*Grenze International Journal of Engineering and Technology, June Issue*

# Social Engineering and its Importance

Yash Watarkar[1], Archis Gokhale[2], Zuhayr Khalid[3], Aditya Inamdar[4] and Prof. Vandana Jagtap[5]
Dr. Vishwanath Karad MIT World Peace University, Pune, India
Email: 1032202182@mitwpu.edu.in,{1032200598, 1032201162, 1032212206, vandana.jagtap}@mitwpu.edu.in

***Abstract* - The use of modern technology is inevitable in today's world. On one side of the coin, it has many advantages and it is proving to be beneficial in day-to-day life, making life easier. On the other hand, it is serving as a great platform for evil-minded people to threaten, attack and steal other people's data and identity. Hence, it gives rise to the term "Social Engineering". It is a form of psychological manipulation that is used to trick or deceive individuals in order to obtain sensitive information, gain unauthorized access to systems, or commit fraud. Social Engineering attacks are increasing day by day. The main reason is that people are still unaware about it and attackers are taking advantage of that fact. This paper focuses on the concept of social engineering and its importance in our life. In most of cases, reverse engineering works the best, that's why this paper also covers social engineering from the attacker's point of view, how the attackers attack and the methodology that they use. A basic Machine Learning model that covers most of the attacking techniques that attackers implement while at work.**

***Index Terms* - Social Engineering, Deception, Phishing, Baiting, Machine Learning**

## I. INTRODUCTION

The twenty-first century has quickly become a boon for businesses, media, social interaction, education, etc. with the help of numerous platforms over the Internet. As a result, the flow of information has increased colossally through the digital landscape and is getting more important day by day. People are relying more on digital devices and are finding ease and comfort along with it. The impact that these devices have on their life is extraordinarily amazing. With this newest technology, people can achieve almost everything. Although it appears that this technology has no side-effects, risks, it is not the case. There is a huge risk if people aren't educated on how to go about these newer technologies.Due to people's heavy reliance on digital technology, there has been an increase in cybercrime, including data breaches, malware, ransomware, and phishing-style attacks. Data is so important in today's world that many people can go to high extremities in order to acquire it. Social engineering is such a concept which has changing definitions till today. The methods used to trick people into disclosing private information or acting against their best interests are described in the aforementioned definitions. It heavily utilizes human psychology, which leads to winning someone's trust and then utterly manipulating them. According to this definition, humans are considered as a vulnerability to extract confidential information. This only happens when people are not aware of the risks that they might encounter while being engaged on social media platforms, surfing on the Internet, etc. It is of utmost importance that people become completely aware of this, thereby reducing the success rates of attackers. High levels of technical competence are not necessarily necessary for social engineering to be effective. Instead, social engineering preys on innate human psychological characteristics including avarice, politeness, gullibility, empathy, and curiosity. It has been established that many effective attacks occur as a result of digital users' ignorance. This paper explains the common social engineering techniques that attackers use to

*Grenze ID: 01.GIJET.9.2.431_1*

implement social attacks. It also makes use of the most common approach of reverse engineering to think from the point of view of an attacker and analyze the most common and frequent attacking patterns that make social engineering successful. With this, this paper aims to reduce the frequency and success rate of social engineering attacks by highlighting its importance in our life.

### *A. Need, Concerned Background And Motivation*

With the rise of digital technology and the increasing reliance on computer systems, there is huge scope for social engineering. From layman to professionals, almost everyone is falling prey to social engineering. Losing personal data and getting hacked by an"outsider" highly affects your life/psychology/businesses and what not! Little research has been done on social engineering in the potentially lucrative sector of social networks, despite increased awareness of this issue. First and foremost, social engineering is a serious problem in today's world because hackers use psychological deception techniques to access networks and sensitive data. It is crucial to comprehend the different kinds of social engineering attacks and how to defend against them. Second, social engineering is a fascinating and complicated topic that crosses several disciplines, including computer science, psychology, and sociology. One can learn more about the fundamental concepts and ideas governing human behavior and computer security by studying social engineering and writing about it.

### *B. Research Objective*

The main objective of this research paper is to spread awareness about social engineering and its importance, and analyze the main causes of social engineering attacks, thereby finding out the reasons behind the huge success rate of attackers by applying the concept of reverse engineering.

## II. LITERATURE REVIEW

| SR. NO. | RESEARCHERS | YEAR | METHODOLOGY | MAJOR CONTRIBUTION TO RESEARCH |
|---|---|---|---|---|
| 1 | Joseph M. Hatfield | 2018 | Descriptive | Conceptual array of contemporary connotations of SE |
| 2 | Aldawood and Skinner | 2018 | Descriptive | Increases user/employee awareness to reduce cyber security incidents |
| 3 | Shivam Lohani | 2018 | Descriptive | Impact of SE and ways to prevent its attacks |
| 4 | Salahdine and Kaabouch | 2019 | Descriptive | In-depth survey about SE attacks, classification, detection and prevention |
| 5 | Li, Wang and Horkoff | 2019 | Descriptive | Details critical research directions and discusses corresponding challenges of SE |
| 6 | Wang, Sun and Zhu | 2020 | Descriptive | Eliminates conceptual deficiencies by proposing compatible definitions of SEiCS. |
| 7 | Kathleen M. Carley | 2020 | Descriptive | Defines social cybersecurity as a new scientific discipline. |
| 8 | Jones et al. | 2020 | Descriptive | How vishing attacks employ persuasive techniques by social engineers |
| 9 | Odeh and Eleyan | 2021 | Descriptive | Methods and procedures for reducing SE attacks. |
| 10 | Syafitri et al. | 2022 | Descriptive | A systematic literature review to prevent SE attacks. |
| 11 | Biyani and Khan | 2020 | Conceptual | To detect spam content on social websites |
| 12 | Kangane et al. | 2022 | Conceptual | Detection of cyber-bullying attacks on social media |
| 13 | Xiaoyu Song | 2022 | Conceptual | Identifying abnormal user activity |
| 14 | Yan et al. | 2022 | Descriptive | Survey to show how ML is beneficial in cybersecurity |
| 15 | Wendy Durica | 2017 | Conceptual | Recognizing patterns in questioning |
| 16 | Charu Singh | 2020 | Conceptual | Detection of web links for phishing |
| 17 | Shelke and Achary | 2021 | Conceptual | An expert system for SE attack detection and defense |
| 18 | Ren et al. | 2022 | Conceptual | Machine Learning model for threat detection of general SE attacks |
| 19 | Lopez and Camargo | 2022 | Conceptual | ML classification algorithms to detect SE attacks |
| 20 | Alsufyani et al. | 2021 | Conceptual | Classifying text messages as phishing legitimate |

## III. Research Methodology

Cybersecurity is a concept which means a secure and safe group of networks. It has become increasingly important as there are an increasing number of digital users in today's world. Social Engineering attacks have many aspects and influencing factors. There are several ways in which one can defend efficiently against any SE (Social Engineering) attack.

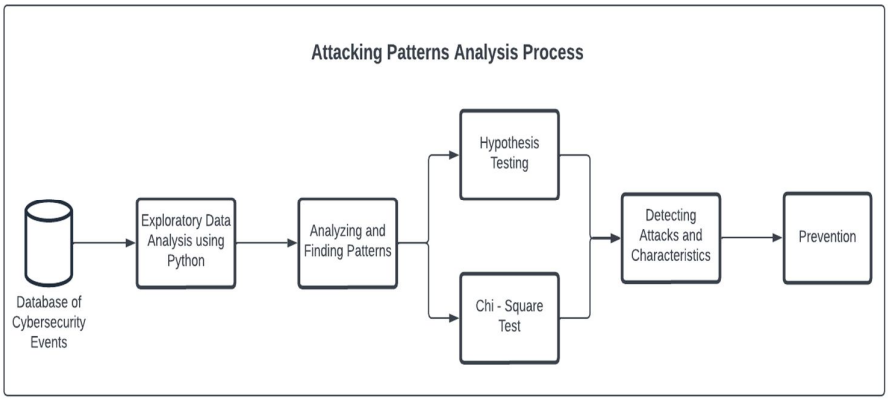


Fig 2. Attacking Patterns Analysis Process

A database consisting of the cybersecurity events of a company is taken into consideration. Features or attributes like attack types, attack categories, source and destination ports as well as source and destination IP addresses, duration of the attack, etc. form the database. EDA (Exploratory Data Analysis) is performed on the dataset. It gives an overview of the dataset, its features, and tuples, the type of data taken into consideration, etc. Python is chosen as the programming language to perform the analysis. After EDA, data is visualized using a variety of visualization approaches. The visualizations are thoroughly examined to discover specific redundant patterns that frequently recur. In order to be more specific and dive deeper into the details of those patterns, mathematical models like Hypothesis testing and Chi-square test are used. They help in determining the accurate pattern recognition task. A combination of the results of the two tests ultimately leads to the detection of attacking patterns. Using this we are able to prevent them by taking necessary precautions.

## IV. Problem Description

The objective of this experimental work is to analyze SE attacks from the attacker's point of view by applying the concept of reverse engineering. The ultimate goal is to reduce the success rate of attackers. Hence, this can be done either by educating the digital users and inculcating into them certain ways of precaution against these attacks or by analyzing the attacking patterns of attackers and working against them. This gives rise to certain specific questions which are listed below:

Q1. What types of attacks do attackers frequently carry out?
Q2. What specific patterns do the attackers follow while implementing an attack?
Q3. Why do digital users fall prey often to SE attacks?

Machine learning in cyber security is an efficient way that can help detect and respond to threats more quickly and accurately than traditional methods. Machine learning involves using algorithms to automatically identify patterns and anomalies in data, which can be applied to cybersecurity to detect and respond to threats. Machine Learning is used to detect potential threats and abnormalities, and analyze vast amounts of data from numerous sources, such as network logs, endpoint data, and threat intelligence feeds. Security personnel may be able to react to attacks more quickly and correctly as a result. ML (Machine Learning) techniques can be used to spot odd behavior on a network or system, including odd login habits or aberrant network traffic, which could point to a security breach. By analyzing user behavior and spotting deviations from the norm, ML can be used to spot potential insider threats or compromised accounts. Even if malware has never been seen before, it can be found using file analysis and ML. By analyzing vast volumes of data and finding patterns of suspect behavior, ML can be used to identify fraudulent conduct, such as credit card fraud or identity theft.

*A. Hypothesis Testing*

A statistical hypothesis test is a technique for determining if the available data are sufficient to support a specific hypothesis.

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

1. We reject the null hypothesis $H_0$ and confirm that the observed difference is statistically significant if the p-value is smaller than our significance level ($p<\boldsymbol{\alpha}$).
2. If the p-value exceeds our threshold for statistical significance ($p>\boldsymbol{\alpha}$), we must maintain H0 and draw the conclusion that the observed difference is not statistically significant.

*B. Chi- Square Test*

The following describes the chi-square test's null hypothesis:
*$H_0$: The destination port has no bearing on the attack category.*
It makes sense that "independence between two variables" refers to the fact that the distribution of values for one variable doesn't change when the value of the other one does (and vice versa).

*C. Pearson's Correlation*

The linear associations between two variables are assessed. A weak or nonexistent linear relationship between the variables exists if the value is close to 0.

*D. Spearman's Correlation*

The monotonic relationships between two variables are assessed. A weak or nonexistent monotonic relationship between the variables exists if the value is near to 0.

## V. Analysis Discussion

The analysis conducted in this research paper showed the common trends and attacking patterns that an attacker follows based on a database of cybersecurity events of a certain company. The results showed the most targeted destination IP Address, the most logical ports attacked, the most frequently implemented / common type of attack.

*A. Findings*

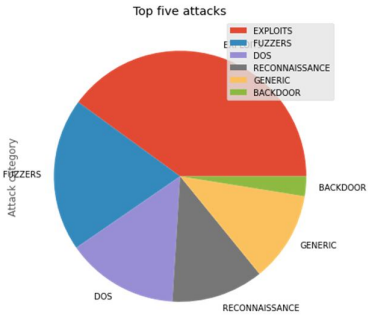


Fig 1. Top Five Attacks

Out of a vast variety of attack categories, a specific number of attack categories were chosen according to the collected data. The majority of attacks are direct exploits. Those include all the SE attacking techniques. Fuzzers, generic attacks, and reconnaissance come next. A denial of service (DoS) attack aims to block access to a website, server, or network by its intended users by saturating it with traffic or other types of data.
Despite the fact that there are more than 170,000 records of cybersecurity assaults, 10 servers with IPv4 addresses between [149.171.126.10 and 149.171.126.19] are the target of these attacks. This indicates a non-accidental and deliberate series of actions, as the company was assaulted on various occasions over the observed time period on a particular subnet.

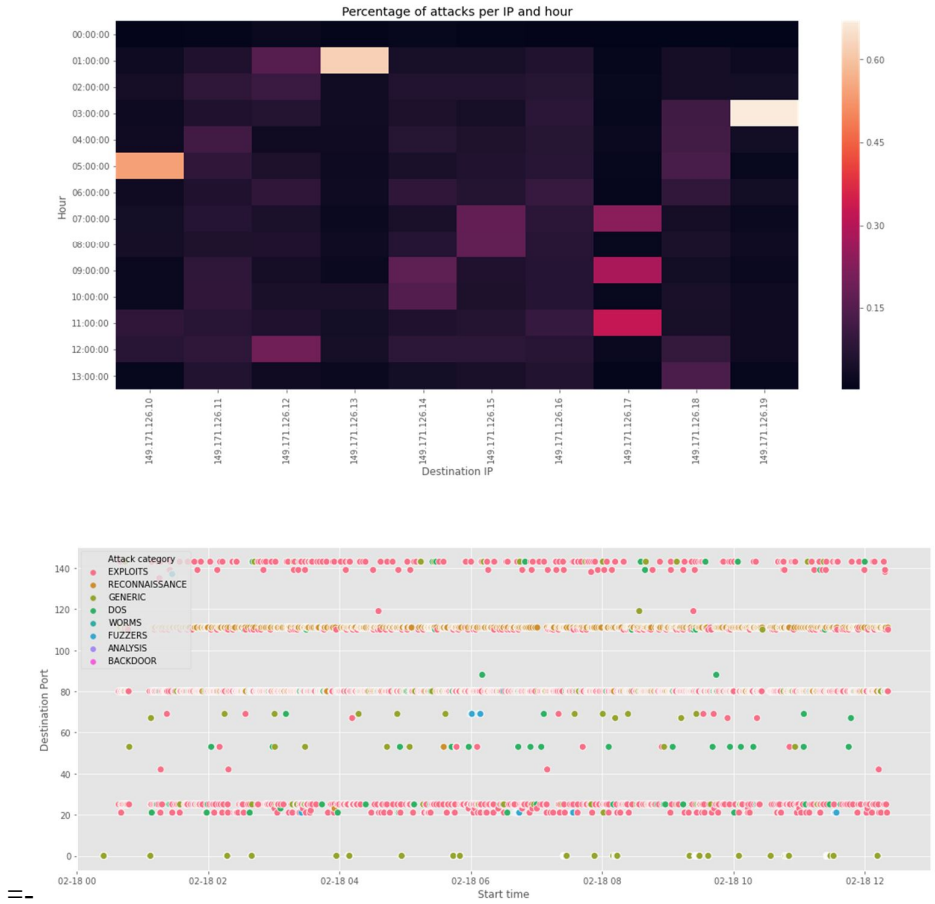


Fig 3. Most Attacked Logical Ports

As seen in the above figure, a range of ports have been attacked with a specific type of attack. Port 0 is likely attacked by generic attacks. Ports in the 20s and 140s are dominated with exploits. Combinations of reconnaissance, worm assaults, and denial-of-service attacks target ports between 100 and 120. Ports from 40 to 80 experience a variety of attacks including worms, fuzzers, and backdoors.

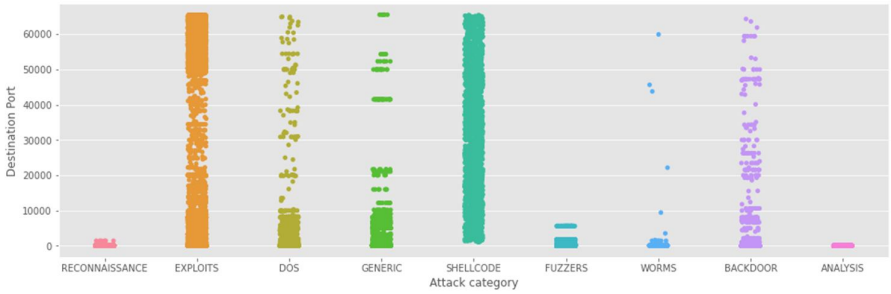


Fig 4. Attack Categories w.r.t. Ports

The graph exhibits a particular pattern, where attacks typically take place on ports between 10,000 and 50,000, with the exception of shellcode-type attacks, as predicted by the hypothesis test.

## VI. Contributions And Implications Of The Research

The research conducted in this research paper will be beneficial for the analysis and in-depth study of hackers, their techniques and how they pose a threat to the digital users of today. By reverse engineering we put ourselves into the shoes of an attacker, thereby being a step ahead of the attacker. Hence, we will be able to plan the precautions and preventive measures beforehand and way ahead before an attacker decides to implement an attack. The study conducted above will also help in creating awareness about Social Engineering among people, thus creating a society with educated people and minimizing the potential risks and threats they face everyday.

## VII. Conclusion

This paper has discussed the importance and gravity of Social engineering in an individual's life. It has highlighted the need to know the scientific discipline of Social engineering and be aware of that. It has also discussed the various types of SE attacks that have been/are implemented by attackers. A general insight was explained through which an individual can become aware of various SE attacks. There are some common and easy ways in which one can defend against these attacks, which were discussed in this paper. As previously mentioned, social cybersecurity is a young field of science and engineering. This work's initial interdisciplinary focus is giving way to transdisciplinary inquiry. This paper also put light on SE attacks from an attacker's perspective. The way an attacker executes an attack and the pattern that the attacker follows is worth analyzing as it helps individuals pre-plan their defense strategies and hence be in a safe environment away from any threats by completely mitigating them. An ML model was implemented and performed to analyze the attacking patterns of several cyber security events. The model was fed with a company's data of cyber security events which met the current market standards. It was an apt example to demonstrate the attacking patterns and redundant use of similar attacks by the attacker. Hence, the process of reverse engineering was also applied. Finally, using hypothesis test and chi-square test various results related to the attacker's behavior were inferred. The results included the most targeted destination IP addresses, most logical ports which were likely to get attacked, the most frequent type of attack, and at what time of the day was the attacker most active. A comprehensive analysis of all these factors led to the conclusion of this research paper thereby spreading the importance of social engineering.

## VIII. Limitations Of The Study And Future Scope

The study conducted in this research paper is limited to the attacking patterns. A more detailed analysis can be done on the user-centric frameworks which evaluate the user behavior. Social engineering is a constantly evolving field that adapts to new technological advancements and changes in society. The future scope of social engineering is wide-ranging and includes a variety of potential developments and trends. Social engineering attacks could become more sophisticated and automated with the use of AI and machine learning technologies. Attackers could use these technologies to craft more convincing phishing emails, create more realistic social media profiles, and develop more effective attack strategies.

## References


[1] Joseph M. Hatfield (2018), Social engineering in cybersecurity: The evolution of a concept https://doi.org/10.1016/j.cose.2017.10.008

[2] H. Aldawood and G. Skinner, "Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review," 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Wollongong, NSW, Australia, 2018, pp. 62-68, doi: 10.1109/TALE.2018.8615162

[3] Lohani, Shivam, Social Engineering: Hacking into Humans (February 5, 2019). International Journal of Advanced Studies of Scientific Research, Vol. 4, No. 1, 2019, Available at SSRN: https://ssrn.com/abstract=3329391

[4] Salahdine, F., & Kaabouch, N. (2019). Social Engineering Attacks: A Survey. Future Internet, 11, 89.

[5] T. Li, K. Wang and J. Horkoff, "Towards Effective Assessment for Social Engineering Attacks," 2019 IEEE 27th International Requirements Engineering Conference (RE), Jeju, Korea (South), 2019, pp. 392-397, doi: 10.1109/RE.2019.00051

[6] Wang, Zuoguang & Sun, Limin & Zhu, Hongsong. (2020). Defining Social Engineering in Cybersecurity. IEEE Access. 8. 85094-85115. 10.1109/ACCESS.2020.2992807

[7] Carley, K.M. Social cybersecurity: an emerging science. Comput Math Organ Theory 26, 365–381 (2020). https://doi.org/10.1007/s10588-020-09322-9

[8] Jones, K.S., Armstrong, M.E., Tornblad, M.K. and Siami Namin, A. (2021), "How social engineers use persuasion principles during vishing attacks", Information and Computer Security, Vol. 29 No. 2, pp. 314-331.

https://doi.org/10.1108/ICS-07-2020-0113
[9] Odeh, Noor & Eleyan, Derar & Eleyan, Amna. (2021). A SURVEY OF SOCIAL ENGINEERING ATTACKS: DETECTION AND PREVENTION TOOLS
[10] W. Syafitri, Z. Shukur, U. A. Mokhtar, R. Sulaiman and M. A. Ibrahim, "Social Engineering Attacks Prevention: A Systematic Literature Review," in IEEE Access, vol. 10, pp. 39325-39343, 2022, doi: 10.1109/ACCESS.2022.3162594
[11] McNealy, J. E. (2022). Platforms as phish farms: Deceptive social engineering at scale. New Media & Society, 24(7), 1677–1694. https://doi.org/10.1177/14614448221099228
[12] Biyani Yogita, Dr. Rahat Afreen Khan (2020). Spam Detection in Social Media using Machine Learning Algorithm. International Journal for Research in Applied Science and Engineering Technology. 9. 432-439. 10.22214/ijraset.2021.32832
[13] Kangane, Sara. (2022). Detection of Cyberbullying on Social Media Using Machine Learning. International Journal for Research in Applied Science and Engineering Technology. 10. 1530-1535. 10.22214/ijraset.2022.44094
[14] Salina Adinarayana, E.Ilavarasan, (2022). A Machine Learning Approach to Extract Opinions from Social Media Content,https://doi.org/10.14419/ijet.v7i4.5.20080
[15] X. Song, "Prediction of People's Abnormal Behaviors Based on Machine Learning Algorithms," 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Guangzhou, China, 2022, pp. 406-409, doi: 10.1109/MLISE57402.2022.00087
[16] Yan, F, Wen, S, Nepal, S, Paris, C, Xiang, Y. Explainable machine learning in cybersecurity: A survey. Int J Intell Syst. 2022; 37: 12305- 12334. doi:10.1002/int.23088
[17] Giovanni Apruzzese, Pavel Laskov, Edgardo Montes de Oca, Wissam Mallouli, Luis Brdalo Rapa, Athanasios Vasileios Grammatopoulos, and Fabio Di Franco. 2023. The Role of Machine Learning in Cybersecurity. Digital Threats 4, 1, Article 8 (March 2023), 38 pages. https://doi.org/10.1145/3545574
[18] D. Roshni Thanka, G. Jaspher, W. Kathrine, E. Bijolin Edwin. (2023). Using Machine Learning for Cyber Security, https://doi.org/10.1201/9781003187158
[19] Lorena Montoya, Wolter Pieters, Marianne Junger, Pieter Hartel, "On the anatomy of social engineering attacks—A literature-based dissection of successful attacks". July 2017 https://doi.org/10.1002/jip.1482