Phishing Attacks Detection

Serhii Buchyk, Dmytro Shutenko and Serhii Toliupa

Taras Shevchenko National University of Kyiv, Bohdan Hawrylyshyn str. 24, Kyiv, 04116, Ukraine

Abstract

It is safe to say that the number of phishing attacks is dramatically increasing each year as the world becomes more and more digital. Security specialists are doing a good job developing social engineering countermeasures but there is an increasing workload to be dealt with as attackers constantly come up with more sophisticated ways to deceive workers. This work features approaches used for phishing detection: human and technology-based. It also discusses issues associated with both those methods and addresses the difficulty of eliminating phishing altogether. While using technology can lessen the burden on humans, a balance must be achieved where there is no complete reliance on either humans or technology as both have proven to have their own flaws.

Keywords

social engineering, phishing, mitigation methods, human based detection, technology based detection.

1. Introduction

Phishing is a cyber-criminal activity where a social engineer baits its target for information and passwords by masquerading as a trustworthy party. Before it was popular on the Internet, phishing was performed by phone, and the technique was referred to as vishing [1]. The current method of phishing over the Internet is most often carried out in the form of an e-mail or pop-up directing the target to a page similar to the page target is well familiar with, this page usually will prompt the user to enter their credentials either to log in to engage in a fabricated scenario, the rest is history. Mail-out is another phishing technique social engineers use to gather information. An example of a mail-out is a survey given to employees of an organization where they are asked to answer a few questions of 'their company's IT department'. Mail-out is a technique that can also be used to spread malware, usually attached to the files sent out to the target [2]. Interestingly enough, attached files don't even have to be executable from the first look. It means that a file with any kind of extension can conceal a potential threat of data breach which is a very unwanted event for any business, state institution or simply a private individual.

2. Task formulation

The aim of the study is research and efficiency analysis of known phishing attacks detection methods, display of their flaws.

The object of research is the process of detecting phishing attacks.

The subject of research is methods for phishing attacks detection.

Thus, the task is to analyse known methods for phishing attacks detection both human and technology based, provide method to assess such systems' effectiveness in given scenarious and discuss potential flaws associated with deployment and utilization of such systems.

[1] *Information Technology and Implementation (IT&I-2022), November 30 - December 02, 2022, Kyiv, Ukraine*

EMAIL buchyk@knu.ua (A. 1); dima.shutenko@knu.ua (A. 2); tolupa@i.ua (A. 3)

ORCID: 0000-0003-0892-3494 (A. 1); 0000-0002-6895-0438 (A. 2); 0000-0002-1919-9174 (A. 3)

3. Solving the task

Phishing attacks can be divided into two categories, which are human based social engineering which includes real-life direct physical interaction with its human victim through phone, and also technology-based social engineering such as online social networks impersonation, website phishing scams, and email phishing [3]. For the past decades, numerous researchers all around the world have been studying ways to detect and prevent social engineering attacks. Figure 1 illustrates the taxonomy of phishing detection methods known to date. The following few paragraphs of the article will focus on discussing both types of phishing detection methods as well as the approaches information security specialists could take advantage of on their way to preventing the leak of sensitive data.

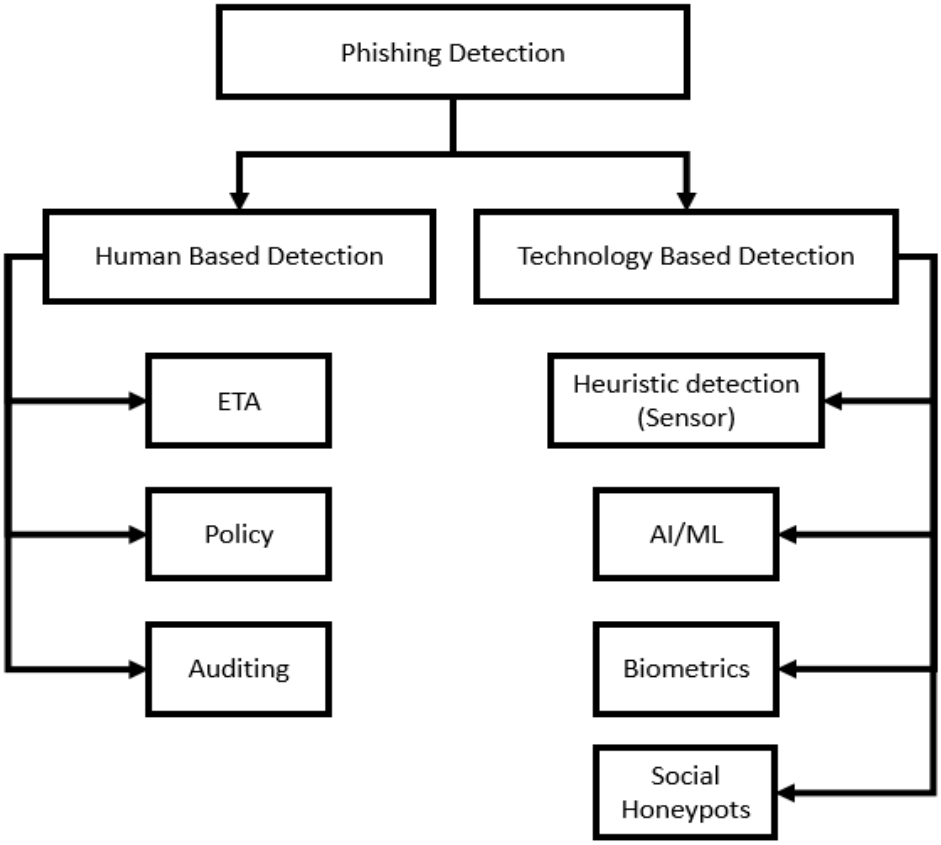

Figure 1: Taxonomy of Phishing Attacks Detection Methods

Human based detection methods involve human intervention in detecting and preventing phishing attacks. Human based detection focuses more on the judgment of humans to determine whether the activities that they have encountered are in any way related to social engineering attacks. Historically it was the first and for some time the only available at the time method to detect social engineering attacks due to the absence of automated systems, it is therefore quite well researched. Currently, there are three approaches that can be classified in human based mitigation. Those are policy, auditing, and also education, training and awareness approaches that will be discussed in the following paragraphs. The importance of these approaches is well highlighted in studies [3-10] that have researched phishing attacks mitigation methods using human decision-making.

Education, training and awareness (ETA) approach in detecting social engineering is one of the best researched approaches in human based mitigation. Many works including [3] emphasized that employee education is important to ensure the policies, procedures and standards that have been developed in the organization are to be deployed effectively. It is also suggested that ETA must be implemented especially for the newly employed staff in their orientation phase right after onboarding ends. ETA is best implemented by developing interactive social engineering awareness websites for staff training promoting personnel awareness of social engineering attack vectors. This interactive learning and

education game-based system proved to be an effective education tool in providing the users of this system with knowledge and experience in spotting social engineering and its attack patterns [4]. It is also the modular-based design of the environment that can be particularly handy as the system is to be updated with the latest trends and additional techniques of social engineering attacks in hours as opposed to days needed to develop a brand new platform. This method proves to be vital as most victims fall for phishing attacks because they lack knowledge about the attack vectors and are ignorant of passive warnings from security tools regarding phishing attacks. There is no doubt security training if done right help employees enhance their classification accuracy and teach them to take necessary action in preventing them more often.

Well established organizations are known to enforce certain rules developed to help personnel in detecting and preventing social engineering attacks including phishing. These rules are usually directed by policies that guide personnel on how to analyze and make a decision whether the situation they encountered is indeed a social engineering attack or a legitimate activity. Their further actions should also be part of a well-documented procedure to make sure the staff doesn't become the reason of data breach or having a malicious actor in the organization's corporate network. Most important bits of data in policies modern organizations simply can't function without are the following. Clear desk policy to prevent password or sensitive information being left lying about; paper shredder usage procedure to undermine dumpster diving attempts; identification checking policy (implementation of caller ID technology for phone calls and service personnel); rules of defining sensitive information in an organization, authorization and access control policy, data classification policy and security policies [5]. As proposed by the Colorado Department of Education in [6], Audit controls and effective security safeguards are part of normal operational management processes to mitigate, control, and minimize risks that can negatively impact business operations and expose sensitive data.

Auditing is complimentary to the policy-based approach as mentioned in works [5,7]. While auditing is often used to inspect or examine processes or systems to ensure compliance with requirements, our interest towards auditing lies in testing the level of user awareness or exposure to social engineering attacks [8]. This approach is frequently used to ensure the effectiveness of policies and ETA conducted in an organization against attacks. One key difference between processes and ETA + Policy auditing lies in the nature of phishing attacks. Since new attack vectors are emerging quite often, it makes perfect sense to carry out auditing procedures with higher frequency than those that assess processes or systems.

Approaches mentioned in previous paragraphs are the most fundamental and common countermeasures in detecting and preventing social engineering attacks including phishing. Policy, auditing and ETA for users and employees in the organizations are a must as social engineering preys on psychological traits in exploiting their victims [9].

We can't argue that other technology based detection solutions help users to recognize the attacks, but at the end of the day, it depends solely on the decision-making and action taken by those very users so that they classify the situation they encountered as suspicious and take necessary measures as pointed out in procedures for dealing with potential social engineering attacks. Human judgment is undoubtedly somehow subjective and even with good knowledge, awareness and policy against this type of threat, social engineers can find multiple ways to convince their victims and exploit human psychology to gain information or access to sensitive information causing data leaks or other malicious activity. Therefore, there is a need for technology based mitigation methods as complimentary to human based mitigation to increase the overall detection and prevention accuracy.

According to Kevin Mitnick, another problem is that the most popular targets for social engineering exploitation are new employees. Mitnick argues this is because new employees and interns are one of the weakest links in an organization. They may not yet have completed ETAs and do not possess sufficient knowledge about the company's sensitive information assets, as well as they are not familiar with all the staff within the organization or relevant business processes they became a part of. As one would expect, they can be easily fooled. What appears even more striking is that according to [10] even with the best security education, awareness and training programs in place, new employees will always represent a threat. Therefore, the best approach information security officers might take to avoid such a scenario would be to limit new employees' access to sensitive organizational assets. The challenge here is that by doing so, staff that has just been on-boarded would simply become incapable of carrying out their duties as there would be an obstacle to accessing the information and resources that are required to get the job done.

Another set of methods used for detecting and preventing phishing attacks takes advantage of a wide range of technology based solutions. Technology based mitigation methods entered the equation as soon as phishing become more widespread and users were in urgent need of both email filtering solutions and website checkers to do some analytical job for them in the background. Since then it's been a constant bout between hackers and security specialists that both developed more and more sophisticated methods to outperform one another. Technology based mitigation methods have been well-researched in detecting and preventing social engineering attacks in the last decade. Several categories represent this method. In the next few paragraphs of this work an overview of the following phishing technology based detection methods will be presented:

1. heuristic detection;
2. artificial intelligence and machine learning powered solutions;
3. biometrics;
4. social honeypots.

First and foremost, there's a variety of heuristic methods that detect phishing patterns with the help of digital signatures or other identifiers, object properties, etc. In [11] phishing detection by heuristics is defined as software that is deployed on the server or client side to inspect payloads of different protocols via diverse algorithms. Protocol list includes HTTP, HTTPS, SMTP, POP3 or any arbitrary protocol used to deliver content/emails to Internet users. Algorithms could be represented as any method to detect and if configured so block phishing attempts automatically.

In [12], the authors suggested an anti-phishing approach that examines webpage irregularities. The method collects anomalies from a variety of sources, including URLs, page titles, cookies, login forms, DNS data, and SSL certificates, among others. If a set of universal heuristic examinations are recognized, as a result of comparing to a massive dataset of known malicious patterns, such software might detect zero-day phishing attacks. Some may say that it doesn't really give this approach a competitive advantage over blacklists. Since blacklists require exact matches to detect phishing websites, the exact same phishing attacks need to be examined first to blacklist them [13].

However, as heuristic methods focus on signatures comprised of similar patterns, they are more prone to identify malicious payload never seen before with higher probability which makes them more flexible but at the same time creates a risk of misidentifying legitimate websites and producing false positives disrupting normal workflow and unwanted system overhead. Mainstream mail clients and web browsers have already begun to equip their services with phishing protection technologies, such as heuristic based detectors that help at identifying phishing attacks. What is more, phishing detection based on heuristics is incorporated in countless antiviruses so in a world where hackers didn't tweak their attacks it would be a matter of time before a perfect set of signatures could be created and used to identify any potential threat. Secondly, there are more modern and powerful methods that only get better with time - artificial intelligence and machine learning powered solutions. Algorithms they use learn from massive databases of known phishing websites, emails or even SMS and therefore can spot a suspicious entry with quite a high accuracy. Study [14] considers the phishing detection problem as an AI-based classification problem wherein the result of the decision-making phase leads to detecting if a given website is either a legitimate or a phishing website. In essence AI's job is to conduct analisys of ever-changing phishing patterns, determine the combinations of characteristics that should be used to successfully identify malicious activity and filter out data that is no longer useful. Thus, consideration of the AI algorithms as the basis for developing viable phishing detection models to combat phishing threats in their evolving nature was made in [15].

In short, many AI-based solutions take advantage of systematized knowledge about significant characteristics that have proven to be efficient in spotting elements phishing websites are prone to possess as [16] suggests. Most commonly used characteristics or features as well as phishing website attributes that can be used for phishing detection with high accuracy are featured in Table 1.

Determining heuristic and artificial intelligence + machine learning powered methods efficiency for phishing detection in theory is determined by correctly weighted set of attributes and their individual entropy. Since it is exactly the job of AI and ML to determine most accurate estimate of each attribute importance and hence assign higer weight to it, for this work we will come up with initial values based on existing knowledge and as such obtained results will reflet the effectiveness of the most primitive heuristic-based phishing detection system in percentage from 0% to 100%.

Table 1
Website attributes

Address Bar based features	HTML and JavaScript based features	Domain based features	Abnormal based features
Use of IP address	Website Forwarding	DNS Record	Request URL
Long URL to hide the suspicious part	Status Bar Customization	Website traffic	Server form handler (SFH)
Adding prefix or suffix separated by '-' to the domain	HTML links to third-party resources like Google Analytics, Facebook, Cloudflare, etc.	Domain registration length	Links in <Meta>, <Script> and <Link> tags
URL's having '@' Symbol	Using Pop-up window	PageRank	URL of anchor
Existence of 'HTTPS' in the domain part of the URL	IFrame redirection	Google Index	Submitting information to Email
Using URL shortening services	<Body> length in tags	Number of links pointing to page	Abnormal URL
Sub domain and multi sub domains	Disabling Right Click	Statistical-reports based feature	
HTTPS (HTTP with SSL)	Missing Title	Using non-standard port	
Redirecting using '//'	Favicon		

To get each individual attribute entropy, slightly modified Shannon's concept will be taken up. Formula 1 is used to calculate it

$$H_{attribute} = p_1 \log_s(1/p_1) + p_2 \log_s(1/p_2) + ... + p_s \log_s(1/p_s), \quad (1)$$

where $p_1, p_2 ... p_s$ are the probabilities of attribute having unique values, s is the number of unique values for a given attribute.

Sample distribution of attributes' weights and unique values with their probabilities can be used as such given in Table 2 and Table 3 respectively. Effectiveness of a system is a product of all weighet attributes' entropies. Formula 2 is used to calculate it.

$$E = \sum_{i=1}^{k} w_i H_i \quad (2)$$

where H_i and w_i are entrophy and a weight of a given attribute respectively, k is the number of attributes considered for a given system.

Eventually, having extended the list of attributes and their probable values and probabilities, the data can be fed into machine learning model for it to determine the weights of each attribute so that peak efficiency can be reached. What is more, this approach of system efficiency assessment can be used to determine attributes which no longer provide any use for phishing elements detection.

The reason ML approaches became popular for phishing detection is because they made it a simple classification problem. Training of ML model for a learning-based detection system requires the data at hand must-have features that are related to phishing and legitimate website classes mentioned earlier. Previous studies show that detection accuracy is high as robust ML techniques are used; those are k-Nearest Neighbor (KNN), Random Forest, and Support Vector Machine (SVM) to name a few. Figure 2 [17] illustrates one of the scenarios used to train a ML model to differentiate phishing websites from legitimate ones. Each algorithm has a little-to-no influence on the common approach taken for machine learning: a dataset of entries containing phishing has to be used anyway. In [18] it was discovered that the Random Forest model proved to deliver the best performance in a given setting. This algorithm is a collection of Decision trees with each tree differing slightly from the other. A prediction is made by averaging the result obtained from all individual Decision Trees. This helps to reduce the problem of overfitting, a problem peculiar to the Decision Tree Algorithm.

Thirdly, since social engineers are frequently attempting to impersonate a trustworthy party by creating a fake profile and mimicking its identity through visual appearance, use of lingo and knowledge of internal business processes a method that can counteract physical impersonation is using biometrics

as it does not rely on the perceived identity of a person, but rather distinguishes someone using their unique biological traits such as fingerprint, voice or facial recognition [19].

Table 2

Sample Distribution of attributes' weights for a given set of attributes

Address Bar based features		**HTML and JavaScript based features**		**Domain based features**		**Abnormal based features**	
Attribute name	**weight**	**Attribute name**	**weight**	**Attribute name**	**weight**	**Attribute name**	**weight**
Use of IP address	1/32	Website Forwarding	1/32	DNS Record	1/32	Request URL	1/32
Long URL to hide the suspicious part	1/32	Status Bar Customization	1/32	Website traffic	1.2/32	Server form handler (SFH)	1/32
Adding prefix or suffix separated by '-' to the domain	1.5/32	HTML links to third-party resources like Google Analytics, Facebook, Cloudflare, etc.	1.5/32	Domain registration length	1.8/32	Links in <Meta>, <Script> and <Link> tags	1/32
URL's having '@' Symbol	1/32	Using Pop-up window	0.5/32	PageRank	0.5/32	URL of anchor	1/32
Existence of 'HTTPS' in the domain part of the URL	1/32	IFrame redirection	1/32	Google Index	1.3/32	Submitting information to Email	1/32
Using URL shortening services	0.8/32	<Body> length in tags	0.5/32	Number of links pointing to page	1.4/32	Abnormal URL	0.7/32
Sub domain and multi sub domains	1/32	Disabling Right Click	0.5/32	Statistical-reports based feature	0.8/32		
HTTPS (HTTP with SSL)	1/32	Missing Title	1/32	Using non-standard port	1/32		
Redirecting using '//'	1/32	Favicon	1/32				

Table 3

Sample set of unique values with their probabilities for a given attribute

Attribute: Links in <Meta>, <Script> and <Link> tags	
Values	**Probability**
No links in any of the tags	0.1/8
Links in <Meta> tag only	0.4/8
Links in <Meta> and <Script> tags	1/8
Links in <Meta> and <Link> tags	2/8
Links in <Script> tag only	0.2/8
Links in <Script> and <Link> tags	1.3/8
Links in <Link> tag only	0.5/8
Links in all the tags	2.5/8

Biometric systems have improved significantly in recent years and visual disguises that may fool a human will not be successful when confronted with them. It's important to state that in comparison with other approaches to phishing detection, this one requires a more specific setting to be of any use. Apparently, biometric detection can only be useful if the attacker is forced to be subjected to this kind of test. An example would be using two or even three-factor authentication for an account one of which is biometrically enforced. Last but not least, there are systems called honeypots that imitate an existing working system to trap attackers and learn their behavior patterns to develop better signatures as well

as relevant phishing patterns for further automatic filtering. Honeypots can be implemented in any form including a website, network, or computer. While it's traditionally used to defend against threats such as malware and database attacks, it can be also utilized to learn more about email attacks and spam attacks.

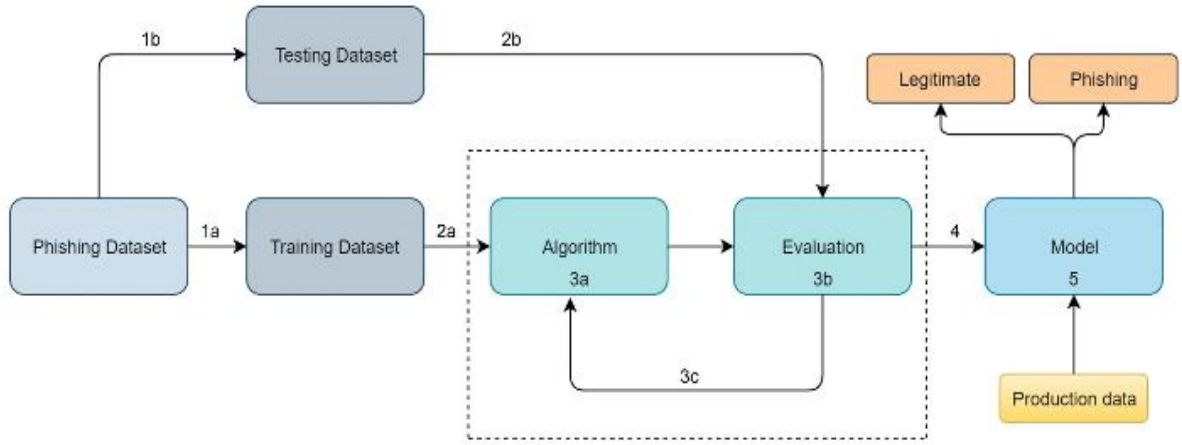

Figure 2: Machine learning model for phishing attack detection

Honeypot is primarily used to gather data points on hacker's actions that will be further used to form a data set that will eventually be fed for training (ML models are discussed in more detail in previous paragraphs). Therefore, the main purpose of using honeypot auto harvest of information based on hacker's activities on the system, filter certain activities and develop a statistical user model.

For social media honeypots, detection for spamming and phishing will need manual work with personnel operating the honeypot profile as many spam works are in form of video, image, text and social network features being manipulated [19]. Having collected enough input data, statistics can be drawn to differentiate between real profiles, fake profiles, spam profiles or bot profiles and automatically filter our unwanted entries in the future. The use of technology is often accompanied by added cost, complexity and overall system overhead. The systems that have been mentioned in this work would require a significant financial investment by an organization in most cases without a clear measure of cost-benefit once they will be deployed. Thus, spending large amounts of money on such systems, their training, deployment, management and maintenance can be irrelevant. The added complexity of the systems also means that there is potential for a business process to be interrupted in case those systems malfunction yielding false positives one after another.

While heuristic detection methods do quite well in terms of spotting known phishing patterns, researchers argue they barely outperform simple blacklists but create substantial system overhead as they require a signature of the entry to be built every single time and then comparing it to countless known signatures in a given database. Such solutions can be useful in case a snippet can be sent to the cloud, analyzed and compared with known signatures and then the result should be sent to the host notifying the user of a potential threat or filtering out such content automatically even before the user can see it (especially useful with emails and their attachments). In case all processes associated with payload analysis take place on the host machine it is expected to overwhelm a system and lead to noticeable freezes interfering with normal workflow.

Artificial intelligence systems usually require large datasets and long periods of training to be effective. In [19] authors argue the issue is that the datasets required are hard to come across unless there have been specific efforts to gather samples. The datasets themselves may also be of limited use as they become outdated if not updated with newly detected samples. Older datasets are becoming obsolete because of the phishing attacks' nature, as soon as defense mechanisms manage to detect and mitigate them automatically, hackers come up with more sophisticated approaches. Another thing worth mentioning is that the information-gathering process itself can be subject to inaccuracy and therefore high false positive rate, making the system inconvenient rather than effective. Not to mention the unpredictability of machine learning algorithms and the time needed to train, test and deploy them.

Biometrics-based systems can be bypassed using piggybacking, tailgating or other social engineering tactics. The attacker can also exploit the technological vulnerabilities of the security

systems on-premise, thus avoiding detection. In [20] it was shown that biometric systems are especially vulnerable to targeted impersonation attacks without manipulating the actual mechanisms of the device. This creates an opportunity for a social engineer to manipulate authentication devices and avoid detection. All things considered, biometrics is a good tool to prevent unauthorized physical access to facilities but offers little-to-no competitive advantages to other methods of phishing detection.

Despite being quite a progressive technological solution honeypots are banned in some countries as it is against user privacy rights to collect data generated by their browsing activity. Even though it sounds ridiculous taking into consideration modern-day cooking policy of Google or Meta products, according to [21] proactive security specialists can be charged with a breach of privacy. Therefore, a thorough analysis of a legal aspect has to be carried out before developing and adding such systems. Another problem is that honeypot is a relatively new approach and not many accurate datasets are collected, which makes the ratio of false positive and false negative quite high, which results in inaccurate system execution. Thankfully, now this data doesn't need to be analyzed by engineers, in contrast, it can be fed into the ML model and further used to determine the most prominent detection attributes and malicious actor behavior patterns to generate warnings or block similar activity in the future. Hence, it is not fair to say that technologies address phishing threats with no complications associated with their utilization and in most cases company's CISO or CEO has a tradeoff between spending little money on security and having a badly protected but fast network that ensures smooth business processes continuity and spending more money on security systems and being sure technical solutions are there to protect company's information assets from expected threats.

4. Conclusions

Our increasing reliance on the Internet for much of our day-to-day operations has created the ideal setting for fraudsters to launch targeted phishing assaults [22]. It is nowhere to hide from so measures must be taken to alleviate risks of data breaches and loss of companies' informational assets as a result of employee carelessness. In this work, phishing was discussed as one of the most sophisticated types of social engineering attacks that takes advantage of human psychological weaknesses. Phishing attacks have been a threat to both organizations and individuals for a very long time and although it has been a known threat with many cases of security incidents involving social engineering, to date there has not been a clear answer on how to deal with this threat and thoroughly mitigate it.

For a long time, it has been proposed that any social engineering threat can be prevented through the use of security policies, ETA of employees and establishing a security culture within the organization and regular audits. Taking into consideration the nature of such attacks nowadays – more complex solutions have to be used. It has been discussed that human based detection methods are simply no longer enough on their own as naive adoption of best security practices does not guarantee good security posture of the organization.

It has been shown that various technology-based solutions exist to automate phishing detection. These technological systems can lessen the impact of human weakness in detecting attacks as they occur. Among the measures presented were those that used heuristic detection, systems powered by AI and ML, biometrics and social honeypots that can be used to progressively learn about and adapt to ever-changing social engineering tactics. However, relying on technology also has its drawbacks in terms of cost, management and maintenance.

An approach to assess technology-based systems effectiveness in terms of website phishing elements detection has been offered. Since it has website attributes that are most often analysed in search of phishing patterns at it's foundation – the system can be improved by extending the list af attributes, their values and probably combinations of other factors.

The threat of social engineering and thus phishing can never be totally eliminated as long as an organization requires human beings to do the job systems are not yet capable of. Using technology-based solutions can lessen the burden on humans in providing security, though a balance must be achieved where there is no total reliance on either of those methods as both have their own issues and weaknesses. Moving forward, the best thing that can be done to combat social engineering and phishing in particular is to continue researching how organizations are being exploited, use honeypots to learn more about attackers' behavior, improve existing security standards and develop new solutions to detect and mitigate existing and emerging threats.

5. References

[1] Terranova Security, WHAT IS VISHING? URL: https://terranovasecurity.com/what-is-vishing.

[2] Mitnick, K. D (2003). Are You The Weak Link. Harvard Business Review, 81(4).

[3] Peltier, T. R. (2007). Social Engineering: Concepts and Solutions. Information Systems Security. Volume 15(5), pp. 13-21.

[4] Khonji, M. et. al. (2013). Phishing Detection: A Literature Survey. IEEE Communications Surveys & Tutorials. Volume 15(4), 2091-2121. IEEE.

[5] Twitchell, D. P. 2006. Social Engineering In Information Assurance Curricula. Proceedings Of The 3rd Annual Conference On Information Security Curriculum Development. 22-23 September. Kennesaw, Georgia, United States: ACM, pp. 191-193.

[6] Information Systems Audit Policy. URL: https://www.cde.state.co.us/dataprivacyandsecurity/informationsystemsauditpolicy

[7] Algarni, A., Xu, Y., Chan, T., & Tian, Y. (2013). Social engineering in social networking sites: Affect-based model. *8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*, 508-515.

[8] Smith, A., Papadaki, M., Furnell, S.M. (2013). Improving Awareness of Social Engineering Attacks. In: Dodge, R.C., Futcher, L. (eds) Information Assurance and Security Education and Training. WISE WISE WISE 2013 2011 2009. IFIP Advances in Information and Communication Technology, vol 406. Springer, Berlin, Heidelberg.

[9] Siponen, Mikko (2006). Information security standards focus on the existence of process, not its content. Commun. ACM. 49. 97-100. 10.1145/1145287.1145316.

[10] Kevin D. Mitnick, William L. Simon, Steve Wozniak, John Willey & Sons publisher (2002), – "The Art of Deception".

[11] M. Khonji, Y. Iraqi and A. Jones (2013). "Phishing Detection: A Literature Survey", IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2091-2121.

[12] J. Poderys, M. Artuso, C. M. O. Lensbl, H. L. Christiansen and J. Soler (2018). "Caching at the mobile edge: A practical implementation", *IEEE Access*, vol. 6, pp. 8630-8637, Feb.

[13] Ahmed Nafies Okasha Mohamed (2017). A New Heuristic Based Phishing Detection Approach Utilizing Selenium Web-driver UNIVERSITY OF TARTU.

[14] B. Wei, R. A. Hamad, L. Yang, X. He, H. Wang, B. Gao (2019). "A deep-learning-driven light-weight phishing detection sensor", *Sensors*, vol. 19, no. 19, pp. 4258.

[15] Y. A. Alsariera, V. E. Adeyemo, A. O. Balogun and A. K. Alazzawi (2020). "AI Meta-Learners and Extra-Trees Algorithm for the Detection of Phishing Websites," in IEEE Access, vol. 8, pp. 142532-142542.

[16] M. D. Bhagwat, P. H. Patil and T. S. Vishawanath (2021). "A Methodical Overview on Detection, Identification and Proactive Prevention of Phishing Websites," Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 1505-1508.

[17] Basit, A., Zafar, M., Liu (2021). A comprehensive survey of AI-enabled phishing attacks detection techniques. Telecommun Syst 76, pp. 139–154.

[18] T.O. Ojewumi, G.O. Ogunleye, B.O. Oguntunde, O. Folorunsho, S.G. Fashoto, N. Ogbu (2022). Performance evaluation of machine learning tools for detection of phishing attacks on web pages, Scientific African, Volume 16, e01165, ISSN 2468-2276

[19] JOURAU, Chizari Hassan, Zulkurnain Ahmad, Hamidy Ahmad, Husain Affandi (2015). Social Engineering Attack Mitigation, VL – 1, JO. International Journal of Mathematics and Computational Science, pp. 10-11.

[20] Bustard, J. D.et. al. (2013). Targeted Biometric Impersonation. International Workshop on Biometrics and Forensics (IWBF). 4-5 April. Lisbon, Portugal: IEEE, pp. 1-4.

[21] Haddadi, H. and P. Hui, P. 2010. To Add Or Not To Add: Privacy and Social Honeypots. IEEE International Conference on Communications Workshops (ICC). 23-27 May. Capetown, South Africa: IEEE, 1-5.

[22] M. M. Uddin, K. Arfatul Islam, M. Mamun, V. K. Tiwari, J. Park (2022). A Comparative Analysis of Machine Learning-Based Website Phishing Detection Using URL Information, 2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), pp. 220-224.