agent, SEDNIT has been used in various cyber espionage campaigns, targeting government and military organizations, as well as critical infrastructure sectors.

Both X-Agent and SEDNIT are considered highly sophisticated malware tools and are often used by APT28 in targeted attacks aimed at stealing sensitive information and disrupting critical infrastructure [5], [6].

The scientists' research revealed that APT28 utilizes a modular framework called backdoor CHOPSTICK as a means of developing a backdoor. The ironic name is due to its flexibility in compiling variants with different capabilities and the ability to deploy additional capabilities during runtime. With this modular design, the developers can create targeted implants, including only the necessary capabilities and protocols required for a specific environment [5], [6].

CHOPSTICK uses these methods in order to move messages and information:

1. Communications with a C2 server using HTTP.

2. Email sent through a specified email server. One CHOPSTICK v1 variant contained modules and functions for collecting keystroke logs, Microsoft Office documents, and PGP files.

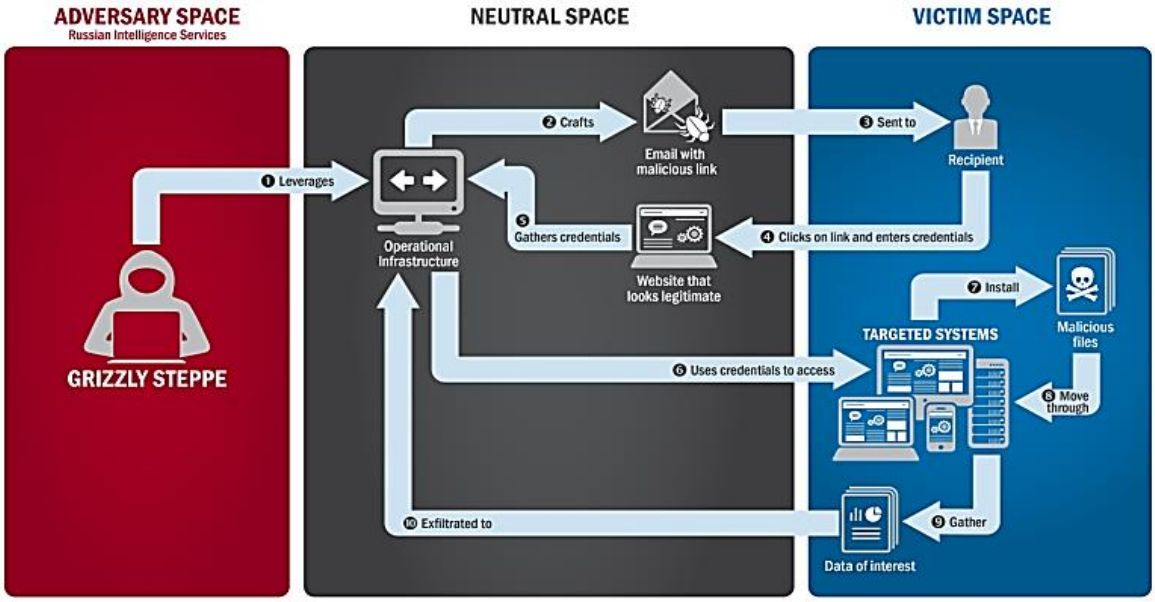


Fig. 3. APT28's use of spear phishing and stolen credentials

Also, during the research, the scientists discovered that APT28 used in the coding two details consistent across the malware samples. The first one was that APT28 had consistently compiled into their malware Russian language settings, which made it harder to read the code sometimes. Also, the second one was the fact that the malware compiled times from 2007 to 2014, and it corresponded with Moscow and St. Petersburg business hours. (Diagram taken from [6])

In conclusion, APT28, also known as Fancy Bear, is a cyber espionage group believed to be based in Russia that has been active since at least 2007. The group is known for using sophisticated and highly targeted techniques, including spear-phishing, malware, and social engineering, to gain access to sensitive information. APT28 has been linked to a number of high-profile cyber-attacks around the world, including the hacking of the Democratic National Committee during the 2016 US presidential election. The group is believed to be politically motivated and sponsored by the Russian government. APT28 utilizes a variety of malware tools, including X-Agent and SEDNIT, to remotely control systems and exfiltrate data. The group also uses a modular framework called backdoor CHOPSTICK to develop a backdoor and communicate with a command-and-control server using HTTP and email. The use of Russian language settings and compile times corresponding to Moscow and St. Petersburg business hours in their malware code indicates APT28's origin and location. APT28 is considered to be one of the most sophisticated and dangerous cyber espionage groups in the world, and its activities have raised serious concerns about the security of sensitive information and infrastructure worldwide [3], [5], [6].

3. Fortifying The Nation: Robust Cybersecurity Protocols for Safeguarding State Security

The idea of fortifying the nation goes above the common sense to fortify something physically. In our field we choose to use," fortify" to refer to the whole action/process which consists of recruiting people to work in the cybersecurity field, training them and gaining advantage in front of other

organisation/states. In this way, we know that the new cybersecurity strategy says that no product is totally secured and cannot be totally secured at least in the near future.

To improve this, new strategies urge the government to consider taking on some responsibility for so-called cybersecurity insurance. It is understandable how many companies and government agencies are reliant on the internet and corporate networks to conduct daily operations. By protecting, or "backstopping," cybersecurity insurers, the administration hopes to prevent a major systemic financial crisis for insurers and victims during a cybersecurity incident. Every country has a National Cybersecurity Strategy institute which is constantly looking forward to doing? continuing research.

Due to the current political situation, with a war ongoing between Russia and Ukraine, the White House led by Joe Biden launched a National Cybersecurity Strategy on March 2023.

The US government is continuingly trying to strengthen the country's cybersecurity safety although its overall technology governance is one of the most powerful in the world.

Earlier that month, President Joe Biden released the new National Cybersecurity Strategy which outlines the steps the government is taking to secure cyberspace and build a resilient digital ecosystem that is easier to defend than attack - and that is open and safe for all [7], [2].

Why does the US need a National Cybersecurity Strategy?

The world is increasingly complex and cyberthreats are growing more sophisticated, with ransomware attacks running into millions of dollars in economic losses in the US. In 2022, the average cost of a ransomware attack was more than $4.5 million, according to IBM.

The greatest risks we face are interconnected, creating the threat of a "polycrisis", whereby the overall combined impact of these events is greater than their individual impact.

This is equally true of technological risks, where, for example, attacks on critical information infrastructure could have disastrous consequences for public infrastructure and health, or where growing geopolitical tensions heighten the risk of cyber-attacks. Cybercrime and cyber insecurity were seen by risk experts surveyed for the World Economic Forum's Global Risks Report as the 8th biggest risk in terms of severity of impact, across both the short term (next two years) and over the coming decade [13].

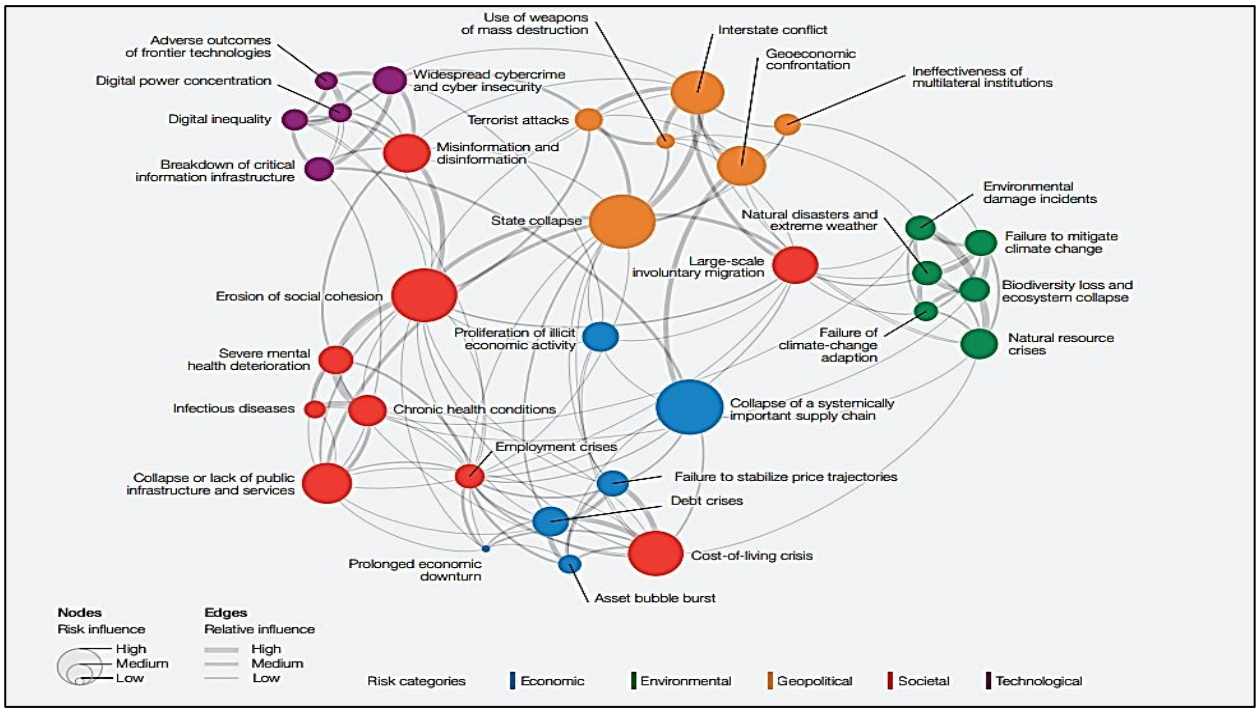


Fig. 4. Google data [13]

In 2022, state-sponsored cyber-attacks targeting users in NATO countries increased by 300% compared to 2020, according to Google data. (Diagram taken from [13])

With cyber-attacks on the rise, experts at the World Economic Forum's Annual Meeting at Davos predicted that 2023 would be a "busy year" for cyberspace with a "gathering cyber storm".

Cybersecurity protocols are the policies, procedures, and guidelines that a state government puts in place to protect its digital assets from cyber threats. These protocols are essential to ensure the security and confidentiality of critical data and systems.

The cybersecurity protocols of a state may include several components, such as access controls, incident response plans, data encryption, network monitoring, and employee training.

Access controls are measures that restrict access to sensitive information and systems to authorized personnel. Access controls can include passwords, biometrics, two-factor authentication, and other security mechanisms.

Incident response plans are procedures that outline the steps to be taken in the event of a cyber-attack or other security incident. These plans may include actions such as isolating affected systems, notifying law enforcement, and implementing backup and recovery procedures.

Data encryption is a technique used to protect sensitive information by converting it into an unreadable format. Encryption can help prevent unauthorized access to data, even if the system is compromised.

Network monitoring involves the use of software tools and techniques to monitor network activity for signs of cyber threats. This includes monitoring for suspicious activity, such as unauthorized access attempts or data exfiltration.

Employee training is a critical component of cybersecurity protocols. Employees are often the weakest link in cybersecurity, and their actions can inadvertently cause security breaches. Effective training can help employees understand the risks of cyber threats and how to identify and prevent them.

In addition to these components, the cybersecurity protocols of a state may also include regular security audits and vulnerability assessments, compliance with relevant laws and regulations, and partnerships with other organizations and governments to share threat intelligence and best practices.

The cybersecurity protocols of a state must be comprehensive, up-to-date, and regularly reviewed and updated to address new and emerging threats. By implementing effective cybersecurity protocols, states can protect their critical data and systems from cyber threats and ensure the safety and security of their citizens.

Another type of cyber-attack

A DDoS (Distributed Denial of Service) attack from one country to another can have serious consequences and is considered a cyber-attack on the targeted country's infrastructure. In a DDoS attack, an attacker typically uses a network of infected computers, known as a botnet, to flood a targeted website or network with traffic, overwhelming its servers and causing it to become unavailable to legitimate users.

A DDoS attack from one country to another can be used as a cyber weapon to disrupt critical infrastructure, such as government websites, financial systems, or healthcare networks. This can cause significant economic damage and impact the ability of the targeted country to function effectively.

In addition to the direct impact on the targeted country, a DDoS attack from one country to another can also have broader geopolitical implications. It can be seen as an act of aggression by the attacking country and may lead to diplomatic tensions or even military conflict.

To prevent DDoS attacks, countries can implement a range of measures, such as using anti-DDoS technologies, increasing network capacity, and improving incident response capabilities. In

addition, international cooperation is important in preventing and responding to DDoS attacks that cross national borders.

Overall, a DDoS attack from one country to another is a serious cybersecurity threat that can have significant consequences for both the targeted country and the broader international community. It is important for countries to work together to develop effective cybersecurity strategies and prevent these types of attacks from occurring [8].

DDoS falls into three types depending on the goal of the attack:

- *Layer Attack* - Sometimes referred to as a layer 7 DDoS attack (in reference to the 7th layer of the OSI model), the goal of these attacks is to exhaust the target's resources to create a denial-of-service.

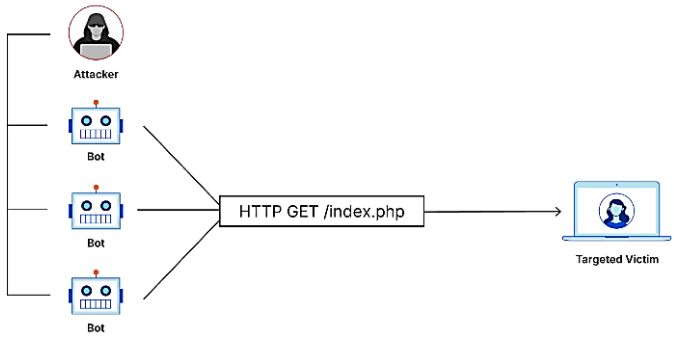


Fig. 5. Layer Attack [8]

- *Protocol Attack* - Protocol attacks, also known as a state-exhaustion attacks, cause a service disruption by over-consuming server resources and/or the resources of network equipment like firewalls and load balancers. Protocol attacks utilize weaknesses in layer 3 and layer 4 of the protocol stack to render the target inaccessible.

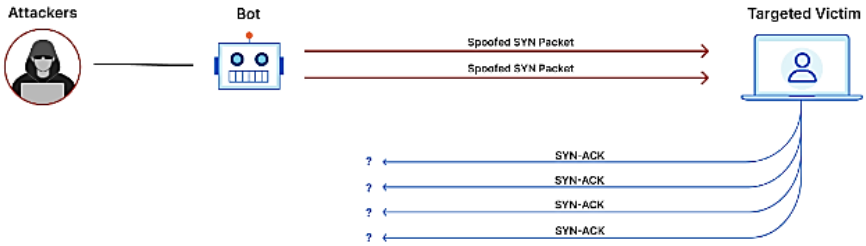


Fig. 6. Protocol Attack [8]

- *Volumetric Attack* - This category of attacks attempts to create congestion by consuming all available bandwidth between the target and the larger Internet. Large amounts of data are sent to a target by using a form of amplification or another means of creating massive traffic, such as requests from a botnet.

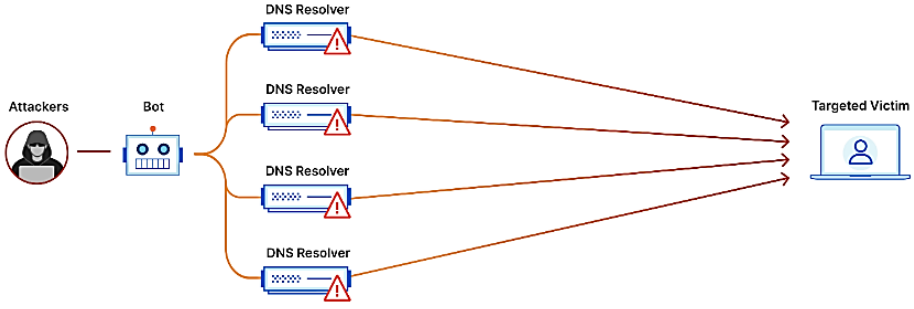


Fig. 7. Volumetric Attack [8]

The 5 pillars of the National Strategy [2] are:

1. Defend critical infrastructure.

To build confidence in the resilience of US critical infrastructure, regulatory frameworks will establish minimum cybersecurity requirements for critical sectors.

2. Disrupt and dismantle threat actors

Working with the private sector and international partners, the US will seek to address the ransomware threat and disrupt malicious actors.

3. Shape market forces to drive security and resilience

Grant schemes will promote investment in secure infrastructure, while liability for secure software products and services will be shifted away from the most vulnerable and good privacy practices will be promoted.

4. Invest in a resilient future

A diverse cyber-workforce will be developed and cybersecurity R&D for emerging technologies including postquantum encryption will be prioritized.

5. Forge international partnerships to pursue shared goals

The US will work with its allies and partners to counter cyberthreats and create reliable and trustworthy supply chains for information and communications technology.

4. Shielding the Nation's Digital Frontline: The Vital Role of the Military

The role of the army in cybersecurity attacks is vital in ensuring the safety and security of a nation's critical infrastructure and sensitive information. The military is responsible for protecting national security interests, which include cybersecurity. As the world becomes increasingly dependent on technology, the military's role in cybersecurity is becoming more important than ever.

One of the primary responsibilities of the military in cybersecurity is to protect the nation's critical infrastructure. This includes infrastructure such as power grids, water treatment plants, transportation systems, and financial institutions. A cyber-attack on any of these systems could have devastating consequences, including loss of life, economic disruption, and damage to the national security.

The military also plays a critical role in defending against cyber threats from other nations or state-sponsored groups. These threats could include attacks on government agencies, military organizations, and other critical infrastructure. The military is responsible for developing and implementing cybersecurity strategies to defend against these threats, as well as responding to and mitigating the impact of successful attacks. In addition to defending against external threats, the military is also responsible for ensuring the cybersecurity of its own networks and systems. This includes protecting sensitive information such as classified documents, personnel records, and communications. The military must also ensure that its own cybersecurity practices are up-to-date and effective, and that its personnel are trained in cybersecurity best practices.

One of the unique aspects of military cybersecurity is the use of offensive cyber capabilities. The military has the ability to use cyber-attacks as a means of disrupting or disabling an adversary's critical infrastructure or communications systems. This type of capability can be a powerful tool in military operations, but it also requires careful consideration and adherence to international law and norms [8], [9].

There are a number of cybersecurity protocols and practices that the army uses in order to protect against cyber-attacks. Here are a few examples:

1. Network Segmentation: This involves dividing a network into smaller subnetworks or segments. This helps to contain a cyber-attack to a single segment, preventing it from spreading throughout the entire network.

2. Firewalls: Firewalls are network security systems that monitor and control incoming and outgoing network traffic based on predetermined security rules. They help to prevent unauthorized access to the network and protect against cyber-attacks.
3. Intrusion Detection and Prevention Systems (IDPS): These systems monitor network traffic for signs of unauthorized access or malicious activity. They can help to identify and block cyber-attacks before they cause damage to the network.
4. Regular Software Updates and Patching: The army regularly updates its software and systems with the latest security patches and updates. This helps to protect against known vulnerabilities and reduce the risk of a successful cyber-attack.
5. Access Control: Access control measures are used to ensure that only authorized personnel have access to sensitive information and systems. This includes the use of passwords, two-factor authentication, and biometric identification.
6. Employee Training and Awareness: The army provides regular cybersecurity training and awareness programs to its personnel. This helps to ensure that all employees are aware of the latest threats and best practices for protecting against cyber-attacks.
7. Incident Response Planning: The army has a detailed incident response plan in place to quickly respond to and recover from cyber-attacks. This includes procedures for identifying, containing, and mitigating the effects of an attack [7].

USA Army is categorised in a list of intelligence gathering disciplines and these are:

- HUMINT
- GEOINT
- MASINT
- OSINT
- SIGINT
- TECHINT
- CYBINT/DNINT
- FININT

These are some of the intelligence categories, working together in order to ensure national state defense.

4.1. Army OSINT: Enhancing Intelligence Gathering with Open Sources

OSINT (Open Source Intelligence) and the army work together in a number of ways to help protect national security and support military operations.

One key area where OSINT can be valuable to the army is in the realm of situational awareness. OSINT can provide real-time information on a variety of topics, including geopolitical events, social media trends, and even weather patterns. By collecting and analyzing this information, OSINT analysts can help the military to better understand the operational environment, identify potential threats, and make more informed decisions.

Another area where OSINT can support the army is in the realm of intelligence gathering. OSINT can provide a wealth of information on a variety of topics, including enemy capabilities, military movements, and other intelligence targets. By collecting and analyzing information, OSINT analysts can provide the military with valuable intelligence that can be used to inform tactical and strategic decisions.

In addition to these areas, OSINT can also be useful in supporting military operations. For example, OSINT can be used to identify key influencers or opinion leaders in a particular region, which can help the military to better understand the local culture and build relationships with key stakeholders. OSINT can also be used to identify potential targets for cyber operations, such as vulnerable computer systems or critical infrastructure.

Also, The US Army is taking the issue of cybersecurity very seriously, as it recognizes the importance of protecting its networks and systems from cyber-attacks. One of the ways the army is combating cyber-attacks is by using OSINT (Open Source Intelligence) to gather information from publicly available sources [7], [9].

OSINT can be used to combat cyber-attacks in several ways. Firstly, it can be used to gather threat intelligence by collecting information about potential cyber threats and attackers. This information can be used to develop threat intelligence that can help the army better understand and mitigate cyber risks. Additionally, OSINT can be used to identify vulnerabilities in army networks and systems, prioritize security measures and remediation efforts, and assist with incident response efforts by providing real-time information about ongoing attacks.

OSINT can also be used to identify potential social engineering attacks against army personnel, such as phishing or spear-phishing attacks. By monitoring social media and other publicly available sources of information, OSINT can provide valuable insights into the tactics and techniques used by attackers to manipulate personnel and gain access to sensitive information.

Moreover, OSINT can help the army monitor the activities of cybercriminals and other malicious actors, including the buying and selling of stolen data and other illicit activities. This information can be used to identify new or emerging threats and to develop effective mitigation strategies to prevent attacks before they occur.

In summary, the US Army is leveraging OSINT to combat cyber-attacks and protect its networks and systems. By using OSINT to gather threat intelligence, identify vulnerabilities, assist with incident response efforts, and monitor cybercriminal activities, the army is better equipped to identify, prevent, and respond to cyber threats. OSINT plays a critical role in the army's comprehensive approach to combatting cyber-attacks and maintaining the security of its information systems.

5. Conclusion

In summary, the role of the military in cybersecurity is crucial in protecting a nation's critical infrastructure and sensitive information, defending against cyber threats from other nations or state-sponsored groups, and ensuring the cybersecurity of its own networks and systems. As technology continues to advance and cyber threats become more sophisticated, the military's role in cybersecurity will only become more important. It is essential that the military stays ahead of these threats and continues to develop and implement effective cybersecurity strategies to keep the nation safe and secure.

References

[1]. Cross-Sector Cybersecurity Performance Goals, March 2023: https://www.cisa.gov/resources-tools/resources/cpg-report.

[2]. Boozallen, What are the 7 types of security? https://www.boozallen.com/expertise/cybersecurity/national-cyber-strategy.html.

[3]. ESET, Part 1: “En Route with Sednit: Approaching the Target” Part 2: “En Route with Sednit: Observing the Comings and Goings” Part 3: “En Route with Sednit: A Mysterious Downloader”, October 2016: https://www.eset.com/afr/about/newsroom/press-releases-afr/research/dissection-of-sednit-espionage-group-1/.

[4]. Editorial Team from Front Lines, CrowdStrike’s work with the Democratic National Committee: Setting the record straight , 5 June 2020: https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/.

[5]. Malpedia, X-AGENT & MALWARE Reports, 2014-2020: https://malpedia.caad.fkie.fraunhofer.de/details/win.xagent.
[6]. CISAGOV & FBI, X-AGENT MALWARE FAMILY (Figures&Working Diagram), 29 December 2016: https://www.cisa.gov/sites/default/files/publications/JAR_16-20296A_GRIZZLY%20STEPPE-2016-1229.pdf.
[7]. World Economic Forum, Global Risk Reports, January 2023: https://www.weforum.org/reports/global-risks-report-2023/.
[8]. Cloud Flare, DDoS Attacks: https://www.cloudflare.com/learning/ddos/what-is-a-ddos-attack/.
[9]. The importance of cybersecurity in military, Nicole Allen, 20 Oct. 2021: https://saltcommunications.com/news/the-importance-of-cybersecurity-in-military/.
[10]. Fireeye APT28 & Programing Lines: http://www2.fireeye.com/rs/fireye/images/rpt-apt28.pdf.
[11]. National Security Technology Accelerator, 26 April 2022: https://nstxl.org/cybersecurity/.
[12]. Akshay Joshi, Daniel Dobrygowki, World Economic Forum, The US has announced its National Cybersecurity Strategy: Here's what you need to know, 9 March 2023: https://www.weforum.org/agenda/2023/03/us-national-cybersecurity-strategy/.
[13]. Cyberscoop, Christian Vasqyez, 16 December 2022: https://cyberscoop.com/apt28-fancy-bear-satellite/.
[14]. Editorial Team, Research & Threat Intel, Who is FANCY BEAR, 12 February 2019: https://www.crowdstrike.com/blog/who-is-fancy-bear/.

Methods for Detecting Malware Using Static, Dynamic and Hybrid Analysis

Alexandru-Radu BELEA
Faculty of Electronics, Telecommunications, and Information Technology,
University POLITEHNICA of Bucharest, Romania
alexandru.belea@stud.etti.upb.ro

Abstract

Malware analysis is the process of locating and examining malicious software or code with the aim of comprehending its operation and developing countermeasures. Malware can take many forms, such as viruses, worms, Trojans, and ransomware, and can cause significant harm to individuals, organizations, and even entire countries. To determine a piece of malware's purpose, potential effects, and capabilities, malware analysis entails examining the behavior, structure, and functionalities of the malware. Malware analysts are essential to the cybersecurity sector because they strive to spot dangers, eliminate them, and defend against online attacks. By using the knowledge gleaned from malware analysis, security solutions can be created that will better protect businesses from dangerous software. Malware analysis is a crucial part of any successful cybersecurity strategy in the continually changing threat landscape of today. In this article, we will explore the key concepts of malware analysis, including its purpose, techniques, and tools and we will contrast methods for detecting malware using static, dynamic, and hybrid analysis.

Index terms: dynamic analysis, hybrid analysis, malware, PE file, static analysis

1. Introduction

Malware analysis is the process of dissecting malicious software to understand its behavior and capabilities. With the increasing prevalence and sophistication of cyberattacks, the need for effective malware analysis has become more critical than ever. Malware analysts use a variety of techniques to analyze malware, including dynamic and static analysis, to identify its purpose and potential harm. The insights gained through malware analysis can be used to develop effective countermeasures and strengthen cybersecurity defenses. In this article, we will explore the importance of malware analysis, the different types of analysis techniques, and best practices for conducting effective analysis. We will also examine real-world examples of malware and how they were analyzed to provide valuable insights into the nature of cyber threats [1] [2].

Malware analysis is a crucial task in cybersecurity because it provides insights into the tactics, techniques, and procedures (TTPs) used by cybercriminals to create and distribute malware. By analyzing malware, cybersecurity experts can develop effective countermeasures to protect against malware attacks and improve the overall security posture of organizations and individuals. There are several types of malware analysis techniques, including static analysis, dynamic analysis, and hybrid analysis.

In addition to analyzing malware samples, cybersecurity experts also use malware analysis techniques to develop and improve security solutions, such as antivirus software, intrusion detection systems, and firewalls. By understanding how malware works and how it can be detected and

prevented, cybersecurity experts can design more effective security solutions that protect against both known and unknown threats [2] [3].

2. Malware Analysis Techniques

Malware analysis techniques are used to understand the behavior and characteristics of malicious software or malware. Malware can be designed to evade detection and analysis, so various techniques are used to overcome these challenges. One of the most widely used and effective techniques is static, dynamic and hybrid analysis, which we will look at in more detail (depicted in Figure 1). Other common malware analysis techniques are:

- Reverse Engineering: This technique involves decompiling the malware's code to understand how it works.
- Sandboxing: This technique involves running the malware in a controlled environment that isolates it from the rest of the system. Sandboxing can help identify the malware's behavior without the risk of infecting the host system.
- Memory Analysis: This technique involves analyzing the contents of a computer's memory while the malware is running. Memory analysis can help identify the malware's activities, such as code injection and network communication.
- Network Analysis: This technique involves analyzing the network traffic generated by the malware. Network analysis can help identify the malware's command-and-control servers and the data it is exfiltrating.
- Behavioral Analysis: This technique involves analyzing the malware's behavior to understand its intent. Behavioral analysis can help identify the malware's target and its objectives [3] [4].

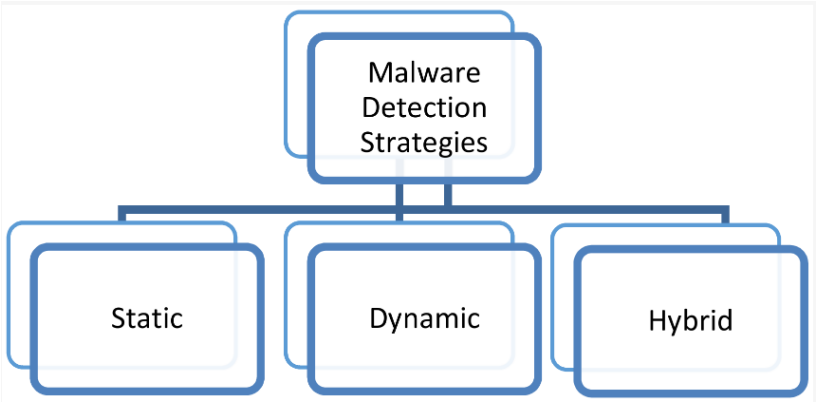


Fig. 1. Malware Detection Strategies [5]

2.1. Static Analysis

Static analysis involves examining the code and structure of the malware without executing it. This technique provides valuable insights into the behavior and potential impact of the malware on a system. Some common static analysis techniques include:

- Disassembling: Converting the binary code of the malware into assembly language to understand the functionality and logic of the code.
- Decompiling: Reversing the compiled code into high-level programming language code to understand the purpose of the malware.
- Debugging: Analyzing the code in a debugging environment to identify vulnerabilities and potential attack vectors [6].

2.2. Dynamic Analysis

Dynamic malware analysis is a technique used to analyze malware by observing its behavior in a controlled environment. This method involves running the malware in a secure and isolated environment, such as a virtual machine, and monitoring its activity to understand its behavior, capabilities, and potential impact. During dynamic malware analysis, the malware is executed and observed as it interacts with the system and network. This allows analysts to identify the malware's functions, including how it spreads, what it communicates with, and what it tries to achieve [6].

2.3. Hybrid Analysis

Hybrid analysis combines the strengths of both static and dynamic analysis. It involves first performing a static analysis to gather as much information as possible about the malware. This can include extracting any embedded files or configuration data, identifying any code obfuscation techniques, and looking for any signs of anti-analysis or anti-debugging techniques.

Once the initial static analysis is complete, the malware is then run in a controlled environment for dynamic analysis. This can include running the malware in a sandbox, using a debugger to step through the code, or running the malware on a virtual machine. The goal of the dynamic analysis is to observe the malware's behavior in a controlled environment and identify any malicious activity that may not have been apparent during the static analysis [3].

3. Types of Malware

Malware comes in various forms, each with its unique characteristics, effects, and methods of delivery. We will discuss some of the most common types of malware:

- Virus: A virus is a type of malware that infects executable files or system boot sectors, making them behave differently from their intended purpose. A virus replicates itself by attaching its code to other executable files, spreading its infection to other systems when the infected file is shared or transferred.
- Worm: Worms are self-replicating malware that spread over networks, exploiting system vulnerabilities to propagate from one computer to another. Unlike viruses, worms do not require a host program to spread, as they can self-replicate and spread autonomously. Worms can cause network congestion and slow down computer systems by consuming system resources, and they can also be used to steal sensitive information from infected systems.
- Trojan: A Trojan, also known as a Trojan horse, is a type of malware that disguises itself as a legitimate program to trick users into downloading or installing it. Once installed, the Trojan can give the attacker remote access to the infected system, allowing them to steal sensitive data, install other malware, or use the infected system as a part of a botnet.
- Ransomware: Ransomware is a type of malware that encrypts the victim's files, making them inaccessible until a ransom is paid to the attacker. Ransomware can be delivered through email attachments, infected websites, or social engineering attacks.
- Adware: Adware is a type of malware that displays unwanted advertisements on the victim's computer, usually in the form of pop-ups or banners. Adware can slow down the victim's computer, consume bandwidth, and track the user's internet activity.
- Spyware: Spyware is a type of malware that secretly monitors the victim's computer activity, collecting sensitive information such as login credentials, banking information, and personal data. Spyware can be used for identity theft, financial fraud, and espionage. Spyware can be delivered through infected websites, email attachments, or social engineering attacks [7].

4. The Advantages of Techniques

4.1. Benefits of Static Malware Analysis

There are several advantages of static malware analysis, including:

- Safe and controlled environment: Static malware analysis provides a safe and controlled environment for examining malware without the risk of infecting a real system.
- Ability to analyze large volumes of malware.
- Detection of hidden or obfuscated code: Malware authors often use obfuscation techniques to hide the true nature of their code. Static analysis can reveal the hidden code and make it easier to identify the malware.
- Identification of malware functionality: Static analysis can reveal the functionality of malware, including its ability to communicate with command-and-control servers.
- Can be used to create signatures that can be used by antivirus software to detect and block malware.
- Static analysis can be used for malware reverse engineering to understand how the malware operates and to develop countermeasures [2] [8].

4.2. Benefits of Dynamic Malware Analysis

There are several benefits to using dynamic malware analysis, including:

- Detection: Dynamic analysis can detect malware that may not be detected by traditional signature-based antivirus solutions.
- Identification: Dynamic analysis can identify the capabilities of malware, such as whether it has the ability to steal data or control a system.
- Rapid Response: Dynamic analysis provides security teams with the ability to rapidly respond to emerging threats.
- Flexibility: Dynamic analysis can be used to analyze a wide range of malware types, including new and unknown threats.
- Forensic Analysis: Dynamic analysis provides forensic investigators with a wealth of information about the behavior of malware, including network activity, system changes, and other actions [2] [8].

4.3. Benefits of Hybrid Malware Analysis

Some of the benefits of hybrid malware analysis include:

- Improved accuracy: Hybrid malware analysis can improve the accuracy of malware detection and analysis by combining multiple approaches, which increases the likelihood of identifying malicious behavior that may have been missed by a single technique.
- Faster analysis: By using multiple approaches simultaneously, hybrid analysis can speed up the malware analysis process.
- Better understanding of malware behavior: Hybrid malware analysis can provide a more complete picture of the behavior and capabilities of malware. For example, static analysis can reveal information about the structure and code of the malware, while dynamic analysis can provide insights into its runtime behavior.
- Effective response to sophisticated threats: Hybrid malware analysis is useful for detecting and responding to sophisticated threats that use advanced evasion techniques, such as polymorphism or obfuscation.
- Enhanced threat intelligence: Hybrid malware analysis can help build a more robust threat intelligence database by identifying commonalities and patterns among different malware samples [2] [4].

5. Malware Analysis on PE Files

PE (Portable Executable) format is a file format used for executables, DLLs (Dynamic Link Libraries), and other Windows operating system components. It was introduced in Windows NT 3.1 and has been used in all subsequent versions of Windows. The PE format is designed to be portable across different Windows systems, with the ability to handle different memory layouts and processor architectures. It contains information about the binary file, such as headers, sections, and resources [9] [10].

The main components of a PE file are: 1.DOS header (includes a small DOS executable to support older versions of Windows); 2.PE header (contains information about the file, such as the number of sections, entry point address, and the file checksum); 3.Section headers (describe the sections in the file, including their size, attributes, and virtual addresses); 4.Import and export tables (contain information about the functions and symbols that are imported and exported); 5. Resource section; 6.Debug information (contains debugging symbols and other information used by debugging tools) [9] [10].

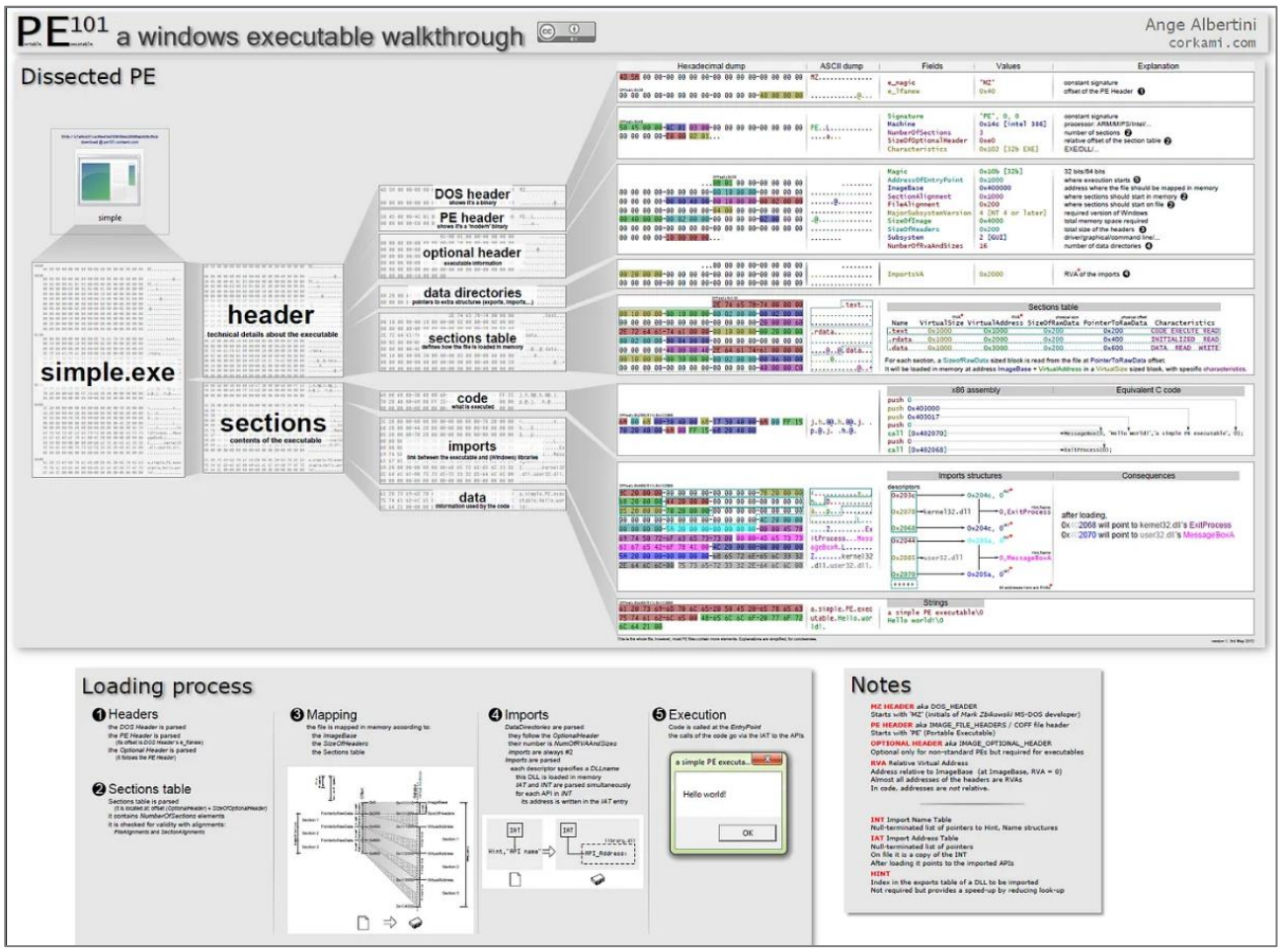


Fig. 2. PE file structure [9]

PE files are the primary format in which malware that targets Windows computers is distributed. Loaded libraries and imported functions are among the most crucial, if not the most crucial, pieces of data that we can statically extract from our malware. We can infer the features of the malware from these imported functions and libraries. For instance, the malware will employ some sort of network capabilities if it references the "ws2 32.dll" file. Another illustration is when malware imports the function "CreateProcessA" from kernel32.dll, indicating that the file will create a process

at some point while it is being executed. We'll be using a program called "Dependency Walker" to search our file and display any imported libraries and functions. Any libraries and imported functions that might be a sign of what the malware will do when run are what we need to be on the lookout for (Fig. 3) [9].

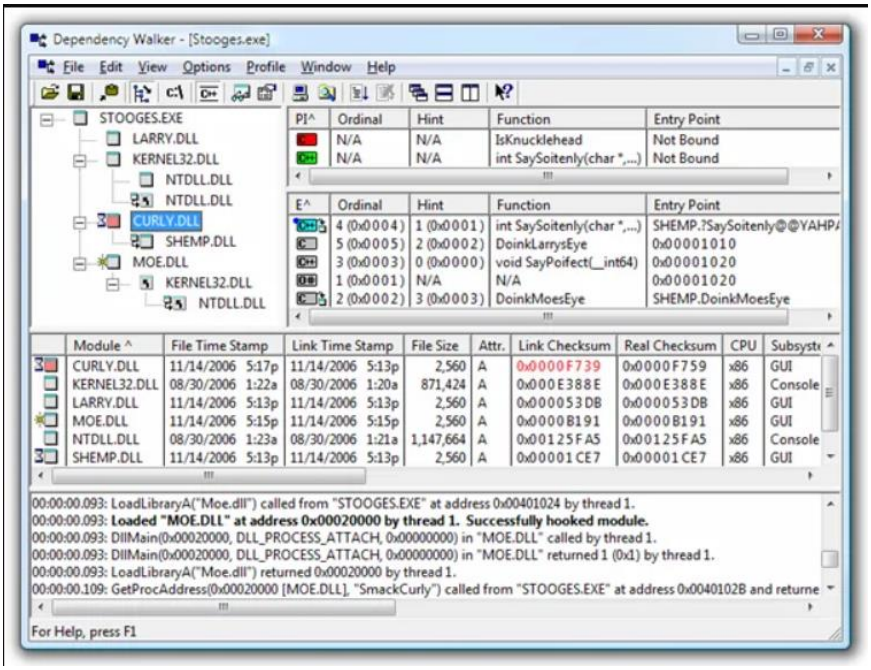


Fig. 3. Analysis of libraries [9]

The PE file format has sections and headers, as we've already explained. The resource part, also known as the.rsrc portion, is among the fascinating sections to examine. This area houses items like pictures, icons, and language strings. Malware writers occasionally use this area to conceal executables that will be used by the main program of the malware at some point during execution. Using Angus Johnson's "Resource Hacker," we may browse the resource area and begin looking for any "suspicious" or "malicious" indications (Fig. 4) [9].

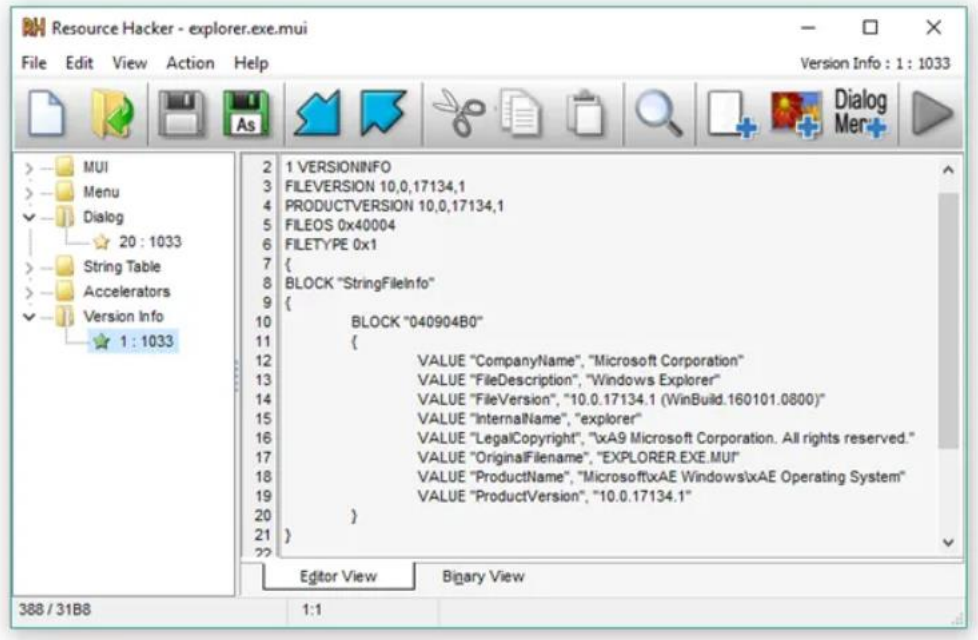


Fig. 4. Analysis of resources [9]

6. Conclusion

Following the static analysis we can see that it is not efficient compared to the dynamic or hybrid analysis. It is necessary to use different tools to analyze all the components of a PE file format, which implies a longer time to generate results.

In conclusion, static, dynamic, and hybrid malware analysis are important methods for detecting and analyzing malware. Static analysis examines code without actually executing it, while dynamic analysis observes the behavior of malware running in a controlled environment. Hybrid analysis combines both static and dynamic analysis techniques to provide a more comprehensive analysis of malware.

Each approach has strengths and weaknesses, and choosing the best analysis method depends on the specific context of the malware being analyzed. Static analysis helps identify known malware signatures and identify patterns that indicate malicious code. Dynamic analysis excels at detecting new, unknown malware that may evade detection by antivirus software.

Hybrid analysis offers the benefits of both static and dynamic analysis, allowing analysts to gain an in-depth understanding of malware behavior while identifying both known and unknown malware.

Overall, the choice of analytical method depends on the specific analytical goals and expertise of the analyst involved. Regardless of the method you choose, the ultimate goal is to detect and analyze malware and protect your system and users from its harmful effects.

References

[1]. Kurt Baker, “Malware Analysis,” 4 January 2022, Crowdstrike, https://www.crowdstrike.com/cybersecurity-101/malware/malware-analysis/.

[2]. Sihwail, Rami & Omar, Khairuddin & Zainol Ariffin, Khairul Akram. (2018). A Survey on Malware Analysis Techniques: Static, Dynamic, Hybrid and Memory Analysis. 8. 1662.10.18517/ijaseit.8.4-2.6827. https://www.researchgate.net/publication/328760930_A_Survey_on_Malware_Analysis_Techniques_Static_Dynamic_Hybrid_and_Memory_Analysis.

[3]. Damodaran, Anusha & Di Troia, Fabio & Visaggio, Corrado Aaron & Austin, Thomas & Stamp, Mark. (2017). A comparison of static, dynamic, and hybrid analysis for malware detection. Journal of Computer Virology and Hacking Techniques. 13. 10.1007/s11416-015-0261-z. https://www.researchgate.net/publication/288905288_A_comparison_of_static_dynamic_and_hybrid_analysis_for_malware_detection.

[4]. Chiradeep BasuMallick, “What Is Malware Analysis? Definition, Types, Stages, and Best Practices”, 19 August 2021, Spiceworks. https://www.spiceworks.com/it-security/data-security/articles/what-is-malware-analysis-definition-types-stages-best-practices/.

[5]. Tayyab, U.-e.-H.; Khan, F.B.; Durad, M.H.; Khan, A.; Lee, Y.S. A Survey of the Recent Trends in Deep Learning Based Malware Detection. J. Cybersecur. Priv. 2022, 2, 800-829. https://doi.org/10.3390/jcp2040041.

[6]. A M. Ijaz, M. H. Durad and M. Ismail, "Static and Dynamic Malware Analysis Using Machine Learning," 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 2019, pp. 687-691, doi: 10.1109/IBCAST.2019.8667136., https://ieeexplore.ieee.org/document/8667136.

[7]. Rabia Tahir, "A Study on Malware and Malware Detection Techniques", Department of Computer Science, Virtual University of Pakistan, https://www.mecs-press.org/ijeme/ijeme-v8-n2/IJEME-V8-N2-3.pdf.

[8]. Shijo, P.V. & Salim, A. (2015). Integrated Static and Dynamic Analysis for Malware Detection. Procedia Computer Science. 46. 804-811. 10.1016/j.procs.2015.02.149. https://www.researchgate.net/publication/276109044_Integrated_Static_and_Dynamic_Analysis_for_Malware_Detection.

[9]. Malware Analysis Techniques - Basic Static Analysis, Nasreddine Bencherchali, https://nasbench.medium.com/malware-analysis-techniques-basic-static-analysis-335a7286a176.

[10]. PE Format, Microsoft Article, 03/06/2023. https://learn.microsoft.com/en-us/windows/win32/debug/pe-format.

Author Guidelines

As an author, you are kindly advised to follow the next instructions. Reading and understanding the requirements before submittal would ensure adherence to the International Conference on Cybersecurity and Cybercrime standards and would facilitate acceptance by the scientific reviewers.

1. Papers must be submitted in English having an even number of pages (minimum 4 pages). At least 50% of the last page should be occupied by text.
2. For papers writing it is recommended the use the text processor Microsoft Word and one of the template models found on the conference website. We will do the final formatting and all necessary format conversions of your paper.
3. The papers will be submitted using our online interface. Please do not send your papers by email.
4. The papers will be reviewed by two scientific reviewers, well-known in their domains of activity. Usually, it takes 1 to 3 months between the moment you finished your submission and a response is given by scientific reviewers.
5. The papers will be sent back to the authors for corrections if the figures, pictures, or tables are not contained in the text or if the reviewers require modifications or supplementary information.
6. The papers will be rejected if their scientific content is not adequate, if they don't contain original elements and if they are not properly written in English.
7. The bibliography must show the authors adequate documentation. At least 7-10 quality references should be cited.
8. Citation standard is IEEE. Please read the IEEE Citation Reference from the website: www.ieee.org/documents/ieeecitationref.pdf.
9. The whole responsibility for the calculation exactitude, experimental data, scientific affirmation, and paper translation belongs to the authors.
10. The authors will declare on their own responsibility that the article or parts of it were not published before in other journals.

More information: proceedings.cybercon.ro/index.php/ic3/author-guidelines

The Romanian Association for Information Security Assurance (RAISA)

The Romanian Association for Information Security Assurance (RAISA) is a professional, non-governmental, non-partisan political, nonprofit, and public benefit association.

RAISA AIM

The aim of the Romanian Association for Information Security Assurance is promoting and supporting information security activities in compliance with applicable laws.

RAISA VISION

The vision of the Association is to promote research and education in information security field and to contribute to the creation and dissemination of knowledge and technology in this domain. RAISA has a strong representation at the national level, bringing together professors and researchers from top universities and Romanian institutions, PhD, master's, and license students, as well as companies in the IT segment.

RAISA OBJECTIVES

To achieve the stated purpose, the Romanian Association for Information Security Assurance proposes the following objectives:

- Collaboration with the academic community from Romania or abroad in order to organize conferences, scientific seminars and workshops for presenting the development and implementation of effective measures to improve information security.
- Collaboration with research centers, associations, and companies from Romania or abroad, to organize informative events in information technology security field.
- To perform specific programs for education and training of personnel involved in electronic information management (data processing, storage, security).
- To ensure the dissemination of notice relating to existing vulnerabilities and nationally and internationally newly identified threats; to provide solutions for data restoration and policies to prevent and combat incidents based on the information provided by suppliers of software solutions.
- To publish scientific journals for university staff, PhD students or master's students, researchers, students, and other professional categories in the field of information security and cybercrime.
- To grant awards, scholarships, or sponsorships to people with outstanding merits in the field of information security.

Website: www.raisa.org
Email: contact@raisa.org

RAISA Members Benefits

RAISA MEMBERS

The Romanian Association for Information Security Assurance (RAISA) is an organization that consists of:

- **Founding members** - are individuals who have participated in the founding process of the Association, have agreed with the Statute of the Association at the date of establishment and are parts of the members' category, with all their rights. The founding members pay annual membership fee and have the right to deliberative vote during the General Assembly.
- **Members** - are individuals who have joined the Association after the date of establishment. The members pay annual membership fee and have all the rights, respecting the obligations stipulated in Statute of the Association. They have the right to deliberative vote during the General Assembly.
- **Honorary Members** - can be scientists, professors, cultural or religious personalities, valuable professionals, who have rendered outstanding services to the Association. They are exempted from contributions and their vote is advisory.
- **Collaborators/Volunteers** - anyone who wants to participate in Association activities without becoming a member. Their collaborations are on no-cost basis; they don't pay a membership fee and don't have the right to vote.

RAISA MEMBERSHIP BENEFITS:

- Free access to RAISA events.
- Discount to workshops and conferences supported by RAISA.
- Discount for professional courses organized by RAISA.
- Possibility to be involved in RAISA projects and campaigns, support offered for research.
- Free publishing for scientific articles in the International Journal for Information Security and Cybercrime (IJISC), indexed in international databases.
- Discount for books and scientific studies promoted by RAISA.
- The possibility of promoting the events on RAISA media channels:
    - www.securitatea-cibernetica.ro
    - www.securitatea-informatiilor.ro
    - www.criminalitatea-informatica.ro

Get the most from your membership!

www.raisa.org/raisa-members/