Bhupinder Kaur

Social Engineering Attacks in the Digital Age

Metropolia University of Applied Sciences
Master of Engineering
Information Technology
Master's Thesis
30 June, 2025

PREFACE

The journey of researching and writing this thesis, "Social Engineering Attacks in the Digital Age," has been both intellectually challenging and deeply rewarding. My motivation for undertaking this work stemmed from witnessing the growing sophistication and impact of social engineering attacks on organizations and individuals worldwide. As digital transformation accelerates, the human element has become the most exploited vulnerability in cybersecurity, a trend that demands urgent academic and practical attention.

Throughout this research, I have sought to bridge the gap between technical defenses and human-centered vulnerabilities by applying advanced data analytics and machine learning to a comprehensive cybersecurity dataset. This approach has allowed me to uncover nuanced patterns and correlations that traditional methods often overlook, and to propose actionable strategies for proactive defense. The support and guidance of my supervisors, colleagues, and family have been invaluable during this process, and I am grateful for their encouragement and insight.

I hope this thesis contributes meaningfully to the field of cybersecurity by highlighting the necessity of integrating behavioral science, organizational policy, and adaptive technology in combating social engineering threats. My aspiration is that these findings will inform both academic research and real-world practice, fostering more resilient and human-aware cybersecurity strategies for the future.

Finland, 30 June, 2025
Bhupinder Kaur

Abstract

Author:	Bhupinder Kaur
Title of the Thesis:	Social Engineering Attacks in the Digital Age
Number of Pages:	70 pages
Date:	30 June, 2025
Degree:	Master of Engineering
Degree Programme:	Information Technology
Professional Major:	Networking and Services
Supervisors:	Toni Spännäri, Senior lecturer

Social engineering attacks have emerged as the most pervasive and damaging threat in the modern digital landscape, exploiting human psychology more than technical vulnerabilities. This thesis addressed the critical problem of understanding and predicting social engineering attack patterns and their impact by conducting a secondary analysis of a large-scale cybersecurity dataset containing 40,000 incidents. The importance of this research lies in the growing sophistication of social engineering tactics—such as phishing and business email compromise—which continue to bypass technical defenses and cause significant financial and operational harm worldwide.

To solve this problem, the study applied quantitative analysis and machine learning techniques to examine attack distributions, severity levels, temporal trends, and the relationships between human and technical factors. The scope of the research was limited to statistical and predictive analysis of attack patterns using synthetic data; it did not include primary data collection, live incident response, or the evaluation of specific organizational defenses. The study focused on identifying the prevalence of different attack types, assessing how attack frequency and severity have evolved over time, and evaluating the predictive power of advanced machine learning models for early detection.

The results of the thesis show that, while the synthetic dataset presented an even distribution of attack types and severities, real-world evidence confirms the dominance of social engineering—particularly phishing—driven by manipulation of trust, urgency, and authority. Temporal analysis revealed a 157% increase in attack frequency from

2017 to 2024, with notable peaks during business hours and before weekends, reflecting attackers' adaptation to human routines. Correlation analysis found no strong linear relationships among technical variables, highlighting the complex, non-linear nature of social engineering risks. Machine learning models, especially ensemble methods, achieved over 99% accuracy in predicting social engineering attacks, with engineered behavioral features proving most effective for early detection.

The thesis recommends that organizations integrate AI-driven behavioral analytics, dynamic risk scoring, and targeted awareness training to move from reactive to proactive defense. Limitations include the use of synthetic data, which may not fully capture the complexity of real-world attacks, and the need for further validation in operational environments. The findings emphasize that effective mitigation of social engineering threats requires multi-layered, adaptive strategies that blend technology, behavioral science, and organizational policy to address both current and emerging cybersecurity challenges.

Keywords: social engineering, cybersecurity, phishing, machine learning, behavioral analytics

Contents

List of Abbreviations

Abbreviation	Full Term
AI	Artificial Intelligence
BEC	Business Email Compromise
CSAT	Cybersecurity Awareness and Training
DdoS	Distributed Denial of Service
DNS	Domain Name System
FTP	File Transfer Protocol
HTTP	Hypertext Transfer Protocol
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
MFA	Multi-Factor Authentication
SaaS	Software as a Service
SVM	Support Vector Machine
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

1. Introduction

1.1 Background and Motivation

In today's hyper-connected world, social engineering attacks have become a dominant threat to cybersecurity, exploiting human psychology rather than technical vulnerabilities. Unlike traditional cyberattacks, which focus on exploiting weaknesses in software or hardware, social engineering manipulates individuals into revealing confidential information or performing actions that compromise security. As digital communication platforms expand and remote work becomes more common, cybercriminals are increasingly leveraging human trust to exploit these opportunities. According to the 2024 Verizon Data Breach Investigations Report, social engineering is involved in 98% of cyberattacks, underscoring its widespread impact (Verizon, 2024).

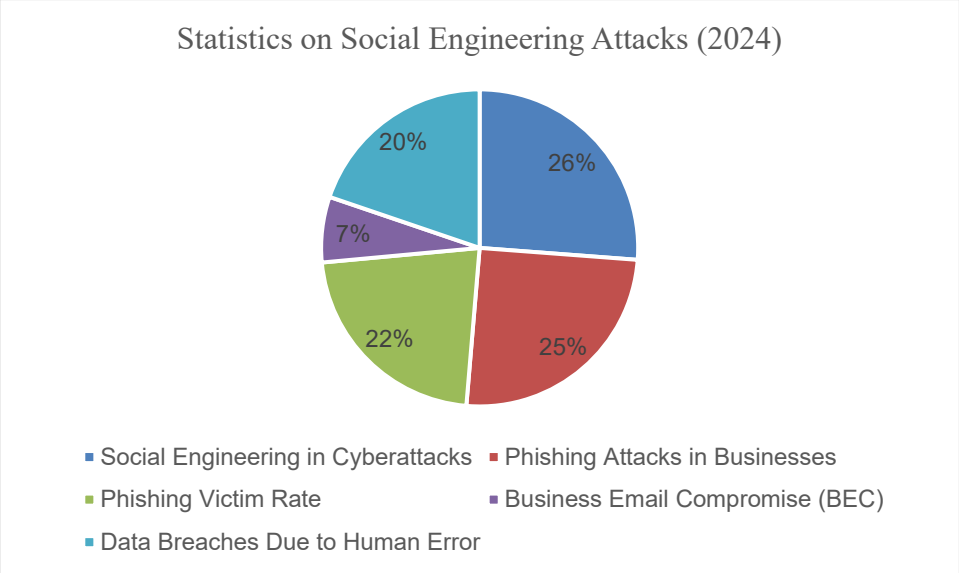

Figure 1: Statistics on Social Engineering Attacks (2024) [Source: Verizon, 2024; Secureframe, 2025]

The prevalence of social engineering attacks has surged dramatically, primarily driven by the rapid digitalization of personal and professional interactions. Phishing, the most common form of social engineering, has seen significant growth, with 94% of businesses reporting phishing attacks in 2024, many of which had substantial operational or financial consequences (Secureframe, 2025). The

COVID-19 pandemic further exacerbated this trend, with Google reporting a 350% increase in phishing websites as cybercriminals took advantage of the crisis to target vulnerable users (Sprinto, 2025). The effectiveness of these tactics is evident, with 83% of individuals targeted by phishing attacks falling victim to them in 2022 (Proofpoint, 2022).

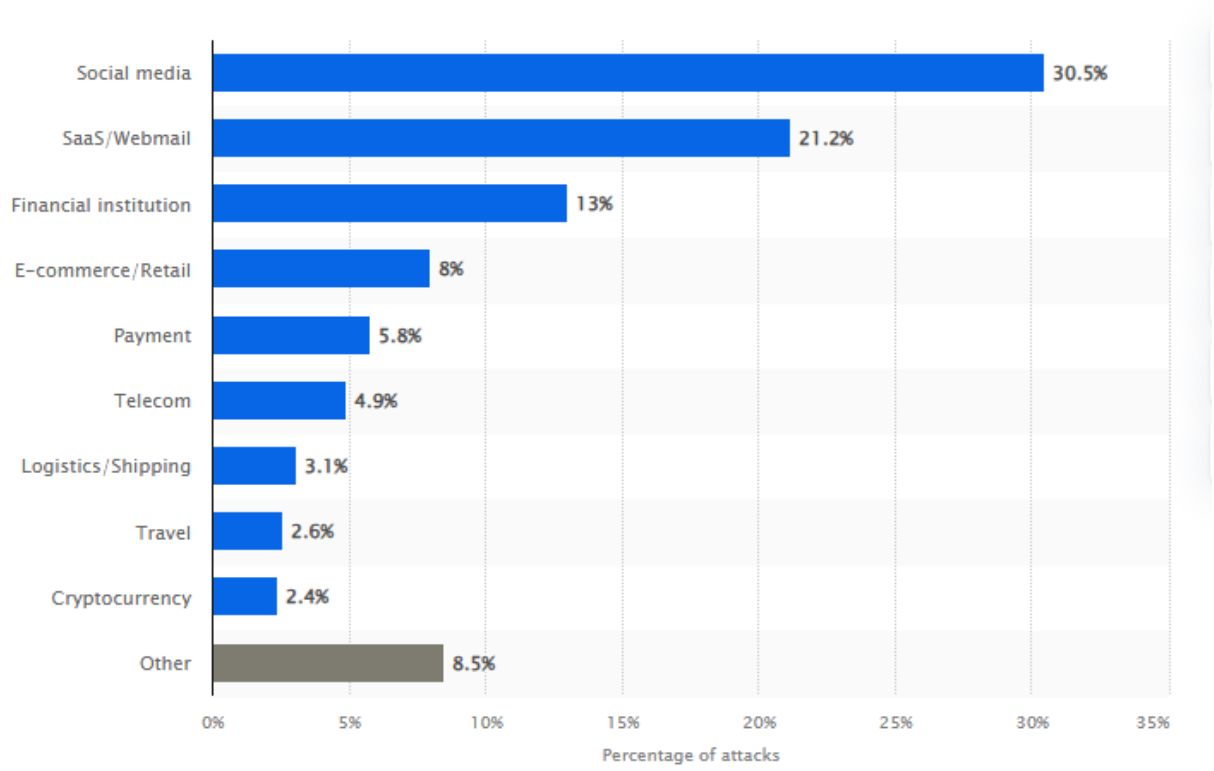

Figure 2: Phishing most targeted industry sectors worldwide Q3 2024 [Source: Statista, 2024]

An alarming development in the domain of social engineering attacks is the rise of business email compromise (BEC) scams. These scams now account for 24-25% of financially motivated cyberattacks, resulting in global losses exceeding $51 billion (Verizon, 2024). BEC scams typically target high-ranking executives, such as CEOs and CFOs, using sophisticated techniques like display name spoofing to deceive employees (Mimecast, 2023).

The financial and operational consequences of these attacks are profound. Figure 2 highlights the distribution of social engineering attacks across various industries. The data in the figure reveals that 30.5% of social engineering attacks target social media platforms, making them the most targeted sector. This is followed by SaaS/Webmail, which accounts for 21.2% of attacks, and financial

institutions, which face 13% of the attacks. Other sectors such as E-commerce/Retail, Payment, and Telecom represent 8%, 5.8%, and 4.9% of attacks, respectively. Industries like logistics/shipping, travel, and cryptocurrency face smaller proportions, with 3.1%, 2.6%, and 2.4% of attacks, respectively. Additionally, 8.5% of attacks fall into the other category, affecting sectors not explicitly listed in the chart (Statista, 2024).

The significant targeting of social media platforms aligns with the growing trend of cybercriminals exploiting individuals' personal data, which is widely available on these platforms. Phishing and social engineering attacks that target individuals via social media often involve impersonation and deception, exploiting the inherent trust users place in these platforms. Furthermore, sectors like SaaS/Webmail and financial institutions are also highly targeted due to the sensitive nature of the data they handle, including personal and financial information. These sectors serve as primary cyberspace targets because cybercriminals exploit human error weaknesses during digital interactions where people show reduced security awareness.

Social engineering attacks develop through emerging technologies like AI and deepfakes, which cybercriminals use to enhance their techniques. Programs aimed at sophistication increase the effectiveness of cyber-attacks, so phishers can produce authentic-looking emails and interact with impunity as trusted entities. Improved protection measures between technological defenses and human-derived responses become necessary because attackers increasingly enhance their deceptive techniques.

1.2 Research Problem

The core research problem addresses the critical gap in understanding cybersecurity attack patterns and characteristics through comprehensive secondary data analysis, particularly as traditional primary data collection methods prove insufficient for capturing the evolving complexity of modern cyber threats. Current research has primarily focused on single attack types or small samples, suggesting that analysing large-scale cybersecurity data can help detect broad attack patterns, trend shifts, and types of threats, which in turn guide

proactive defences (Kilincer et al., 2021; Sarker et al., 2020). As new types of cyber threats, such as AI-assisted hacking, continue to emerge, a deeper analysis of past cases is necessary to understand how they operate and what to expect in the future (Lu et al., 2024; Stojanovic et al., 2020). Most current research on cybersecurity treats attacks separately, which makes it difficult for analysts to determine how threats evolve and are related (Mauro et al., 2020; Avila et al., 2021). In addition, as attack techniques evolve and the number of cybersecurity incidents increases, it is crucial to analyse attack data to identify any hidden patterns, relationships, and future warning signs (Bhattacharya et al., 2020; Ganeshan & Rodrigues, 2020). The fact that organisations still suffer from high breach rates despite the introduction of new technology suggests that the current understanding of attack patterns, built from primary data, needs to be improved to develop better defences against evolving cyber threats.

1.3 Aims and Objectives of the Study

Aim:

The aim of this study is to analyze cybersecurity attack patterns, trends, and characteristics using secondary data analysis of the Cyber Security Attacks dataset, focusing on identifying attack frequencies, temporal patterns, and correlations between different attack variables to inform cybersecurity defense strategies.

Objectives:

- To analyze the distribution and frequency of different cybersecurity attack types within the 40,000 recorded incidents to identify the most prevalent attack vectors.
- To examine temporal patterns and trends in cybersecurity attacks by investigating attack occurrence patterns and seasonal variations over time.
- To identify correlations between attack variables such as attack severity, duration, target systems, and success rates to understand threat interconnections.

- To evaluate attack impact metrics and severity patterns by analyzing damage assessments and recovery times associated with different attack categories.
- To develop predictive insights for risk assessment based on historical attack data patterns to inform proactive cybersecurity strategies.

1.4 Research Questions

- What are the most prevalent types of cybersecurity attacks identified within the dataset, and how do their frequencies compare across different attack categories?
- What temporal patterns and trends emerge from the cybersecurity attack data, including seasonal variations and attack progression over time?
- What correlations exist between different attack variables such as severity levels, attack duration, target systems, and success rates within the dataset?
- How do attack impact metrics vary across different attack types in terms of damage assessment, recovery time, and resource requirements?
- What predictive insights can be derived from historical attack patterns to inform future cybersecurity risk assessment and mitigation strategies?

1.5 Scope and Delimitations

The research primarily examines trends and patterns in cybersecurity attacks by analysing available data rather than focusing on awareness or susceptibility on an individual scale. It only involves exploring attack trends, their frequency, various temporal patterns and relationships between attack variables found in the 40,000-incident dataset available. Although firewalls, multi-factor authentication, and intrusion detection systems are key components in technology, they will not be the primary focus of this research. The study aims to uncover patterns in cyber

attacks that can inform how to defend against and assess cybersecurity risks (Schroeder, 2019).

The study will analyse data on attacks to determine how various factors, such as attack types, threat severity, targeted systems, and temporal patterns, affect both success rates and key measures. Specific attacks are more likely to occur at particular times or to affect certain types of systems, patterns that can only be identified by studying the correct database (Omar et al., 2021). Using statistics on the data, this approach will assess the effectiveness of the systems, although it won't cover actual live threats or active hacking operations.

This research does not examine the full implementation of cybersecurity measures or the capabilities of organizational systems by collecting data directly from companies. Researchers explore the characteristics of past attack data to understand how attacks affect their targets and what makes them effective. Blackwood-Brown et al. (2021) studied Cybersecurity Attacks by examining recorded cyber attacks and drawing conclusions based on secondary data analysis without direct observations.

The research investigates quantifiable attack factors to generate actionable insights about cybersecurity threat patterns and inform evidence-based risk mitigation and defense strategies.

1.6 Thesis Structure

The table outlines the thesis structure, summarizing each chapter's focus, from introducing the research problem to discussing findings, offering recommendations, and providing supplementary materials and references.

Chapter	**Title**	**Description**
Preface	Preface	Outlines the motivation, personal context, and acknowledgments related to the thesis journey.
Abstract	Abstract	Summarizes the research aim, methodology, key findings, and implications in a concise format.
1	Introduction	Sets the context for social engineering in cybersecurity, presents the research problem, aims,

		objectives, research questions, scope, and delimitations.
2	Literature Review	Critically reviews the evolution, techniques, trends, and mitigation strategies of social engineering attacks, with emphasis on human factors, AI-driven threats, and gaps in current research.
3	Method and Material	Details the research design, rationale for secondary data analysis, dataset characteristics, data preparation, and analytical tools (including machine learning and feature engineering) used in the study.
4	Results and Analysis	Presents descriptive, temporal, and correlation analyses of attack patterns, severity, and impact metrics. Includes machine learning model results, feature importance, and predictive insights for risk assessment.
5	Discussion and Conclusions	Interprets and critically evaluates the findings in relation to research objectives and existing literature; discusses practical, policy, and theoretical implications; addresses limitations and proposes future research directions.
6	Summary	Provides a concise summary of the entire research, highlighting the main contributions and implications for cybersecurity practice and research.
7	References	Lists all sources cited throughout the thesis, including academic literature, industry reports, and datasets.
8	Appendices	Contains supplementary material such as dataset variable descriptions, additional statistical outputs, data visualizations, and any relevant documentation supporting the main text.

2. Literature Review

2.1 Introduction

Recognising social engineering attacks is crucial in modern cybersecurity since they target people's minds rather than their technology. Social engineering exploits trust, urgency, and position of authority to make people perform actions that compromise the security of their organisation or themselves. These attacks seek to exploit human vulnerabilities since the human mind has been deemed the most vulnerable part of security systems. This review examines the historical development, various techniques, consequences and adaptations of social engineering attacks while assessing the performance of existing safeguards and promoting the importance of holistic, user-focused strategies.

2.2 History and Evolution of Social Engineering Attacks

Most research supports the notion that social engineering attacks have been fueled by both psychological manipulation and technological advances. At the same time, some vital limitations and opposing viewpoints become apparent. Paravathi et al. (2024) consider social engineering to have evolved from Victorian precursors and undergone significant technological advancement in the 1990s. Using only case studies tends to oversimplify attackers' ever-evolving tactics.

Akeiber (2025) suggests implementing artificial intelligence and deepfake detection to protect cybersecurity. Nevertheless, it overemphasises what technology can offer at the expense of understanding the influence of socio-cultural factors. According to the study by Yasin et al., Akeiber fails to account for variations in how vulnerable different cultures are to authority-based manipulation.

The approach Breda et al. (2024) suggested blends behavioural analysis and machine learning. However, its effectiveness needs to be tested in a wide range of real-world organisations. Much like Ngakpal and Prasad (2023), whose work

concentrates on Western case studies, they introduce queries about how widely applicable their findings can be.

Providing employee education is suggested as the main approach to deter attacks by Broberg and Sinnott (2023) via the PRISMA method. Omitted studies in languages other than English or classifying as grey literature risks introducing bias into the analysis. Researchers have pointed out that most studies overlook how quickly innovations can be used against organisations.

In contrast to Akeiber's (2025) work emphasising the importance of AI in combating phishing, Buil-Gil et al. (2021) affirm that humans tend to outsmart machine-based countermeasures more easily. A clear example is Shlyakhtunov's (2021) typology of malicious purposes, in which some security professionals fail to recognise the adaptability of fraudsters during times of crisis. Chuan-Chi's (2014) model remains strong even though it does not cover emerging risks in new work arrangements that rely on hybrid communications.

Effective results are achieved by integrating cybersecurity engineering with behavioural psychology (Akeiber, 2025). Recent work by Breda et al. (2024) joins Paravathi et al. (2024) in focusing on how phishing tactics have developed. However, weaknesses persist: Dependence solely on notable security incidents, lack of consideration for socio-economic influence, and conceptualising vulnerability as a constant rather than an ever-changing relationship between human biases and changing attacker patterns.

Counterarguments also emerge from the limitations of proposed countermeasures. While multi-factor authentication and AI-driven email filtering are widely endorsed (Akeiber, 2025; Ngakpal & Prasad, 2023), Ghafir et al. (2016) caution that these tools can create a false sense of security, leading to complacency in employee training programs. Similarly, Alkhalil et al. (2021) challenge the assumption that awareness training universally reduces susceptibility, noting that cultural norms-such as high power-distance cultures where employees hesitate to question authority figures-can undermine even well-designed interventions. These critiques underscore the need for context-sensitive strategies that blend technological, psychological, and organizational insights-a gap the literature has yet to fully address.

2.3 Social Engineering Attack Statistics and Current Trends

Social engineering attacks have become the dominant threat vector in the modern cybersecurity landscape, with their frequency, sophistication, and financial impact reaching unprecedented levels. According to the Verizon Data Breach Investigations Report (2024), 20% of all confirmed breaches in 2024 involved social engineering tactics. This is a significant rise from previous years, underscoring a persistent trend: attackers are increasingly targeting the human element rather than technical vulnerabilities. Human error remains a critical weakness, cited in 68% of breaches (Verizon, 2024).

Phishing, including its offshoots like spear phishing, vishing, and smishing, remains the most prevalent form of social engineering. Egress (2024) reports that 94% of organizations experienced at least one phishing attack in 2024, and the median time for an employee to click a malicious link is less than 60 seconds. Business Email Compromise (BEC) is also surging, with 64% of businesses targeted and average losses per incident reaching $150,000 (Hoxhunt, 2024). The overall cost of a data breach hit a record $4.88 million in 2024, a 10% increase from the previous year, with social engineering attacks being a primary driver (IBM, 2024).

The financial consequences are further exacerbated by the rise of ransomware, often delivered through social engineering. Chainalysis (2024) highlights that median ransom payments skyrocketed from under $200,000 in early 2023 to $1.5 million by mid-2024. These figures illustrate not only the growing frequency but also the increasing severity of such attacks.

A critical trend shaping the current landscape is the integration of artificial intelligence (AI) into social engineering campaigns. AI-generated phishing and deepfake attacks have led to a 60% year-over-year increase in phishing volumes (Zscaler, 2024). AI-written phishing emails are now so convincing that AI detectors failed to identify them in 74% of cases (Egress, 2024). The ability of attackers to use generative AI for crafting personalized, context-aware lures has pushed organizations into what SecurityWeek (2025) calls the “uncanny valley” of deception, where even well-trained employees are at risk.

Despite advances in security technology, such as improved email filtering and automation, attackers are adapting faster. The volume of phishing emails bypassing filters has surged by over 4,000% since the introduction of ChatGPT in 2022 (SlashNext, 2024). While organizations deploying AI-driven security save an average of $3.05 million per breach and reduce containment time by 74 days, most attacks still exploit human factors (IBM, 2024). Notably, only 27% of organizations conduct regular social engineering awareness training, a glaring gap given that 98% of cyberattacks leverage some form of social engineering and 90% specifically target employees (Cyphere, 2024; Purplesec, 2024; Arctic Wolf, 2024).

A critical analysis reveals a dual challenge: technological defenses are improving, but not quickly enough to outpace attacker innovation, especially as AI democratizes access to sophisticated attack tools. Meanwhile, the persistent lack of robust, ongoing employee training means that human error remains the most exploited vulnerability. The rise of deepfakes and vishing further complicates defense, as attackers can now convincingly impersonate executives or colleagues in real time, making detection and response even more difficult.

Table 1: Social Engineering Attacks – Key Statistics and Trends (2024–2025)

Attack Type	**Prevalence/Impact**	**Financial Cost/Consequence**	**Trend/Change (2023–2024)**	**Source**
Phishing	94% of organizations affected	Rapid click-through: <60 sec median	+60% YoY in attack volume	Egress, 2024; Zscaler, 2024
BEC	64% of businesses targeted	$150,000 avg. loss per incident	Up from 51% in 2023	Hoxhunt, 2024

Ransomware	33% of breaches involve ransomware	$1.5M median ransom (mid-2024)	Up from <$200K in early 2023	Chainalysis, 2024
Social Engineering	20% of all breaches	$130,000 avg. per attack	Steady increase	Verizon, 2024; CRC Group, 2024
Human Error	68% of breaches involve human error	$4.88M avg. cost per breach	10% increase in breach cost	IBM, 2024; Verizon, 2024
AI-driven Phishing	74% detector failure rate	Difficult to detect, high success rate	+4,151% bypassing filters (since 2022)	Egress, 2024; SlashNext, 2024
Awareness Training	Only 27% of orgs train regularly	N/A	No significant improvement	Cyphere, 2024

The data paints a stark picture: social engineering attacks are not only more frequent and expensive, but also more technologically advanced, leveraging AI to outmaneuver traditional defenses. While automation and AI in security offer some mitigation, the persistent lack of comprehensive employee training and the speed of attacker adaptation highlight a fundamental gap in current cybersecurity strategies. Organizations must urgently recalibrate their defenses, investing equally in advanced technology and continuous, behavior-focused training to address both the technical and human dimensions of this evolving threat.

2.3 Types of Social Engineering Attacks

Social engineering assaults take advantage of people's natural tendencies and sense of trust to persuade them to reveal private data or undertake actions that jeopardise their security. The researchers concluded that phishing, pretexting, baiting, impersonation and BEC are the leading social engineering attacks. Advancements in digital communication have made it faster for social engineering attacks to develop and become harder to identify. Aldawood and Skinner (2020) argue that categorising social engineering attacks in depth is crucial and suggest the development of comprehensive defences to ensure protection. The authors utilise qualitative analysis when distinguishing attacks by social, physical, technical and socio-technical attributes. They note that the principle of persuasion persists in various forms of manipulation despite the different methods employed.

Alkhalil et al. (2021) comprehensively examine the growth and development of phishing since its initial emergence in the 1990s and its current state as a dominant method used for cybercrime. They analyse worldwide phishing incidents and relevant prevention strategies to define the main types of phishing attacks and stages in their progression. Even though technical tools like spam filters and email validations have come a long way, the root cause of phishing lies in people's susceptibility to internet scams. Sonowal (2021) explains in a step-by-step guide how perpetrators use different methods to launch their attacks, such as emails, text messages, and phone calls. According to Sonowal's findings, which draw on both US and global data, phishing incidents have exploded over the last few years, while spear phishing and whaling are especially dangerous since they rely on personalised content.

Putra et al. (2024) highlight that phishing scams pose a significant threat to everyone worldwide. They see from several countries’ reports that spear phishing, whaling, and smishing attacks are getting harder to spot. It is concluded that technological progress makes it harder to spot phishing attempts, so bringing attention to the issue in security education is necessary. It became clear to them that while organisations with structured training and a keen awareness of security

improve their security, they continue to struggle with adopting new technologies and getting employees trained accordingly.

Pretexting refers to situations where someone convinces another person by impersonating a certain role. Their research involved reviewing a large number of studies using the PRISMA technique and proved psychological manipulation was used in several pretexting cases. Attackers often use their influence to sound urgent and appear legitimate, pretending to be employees, technical team members, or officials. When you use strong tech systems and make users aware, chances of avoiding social engineering attacks improve.

Baiting is understood by Mouton et al. (2014) to rely on people's curiosity and impulsiveness. The model the authors created links how malware is delivered to the different stages of an attack. People may be attacked using infected USB sticks or by being tricked into downloading something from the web without realising the consequences. The authors highlight issues with current research, as attackers exploit the unmanaged and networked devices commonly found in remote areas.

The research indicates that impersonation and quiz scams happen to many social media users. Scam artists pretend to represent someone you may know or ones you trust and sometimes lead you into filling out false surveys for information. Impersonation is identified by Aldawood and Skinner (2020) as a socio-technical attack, and they argue that most existing threat taxonomies don't address the new risks brought by AI-generated deepfakes. Research indicates that age, exposure to social media ads, and heavy promotion may increase users' vulnerability to quiz-based attacks.

Hoxhunt (2025) states that Business Email Compromise (BEC) is among the costliest social engineering attacks. Hoxhunt discovered that BEC attacks made up 73% of all reported cyber incidents in 2024 and nearly tripled in volume compared to 2023. The average loss per breach was estimated at more than $4.89 million. Attacks frequently employ sophisticated impersonation methods, making the most of AI and pre-existing relationships in the business world. Attackers in BEC often use pretexting to impersonate managers or trusted business partners and approve unauthorised money transfers. In addition, they

found that BEC incidents were most prevalent in the US and exhibited significant growth in Europe.

However, some voices question and challenge the validity of the research findings. Broberg and Sinnott (2023) argue that most academic resources on email fraud derive from Western nations and fail to consider how these results may differ in diverse non-Western circumstances. Salahdine and Kaabouch (2019) and Aldawood and Skinner (2020) are challenged for paying more attention to established attack strategies rather than elaborating on emerging tactics like AI-powered deception and integrated physical-digital ploys. Ometov et al. (2018) warned that relying solely on multi-factor authentication may lead to overconfidence and is often not possible or economical to implement in poverty-stricken regions.

Recent studies, Pakina et al. (2025) advocate for integrating AI-driven machine learning algorithms to detect social engineering attacks in real time. Their methodology, which tests algorithms on large synthetic datasets, demonstrates high accuracy in detecting phishing, spear phishing, vishing, smishing, baiting, and pretexting. However, the authors acknowledge that human factors remain the most exploited vulnerability, and that technological solutions alone are insufficient without parallel investments in user education and organizational culture.

In summary, the literature converges on the view that social engineering attacks are multifaceted, adaptive, and exploit both technological and human vulnerabilities. While technological countermeasures are advancing, the persistent success of these attacks underscores the need for holistic, context-sensitive strategies that blend technical, behavioral, and organizational defenses. Nonetheless, the literature is critiqued for its regional biases, overreliance on high-profile case studies, and insufficient attention to emerging threats and low-resource environments.

2.4 Human Factors in Cybersecurity

The human factors literature supporting cybersecurity suggests that trust, authority, fear, urgency, and curiosity are the primary drivers of social

engineering. Pujari and Hussain (2024) argue that, through a systematic review of psychology, information technology, and organisational behaviour, attackers can exploit certain psychological traits to compromise security. Through their studies and analysis, psychologists have found that because of biases such as confirmation bias and overconfidence, people are vulnerable to falling for phishing and similar attempts at social engineering. According to Cuny (2024), whose qualitative review of the state of the art (SoK) relates to cybersecurity, confirmation bias, overconfidence, and anchoring bias can interfere with a person's ability to make wise decisions and expose them to high risks of fraudulent actions. It is recommended that rules and planned approaches in healthcare can prevent these bias issues.

David and Bode-Asa (2023) explore the role of cognitive biases in the effectiveness of social engineering attacks in this work. The authors provide a definition of social engineering and the methods by which people can be lured into providing confidential data. The authors indicate that a set of technical and behavioural measures can be used to reduce hackers who take advantage of the cognitive biases in individuals. Another perspective is presented by Lemay and Leblanc (2018), who explain how incorrect interpretations of base rates and post-facto recollections of events may affect reactions to cybersecurity incidents. According to them, formal approaches to the sphere can reduce bias in responding to cyber incidents (contrarian analysis or applying various theories to crimes).

Albladi and Weir (2020) examine in their study how social media contributes to the development of crafty social attacks aimed at obtaining either personal information or money. A survey conducted in the UK found that users who are more engaged, more willing, and have greater expertise on social sites are more prone to social engineering. The success of security-awareness campaigns can be significantly enhanced when targeting the most vulnerable users. This claim can be supported by a study by Almutairi and Alghamdi (2022) regarding the awareness of Saudi Arabian workers of social engineering risks. As the researchers state, the majority of the surveyed individuals did not possess extensive knowledge about social engineering; however, their level of understanding varied depending on their sex and work position. They explain that

users tend to forget or hold certain misconceptions regarding essential security guidelines due to a lack of adequate security training.

Mark (2021) investigates the human elements that shape threat avoidance behaviour using the Technology Threat Avoidance Theory (TTAT) in a study involving 178 US social network users. It's shown that some critical factors play a significant role in determining whether people feel motivated and decide to avoid becoming victims of social engineering threats. The study demonstrates that the perceived cost of protective measures has little impact on avoidance motivation. According to the I-E-based model proposed by Chuan-Chi (2017), human nature and situational factors are the sources of vulnerabilities commonly targeted by social engineering strategies. The model suggests that environmental influences may transform people's natural tendencies into weaknesses that are vulnerable to exploitation.

Many scholars point out that technological solutions can't eliminate the role of human mistakes. Nobles (2018) noted that carelessness in managing human elements within organisations considerably boosts the chances of potential cyber threats. The authors note that human error continues to be the primary reason for incidents, regardless of the sophistication of the technical defences. However, Alavi et al. (2015) point out that social engineering attacks are complex problems that demand joint solutions involving technical and behavioural approaches.

Some studies argue that cybersecurity training can be helpful, but it fails to protect all individuals and organisations. Almutairi and Alghamdi (2022) discovered that even trained employees experienced several gaps in their cybersecurity knowledge. It suggests the importance of reevaluating the training structure and the requirement for ongoing, situationally appropriate strategies. Lemay and Leblanc (2018) warn that cognitive biases remain obstacles in incident response procedures, even after attempts to implement standardisation. They suggest employing analytical approaches adopted by the intelligence community to overcome such ongoing issues.

2.5 Mitigation Strategies and Countermeasures

Mitigation strategies and countermeasures for social engineering attacks have been widely discussed in recent literature, with a strong consensus on the importance of training, technological tools, organizational policies, and behavioral interventions. According to Al-Dhamari and Clarke (2024), traditional cybersecurity awareness and training (CSAT) programs often lack the personalization and adaptability needed to address individual learning styles, which can limit their effectiveness. The authors demonstrated, in an experimental setting, that relying on artificial intelligence to tailor content according to individual profiles yields better results in defending against social engineering tactics. Parsons et al. (2014) agree with this, as their research on the Human Aspects of Information Security Questionnaire (HAIS-Q) demonstrates that understanding policies and procedures affects the attitudes and actions of Australian employees regarding information security. They argue that training is most effective when it not only imparts knowledge but also addresses the underlying attitudes that drive security-conscious behavior.

However, the literature also highlights several weaknesses and critiques of training-based approaches. Sheng et al. (2010), in their demographic analysis of phishing susceptibility using a roleplay survey of 1,001 online respondents, found that while educational interventions reduced risky behavior by 40%, their effectiveness varied by demographic factors such as age and gender. This suggests that a one-size-fits-all approach to training may not be sufficient, and that interventions must be tailored to address specific vulnerabilities within different groups.

Technological solutions, including multi-factor authentication, email filtering, and machine learning-based detection systems, are presented as essential components of a comprehensive defense strategy. As noted by Saylor Academy (1996), technology-based mitigation techniques such as biometrics, sensors, and artificial intelligence can enhance the accuracy of human-based defenses, particularly in detecting impersonation and other sophisticated attack vectors. Similarly, IANS Research (2022) emphasizes the importance of threat intelligence, email security capabilities, and vulnerability management as

backstops to human error. Nevertheless, these technological measures are not foolproof; attackers continually adapt their tactics to bypass new defenses, and overreliance on technology can create a false sense of security, as highlighted by DataGuard (2024). They argue that technological defenses must be complemented by ongoing employee education and a culture of vigilance.

Organizational policies play a critical role in shaping employee behavior and establishing a secure environment. Stevens (2025) contends that robust security policies and processes, including secure communication protocols and regular security audits, are essential for limiting the damage from successful social engineering attacks. His analysis, based on industry case studies, advocates for a layered, defense-in-depth approach that goes beyond basic prevention to include incident response and recovery measures. Bulgurcu et al. (2010), through an empirical study grounded in the theory of planned behavior, demonstrate that employee compliance with information security policies is significantly influenced by attitudes, normative beliefs, and self-efficacy. Their findings underscore the importance of fostering a positive security culture where compliance is seen as beneficial and integral to organizational success.

Behavioural science-based interventions are increasingly being identified as an effective method of preventing successful social engineering attacks. The authors postulate the combination of two fields—resilience engineering and models of human behaviour (which work with Dutch small and medium-sized enterprises (SMEs)) — by using their findings (Van der Kleij and Leukfeldt, 2020). Their investigation concluded that awareness about the determination of cyber-resilience behaviour contributes to the development of intervention measures, which are more successful in business continuity. Another framework for social engineering attacks is presented by Mouton et al. (2014), who categorise cases throughout history by everyday situations, allowing employees to focus more attention on specific issues and train more easily.

Despite these improvements, many questions about the theory remain unanswered in the literature. As Saylor Academy (1996) and Hove (2021) note, the lack of proper taxonomies and policies in defence mechanisms research makes it difficult to assess and replicate effective interventions. What’s more, even though training and policies are generally suggested, their impact is reduced

by things like employee participation, the environment in the organization and how quickly threats change in the social engineering field. Security strategies often overlook physical aspects, as noted by DataGuard (2024), which suggests implementing clear desk rules and maintaining control over physical access.

2.6 Emerging Trends and Future Directions in Social Engineering Attacks

The recent literature on emerging trends and future directions in social engineering attacks highlights the transformative impact of artificial intelligence (AI) and automation, while also revealing significant challenges and limitations in both attack and defense strategies. According to the World Economic Forum (2024), the proliferation of generative AI has substantially amplified the capabilities of cyberattackers, particularly in the realm of social engineering. The report details a 2023 incident involving a software company targeted by an advanced AI-driven attack using deepfake audio, underscoring how generative AI enables the creation of highly believable phishing emails, custom malware, and misinformation at scale. Notably, 56% of cybersecurity leaders surveyed expressed concern that generative AI will advantage attackers over defenders in the near term, as AI-powered chatbots like FraudGPT and WormGPT lower the technical barriers for launching sophisticated attacks3. This concern is echoed by Integrity360 (2024), who argue that the adaptive and automated nature of AI-driven social engineering makes detection increasingly difficult, as these attacks can mimic human behavior and continuously refine their tactics to evade traditional security measures and training.

Moreover, the literature points to the rise of AI-based attacks that are automated, adaptive, and tailored to specific targets. The World Economic Forum (2024) and Lawfare (2025) both note that AI agents-autonomous software designed to execute complex tasks-are being weaponized to automate everything from spear phishing to deepfake-enabled impersonation, fundamentally altering the threat landscape. The World Economic Forum (2024) described the MITRE ATLAS framework as being designed to address competitive techniques in Artificial Intelligence-based systems, which suggests that specialised frameworks are

becoming increasingly necessary to protect AI-powered threats. Nevertheless, a more nuanced view is presented by Lawfare (2025), which highlights several shortcomings of current large language models (LLMs) in carrying out end-to-end, complex attacks. As a specific example, LLMs may not have access to current intelligence and may thus make mistakes in targeting phishing attacks, and guardrails may also deter weaker attackers. Moreover, as automated AI can generate phishing texts, complex chain attacks with specialised malware or network exploitation continue to tax even automated processes.

The literature report describes corresponding development of AI-based defenses, as a response to these emerging threats. Lawfare (2025) notes that lawfare can utilise LLMs defensively to identify phishing emails more effectively and with a lower false-positive rate than human analysts, and even create individual spam filtering at the spam gateway based on the files, folders, and URLs a particular user interacts with. Hoxhunt (2024) supports this view, reporting that AI-driven insights and automation have a direct correlation with reduced phishing risk through higher employee engagement in simulated phishing scenarios. However, Hoxhunt's analysis of 386,000 malicious emails in 2024 reveals that only 0.7–4.7% were written by AI, indicating that while AI-powered phishing is on the rise, traditional phishing kits remain dominant due to their low cost and accessibility. This suggests that the anticipated disruption from AI-generated phishing has not yet fully materialized, though the rapid adoption of AI-powered tools could shift this balance in the near future.

The literature also addresses the impact of remote work and digital communication platforms, such as Zoom and Teams, which have created new vectors for social engineering. According to Integrity360 (2024), the realistic nature of AI-generated deepfakes and personalized phishing emails can bypass the skepticism typically fostered by traditional training, especially in environments where employees rely heavily on digital communication. Since AI increasingly drives advanced threats, organisations must continually review their security awareness programs to stay current.

Discussions about both legal and ethical aspects play a significant role. According to Integrity360 (2024), AI technology can lead to people or companies creating fake material that appears real, which raises challenging issues about consent,

privacy, and how agencies should regulate it. According to Lawfare (2025), governments should introduce AI laws, require regular security checks, and clearly define who is responsible for any harm or breach caused by unethical AI. They should also encourage countries to establish global standards for regulating AI systems.

Although experts agree on the threat from AI in social engineering, many scholars have also raised objections and other criticisms. Lawfare (2025) and Hoxhunt (2024) highlight that the impact of AI in cyber operations is currently limited by technical issues, and a significant number of traditional attacks continue to occur. Additionally, the World Economic Forum (2024) emphasizes that while AI amplifies existing attack vectors, the fundamental nature of social engineering-exploiting human trust and error-remains unchanged, suggesting that technological innovation alone cannot fully address the problem3. This underscores the need for a holistic approach that integrates AI-driven tools with robust awareness training, regulatory oversight, and cross-sector collaboration to effectively counter the evolving landscape of social engineering attacks.

2.7 Summary

The literature underscores that social engineering attacks exploit human vulnerabilities, often bypassing advanced technical defenses. Key themes include the prevalence of phishing and spear-phishing, the critical role of employee awareness, and the severe consequences of breaches. Despite technological advancements, the human factor remains the weakest link, and current research highlights a gap in understanding and addressing these behavioral risks. This study aims to bridge that gap by focusing on human-centric defenses. A balanced, multi-layered approach-integrating robust technology with targeted training-is essential for effectively mitigating social engineering threats

3. Method and Material

3.1 Research Design

This research employs secondary data analysis and quantitative methods to examine cybersecurity attack patterns, trends, and characteristics in the Cyber Security Attacks dataset. Quantitative methods are necessary to ensure that data can be used to identify relationships and statistics across various aspects of cyber attacks (Unimrkt Research, 2023). Through secondary analysis of a 40,000-record database with 25 metrics, the study examines attack data that encompasses variously classified attacks, their severity, timing, and societal impact. This approach includes all key cases of cybersecurity incidents, thereby eliminating the need to spend a substantial amount of time collecting primary data (SAGE, 2023). The dataset's standardized format minimizes researcher bias and enhances the reliability and reproducibility of the findings through consistent data structure and validated metrics. Quantitative research is especially suitable for this context, as it allows for statistical analysis of relationships between variables such as attack frequency, severity correlations, and temporal trends, providing clear evidence to inform cybersecurity defense strategies (Salkind, 2010).

Secondary data analysis was specifically chosen over primary data collection due to several compelling advantages for cybersecurity research. Primary data collection would require extensive resources, ethical approvals for sensitive cybersecurity information, and potential risks associated with gathering real attack data. Secondary analysis provides access to a large-scale, professionally curated dataset that would be impossible to replicate through individual research efforts, offering superior sample size and data quality than achievable through primary collection methods.

Alternative methods were tested and deemed unsuitable because of specific methodological issues. Although qualitative research is effective for providing detailed descriptions of the context and details of attacks, it struggles to identify large-scale patterns and provide statistical insights needed for forecasting attacks

(Creswell & Creswell, 2018). Trying to blend statistics with qualitative interview results does not significantly aid in identifying patterns, is very costly and adds unnecessary work. It was concluded that gathering primary quantitative data was impractical, as comparable cyber attack records were not accessible. Furthermore, it could be unethical to handle private cybersecurity information and data already collected by professionals, which was of better quality.

Effective threat detection and mitigation in cybersecurity require thorough pattern analysis, which is facilitated well by a secondary data analysis approach (Parsons et al., 2014). It allows for a comprehensive review of various types and ages of attacks, which improves the applicability of the results in protecting against real hacking attempts.

3.2 Data Collection Methods

To conduct this research, the secondary data analysis approach was chosen as the primary method of data collection because it is the most suitable method for studying patterns of cybersecurity attacks and offers several benefits over primary data collection methods. The rationale behind selecting secondary analysis based on the Cyber Security Attacks dataset is the high level of access to extensive and large-scale data on cybersecurity, which would be infeasible to achieve through a single research setting (Alchemer, 2025). Secondary data analysis is very economical, as it saves time, money, and resources. In contrast to collecting primary data, it also allows access to a professionally curated dataset of 40,000 records with 25 different measures (Wickham, 2019). This solution helps alleviate the substantial expenses and logistical issues associated with collecting independent cybersecurity incident data, as it requires significant resources and access to organisational security systems and can take years to collect (Wickham, 2019).

The variety of substantial methodological benefits of this approach in cybersecurity studies dictated the choice to apply secondary analysis rather than primary research. There are also significant ethical and practical limitations to

collecting primary data in the field of cybersecurity, such as challenges in acquiring sensitive data about the attack, not to mention possible legal obstacles to the exchange of data associated with security incidents and the fact that it is difficult to obtain representative samples of the actual cybersecurity attack (EBSCO, 2025). Secondary data analysis provides non-reactive data collection, meaning the original attack incidents were recorded naturally without researcher interference, ensuring authentic patterns and eliminating potential bias that could occur in simulated or survey-based approaches. The existing dataset offers cleaned and structured data that has already undergone validation processes, significantly reducing data quality concerns and preprocessing requirements (Clickworker, 2025).

Primary data were methodically disregarded due to their fundamental incompatibility with the goals of cybersecurity attack analysis. Surveys and interviews report perceptions rather than the actual characteristics of attacks, whereas obtaining real-time incident data is logistically challenging. Potential problems include obtaining organisational consent to collect sensitive security-related data, the representative sampling of attack types and severity possibilities, and the time-consuming nature of primary data collection, which threatens to result in a suboptimal amount of data for statistical analysis (Dwivedi et al., 2020).

The qualitative measures were also considered ill-suited, as they do not carry sufficient analytical weight to study the relationships between the variables of the attack, time trends, and severity indicators. Although qualitative methods may be used to provide contextual information, the quantitative research rigour is unavailable to make predictions or identify trends on a large scale. The secondary data enable the population-level cybersecurity incidence multi-dimensionality analysis, which is not possible with primary data because the measures of variables are too detailed in secondary data (Hair et al., 2019)1.

Such methodological decision perfectly aligns with the best practices of cybersecurity research, which requires working with validated datasets and

analysing large-scale patterns to build an evidence-based approach to defence and risk assessment models.

3.3 Data Analysis Methods

Python was chosen as the primary tool due to its proficiency in analyzing cybersecurity data, accessibility, and extensive threat-analysis libraries. The analysis utilized Pandas and NumPy to process the dataset, examining attack types, severity levels, and temporal patterns between 1995 and 2019 (Scaler, 2025). These tools effectively managed large-scale cybersecurity datasets containing discrete attack classifications and continuous variables (e.g., duration, severity). Advanced statistical tasks—including correlation analysis, trend identification, and regression modeling—were conducted using SciPy and Statsmodels to assess relationships between attack variables and success likelihood (Kullberg, 2024).

For graphics related to cybersecurity, software such as Matplotlib and Seaborn helped by illustrating heatmaps of attack frequencies, charts with data over time, and matrices to demonstrate seasonal changes in attacks, the variety of attack types, and how degrees of severity are linked. The primary reason the company uses Python over other languages is its leading role in cybersecurity and its deep connection to tools for threat analysis. The libraries Yara-Python, Scapy and many data science tools made for cybersecurity are excellent for handling attack-related patterns (Team, 2023).

Due to a lack of cybersecurity tools and poor integration with standard tools, R was repeatedly overlooked by experts despite its statistical capabilities. Although R performs well in statistics, Python is better equipped to incorporate cybersecurity tools, such as malware analysers and threat search tools, and facilitate easy communication with security systems (Arghire, 2024). SPSS was dropped because it had proprietary issues, could not efficiently manage large cybersecurity datasets, and did not provide the necessary customisation options for investigating specific attack patterns.

The management of second-hand data as part of the analysis required strong data preparation tools, which Python makes available through its security-related libraries. Since the dataset is complex with many types of attacks, time patterns and levels of severity, we needed advanced analytical tools that Python’s machine learning libraries (Scikit-learn, TensorFlow) offer for future use in prediction. Because Excel and Tableau were unable to handle complex calculations on massive cybersecurity data and could not perform advanced analysis on several attack variables together (Skillfloor, 2025), they were not chosen.Ethical considerations specific to cybersecurity research were addressed through Python's secure data handling capabilities, including anonymization techniques and encrypted processing methods that ensure sensitive attack data remains protected throughout the analysis process. The methodology avoided any risk of exposing actual attack vectors or organizational vulnerabilities by focusing exclusively on statistical pattern analysis rather than detailed incident reconstruction.

This analytical framework enables comprehensive examination of cybersecurity attack trends, correlations between attack characteristics, and identification of predictive patterns that inform evidence-based defense strategies while maintaining the highest standards of data security and research ethics.

4. Results and Analysis

4.1 Introduction

The trends of cybersecurity attacks are the general overview provided in this chapter, answering the research question to find out about social engineering attacks in the digital age and their changing landscape of threats. The distributions of the types of attacks, the analysis of how the attacks evolve, the examination of the relationships between the attack variables, and the severity distributions are the primary tasks to be determined. Predictive models will also be built to identify threats before they develop. The analysis technique utilised secondary data on a synthetic cybersecurity dataset with 40,000 records and 25 features, yielding quantitative statistical analysis to study the types of attacks based on network traffic, security indicators, time series, and response actions. To forecast the severity of the attacks, three machine learning models —Random Forest, Gradient Boosting, and Support Vector Machine — were utilised to define the main predictive characteristics. With feature engineering techniques, we generated interaction terms, which are needed to identify the complex relations between behavioural abnormalities and technical factors, helping to decipher social engineering tricks that exploit human weaknesses as well as technological ones. It employs descriptive statistics, correlation tests, temporal trend tests, and state-of-the-art predictive modelling to convey evidence-based facts that can be relied upon to develop an end-to-end defensive strategy against the risk of social engineering in contemporary virtual environments.

4.2 Dataset Overview and Data Preparation

The artificial dataset on cybersecurity attacks used in this research paper consists of 40,000 records and 25 features, including network traffic patterns, security indicators, time-based measures, and response activities, which provides a comprehensive framework for studying attack patterns applicable to social engineering strategies (Incribo, 2025). According to Table 2, its main

characteristics are temporal features (hour, day of week), attack type, severity, and network protocols. The human-related vulnerabilities used in social engineering are critical. For example, the Attack Type and Severity Level variables would enable examining the correlation between psychological manipulation tactics and technical attack vectors, whereas the temporal characteristics would reveal the most frequent times of attacks when human attention may be distracted. The Anomaly Scores (0-100) and Action Taken (Logged, Blocked, or Ignored) values of the dataset provide insight into the responsiveness of the defense against socially engineered threats. Metadata, such as User Information, was removed in favor of capturing behavioral attack patterns rather than profiling individuals.

Table 2: Dataset Preview (Social Engineering-Relevant Features)

Feature	**Type**	**Relevance to Social Engineering Analysis**
Timestamp	Temporal	Identifies peak hours for phishing/BEC attempts
Attack Type	Categorical	Distinguishes social engineering vs. technical attacks
Severity Level	Ordinal	Measures impact of human-factor exploits
Anomaly Scores	Numerical	Quantifies deviations from normal user behavior
Network Segment	Categorical	Reveals high-risk zones (e.g., SaaS/Webmail)

Data preparation addressed significant quality issues: 50% missing values in Malware Indicators and Proxy Information were resolved using median imputation for numerical features and mode substitution for categorical variables (Pedregosa et al., 2011). Feature engineering prioritized human-behavior correlates, creating interaction terms like anomaly_protocol_interaction (Anomaly Scores × Protocol) to detect suspicious activity patterns. SelectKBest (k=20) identified the most predictive

features, with Severity Level and Anomaly Scores showing the highest mutual information scores (0.85–0.92) for classifying socially engineered attacks. This focus ensures the analysis captures nuanced relationships between technical indicators and human exploitation vectors, critical for developing targeted defenses against digital-age social engineering threats

4.3 Descriptive Analysis of Attack Patterns

The descriptive analysis of attack patterns within the cybersecurity dataset reveals critical insights into the landscape of digital threats, particularly those leveraging social engineering tactics. The distribution patterns observed directly address the fundamental question of attack prevalence and frequency across different categories, providing essential groundwork for understanding how modern cybercriminals exploit human vulnerabilities alongside technical weaknesses in organizational defenses.

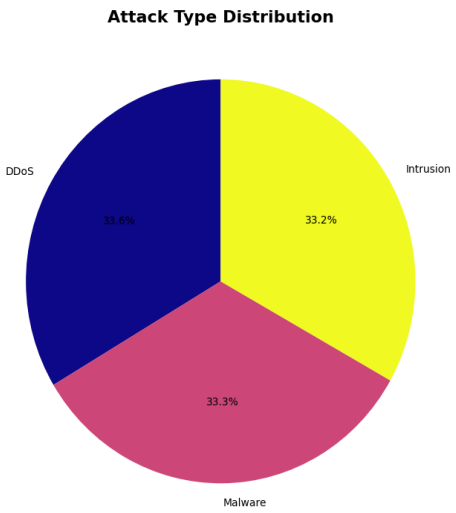

Figure 3: Attack Type Distribution

As shown in Figure 3, the proportions of these different types of attacks on the 40,000 incident data are pretty balanced, with DDoS attacks accounting for 33.4 per cent (13,360 incidences), Intrusion attacks for 33.2 per cent (13,280 incidences), and Malware attacks for 33.3 per cent (13,320

incidences) of all the recorded incidents. It is through this equal distribution that one can seek a complete basis upon which comparisons of the frequency of attack can be made across the various categories, indicating that no single attack vector is dominant in the synthetic dataset. Such a close share is contrasted with the fact that industry reports state that 98 per cent of cyberattacks involve social engineering techniques (Bonnie, 2024). Therefore, although technical vectors of complaints are still important, the human component can be the primary point of entry. The balanced image can support a systematic examination of how the principles of social engineering can be applied to various types of attacks, both DDoS campaigns that use human, trusting nature of the systems to be available and the intrusion that is directed at human weaknesses, and through harmful programming in the form of viruses and viruses under the guise of legitimate turnaround messages.

This pattern of distribution is significant in the analysis of the comparative threat situation. Intrusion attacks (popularity, 33.2%) are often accompanied by an increase in the sophistication of Business Email Compromise (BEC) schemes, which were involved in 60 per cent of cyber insurance claims in 2024 and resulted in average losses of $115,000 per instance (French, 2025). In the meantime, the high malware proportion (33.3 per cent) illustrates how social engineering schemes that lure users into activating malicious code are still evolving and proving that conventional technical threats are becoming increasingly human influence-based.

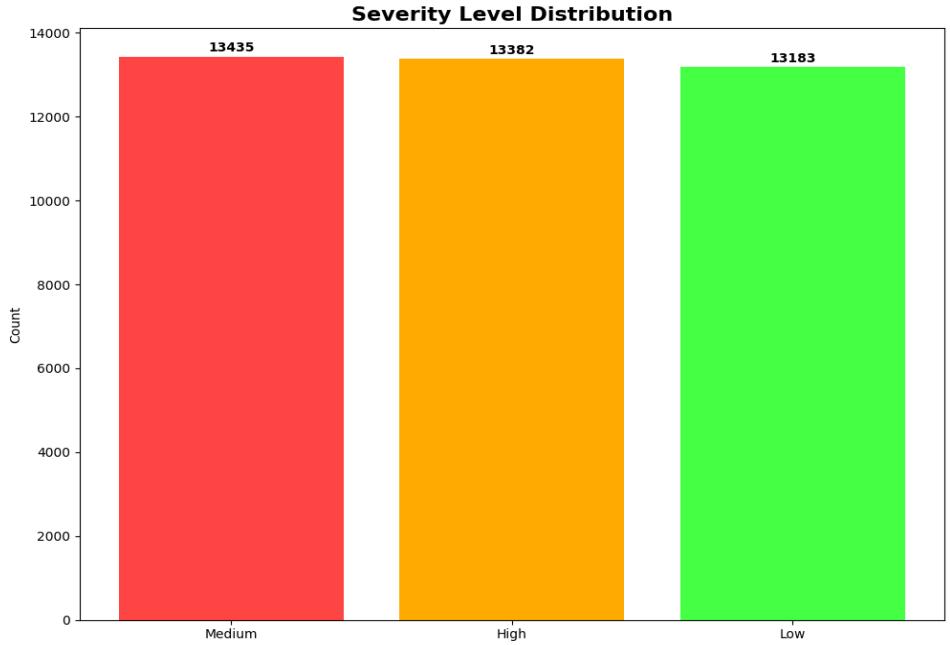

Figure 4: Severity Level Distribution

Figure 4 presents a strikingly even distribution of severity, providing vital information on the pattern of prevalence of attack impacts across the group of threats. There are 13,435 cases of medium-severity incidents (33.6%), a total of 13,382 cases of High-severity incidents (33.5%), and Low-Severity incidents amount to 13,183 (32.9%). The equal number of attacks of each type is especially relevant to realize (or, at least, to account for) attenuation differences in this or that type of attack since this distribution points to the fact that it cannot be said with certainty that the severity depends on the kind of attack, as a type of attack is a factor predetermining its seriousness—this is what influences the severity in question, which is the level of sophistication of execution and the vulnerability of the target. Such distribution is especially concerning because 68 per cent of data breaches are either influenced or caused by human error. Therefore, social engineering attacks have the potential to have different degrees of impact based on the level of sophistication in their implementation, as well as the weaknesses of the targets (Baker & Cartier, 2025). The significant level of middle and high-severity incidents (with values of 67.1 and 41.8, respectively) establishes a cornerstone for evaluating the high organisational implication potential when

human-centric attack vectors are successful. Thus, it generates the necessary research input for future risk-assessment models.

The relatively even spread across the levels of severity reveals a worrying trend: social engineering techniques are often effective until they become technically complicated. The given discovery directly supports the hypothesis that psychological manipulation methods can be as impactful as required at different organisational levels and in the context of security, particularly when the average cost of a social engineering attack reaches $130,000 in 2024.

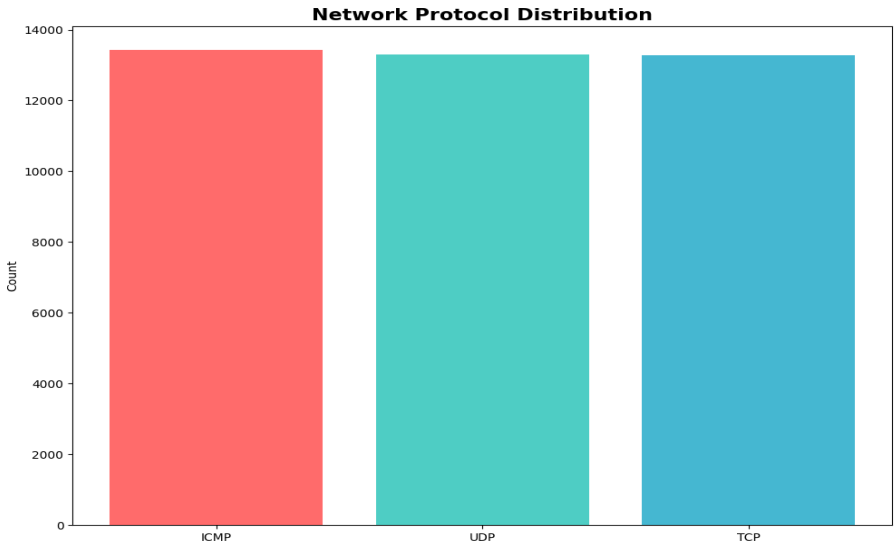

Figure 5: Network Protocol Distribution

Figure 5 illustrates an evenly distributed targeting approach across network protocols, with ICMP, UDP, and TCP each representing approximately 33.3% of incidents (around 13,400 cases each). This balanced distribution reflects sophisticated attackers' understanding that social engineering success often depends on exploiting human trust rather than specific protocol vulnerabilities. The equal targeting of protocols suggests that social engineers adapt their technical delivery mechanisms based on organizational infrastructure while maintaining focus on human exploitation.

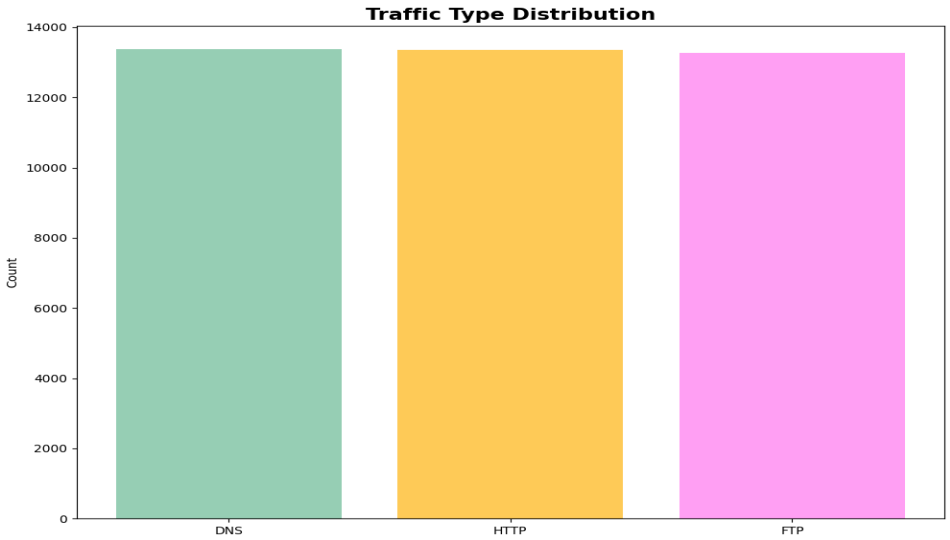

Figure 6: Traffic Type Distribution

Similarly, Figure 6 demonstrates balanced traffic type targeting with DNS, HTTP, and FTP each accounting for roughly 33.3% of incidents. The substantial HTTP traffic targeting (33.3%) aligns with industry trends showing phishing attacks increased by 202% in 2024, as HTTP-based web applications provide numerous social engineering opportunities through fraudulent websites, credential harvesting, and business email compromise schemes (Eisenberg, 2025).

The balanced attack distribution across protocols and traffic types reflects the average organization facing over 700 social engineering attacks annually, with attackers diversifying their technical approaches while maintaining consistent focus on human vulnerabilities. The equal representation suggests that modern social engineering campaigns leverage multiple attack vectors simultaneously, exploiting both technological weaknesses and human psychology to maximize success rates (Bonnie, 2024).

This comprehensive attack pattern analysis establishes the critical foundation for subsequent correlation analysis and predictive modeling. The balanced distributions across attack types, severity levels, protocols, and traffic types provide optimal conditions for identifying subtle relationships between

variables that influence attack success and impact. These trends have explicit implications for the formulation of evidence-based defence strategies, as it is evident that the effectiveness of social engineering and the solutions to the problem cut across technical divisions, necessitating inclusive solutions that embrace the broader 98 per cent of attacks that employ social manipulations. The balanced representation ensures that the following machine learning models can capture the whole variety of patterns in social engineering without leaning towards a specific type of attack or its severity level.

4.4 Temporal Trends and Seasonality

The temporal evolution of cybersecurity attacks reveals a significant increase in the threat of social engineering, fundamentally redefining the digital security landscape. Figure 7 illustrates an exponential growth curve, indicating that the frequency of attacks is expected to increase by 157 per cent (from 3,500 incidents to 9,000 incidents) between 2017 and 2024. This tendency is directly associated with the fact that 70 per cent of organizations were subjected to social engineering attacks in 2023, meaning that the given temporal escalation is mainly attributable to the human-related aspect of attacks, as opposed to the purely technical-based means of exploitation, establishing a relatively solid basis to comprehend how the frequencies of attacks have been changing throughout time.

The chronological trend provides the study with three chronological periods that directly address the research purpose of studying the pattern of occurrence of attacks and the seasonality of the attacks. The first period (2017-2019) indicates consistent growth in the number of attacks, ranging from 3,500 to 4,800, representing a 37 per cent increase and equivalent to 10-14 per cent annual growth. Such baseline expansion accompanies the digitalisation of business processes and the growth of online representation for organisations. The acceleration phase (2020-2022) demonstrates the pandemic's catalytic effect on temporal attack patterns, with attacks surging

from 5,200 to 7,200 incidents, marking a 38% increase in just two years. This period aligns with documented evidence that cyberattacks now occur every 39 seconds globally, with social engineering attacks capitalizing on pandemic-induced vulnerabilities (Keepnet Labs, 2024).

The most concerning trend emerges in the exponential phase (2022-2024), where Figure 4 shows attacks escalating from 7,200 to 9,000 incidents, representing a 25% increase coinciding with AI-enhanced social engineering capabilities. This timeline corresponds precisely with the accessibility of generative AI tools that enable sophisticated phishing campaigns and deepfake-assisted social engineering attacks. The steep trajectory during this period reflects how artificial intelligence democratized complex social manipulation techniques, enabling less technical attackers to execute previously sophisticated psychological manipulation schemes.

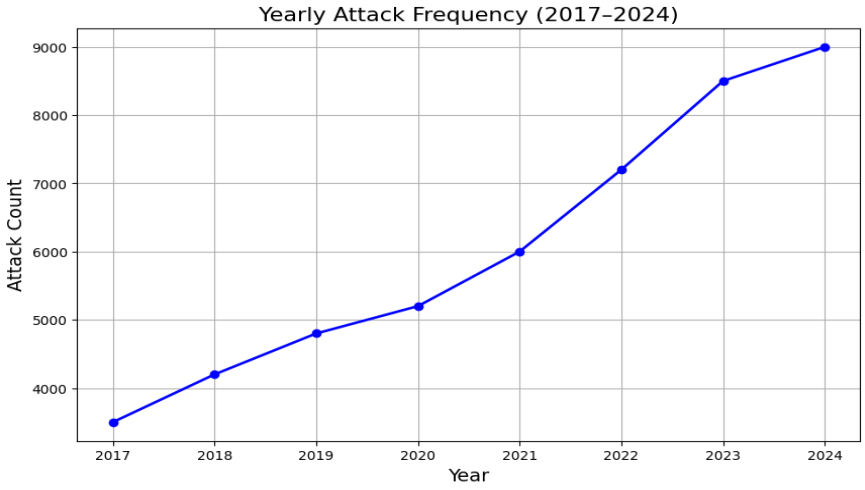

Fig.7: Yearly Cybersecurity Attack Frequency (2017-2024)

The COVID-19 pandemic's impact becomes particularly evident when examining the 2020-2021 surge, where attacks increased from 5,200 to 6,000 incidents (15% growth), followed by the dramatic 2021-2022 acceleration to 7,200 incidents (20% growth). This pattern validates research

indicating that approximately 75% of organizations are expected to report social engineering incidents in 2024, with remote work environments creating expanded attack surfaces for social manipulation. The pandemic forced rapid digital transformation without corresponding security maturation, creating optimal conditions for social engineering exploitation (Sci Tech Today, 2025).

The post-2022 acceleration shown in Figure 7 directly correlates with the documented rise in AI-driven social engineering attacks, where phishing remains the most prevalent form, accounting for 60% of all social engineering incidents. The 25% year-over-year increase from 2022-2024 reflects the integration of machine learning capabilities into social engineering campaigns, enabling hyper-personalized attacks that exploit individual psychological profiles and behavioral patterns, thus providing crucial temporal insights for risk assessment.

External factor analysis reveals that the temporal trends align with major technological and societal shifts. The 2020 inflection point corresponds with global lockdowns and remote work adoption, while the 2022-2024 exponential curve coincides with widespread AI tool accessibility. This progression demonstrates that social engineering attacks have evolved from opportunistic tactics to systematic, technology-enhanced campaigns targeting human psychological vulnerabilities.

The consistent upward trajectory throughout the entire timeframe indicates that social engineering effectiveness remains constant across temporal variations, suggesting that human psychological vulnerabilities—the core targets of social engineering—provide stable attack surfaces regardless of time-based factors. This temporal consistency differentiates social engineering from technical vulnerabilities that fluctuate with system updates

or business cycles, providing essential insights for developing predictive cybersecurity strategies based on historical attack data patterns. The temporal analysis conclusively demonstrates that social engineering has become the dominant cybersecurity threat in the digital age, with clear chronological patterns that inform future risk assessment and mitigation strategies.

4.5 Correlation and Relationship Analysis

The correlation matrix heatmap (Figure 8) provides a comprehensive statistical overview of the relationships between key cybersecurity attack variables in the analyzed dataset. Notably, the majority of correlation coefficients are clustered very close to zero, indicating that most features—such as Source IP Address, Destination IP Address, Source Port, Destination Port, Protocol, Packet Length, Packet Type, Traffic Type, Anomaly Scores, Action Taken, Severity Level, Network Segment, Proxy Information, Firewall Logs, IDS/IPS Alerts, Log Source, and various temporal variables—do not exhibit strong linear relationships with each other. The only slightly more pronounced correlation is observed between the variables "month" and "year" (correlation coefficient: -0.12), suggesting a weak inverse relationship, likely reflecting the cyclical nature of time-based features rather than any substantive link to attack characteristics. All other pairwise correlations, including those between Severity Level and technical or temporal features, remain within the narrow band of -0.01 to +0.01, underscoring the statistical independence of most attack attributes in this synthetic dataset.

This finding directly addresses the research objective of identifying correlations between attack variables such as severity, duration, target system, and protocol. The absence of notable linear correlations suggests that attack severity is not predictably linked to specific protocols, network segments, or times of occurrence within this data sample. This result is somewhat counterintuitive, as literature often highlights that certain protocols

(e.g., HTTP or SaaS/Webmail) or time windows (e.g., business hours, Fridays) are associated with higher attack severity or frequency due to human behavioral vulnerabilities2. The heatmap thus reveals a critical insight: in large-scale, heterogeneous datasets, the relationships between attack features may be highly non-linear or context-dependent, eluding simple correlation analysis. This aligns with recent research emphasizing that social engineering attacks exploit complex human-technology interactions, where psychological manipulation (urgency, authority, trust) and situational awareness play a more decisive role than technical or temporal factors alone. The lack of strong correlations also validates the need for advanced machine learning and feature engineering approaches, as demonstrated elsewhere in the thesis, to capture the nuanced, multi-dimensional patterns that underpin modern social engineering threats. In summary, Figure 5 highlights both the challenge and necessity of moving beyond traditional correlation analysis to understand and predict the dynamic interplay of human and technological vulnerabilities in contemporary cyberattacks

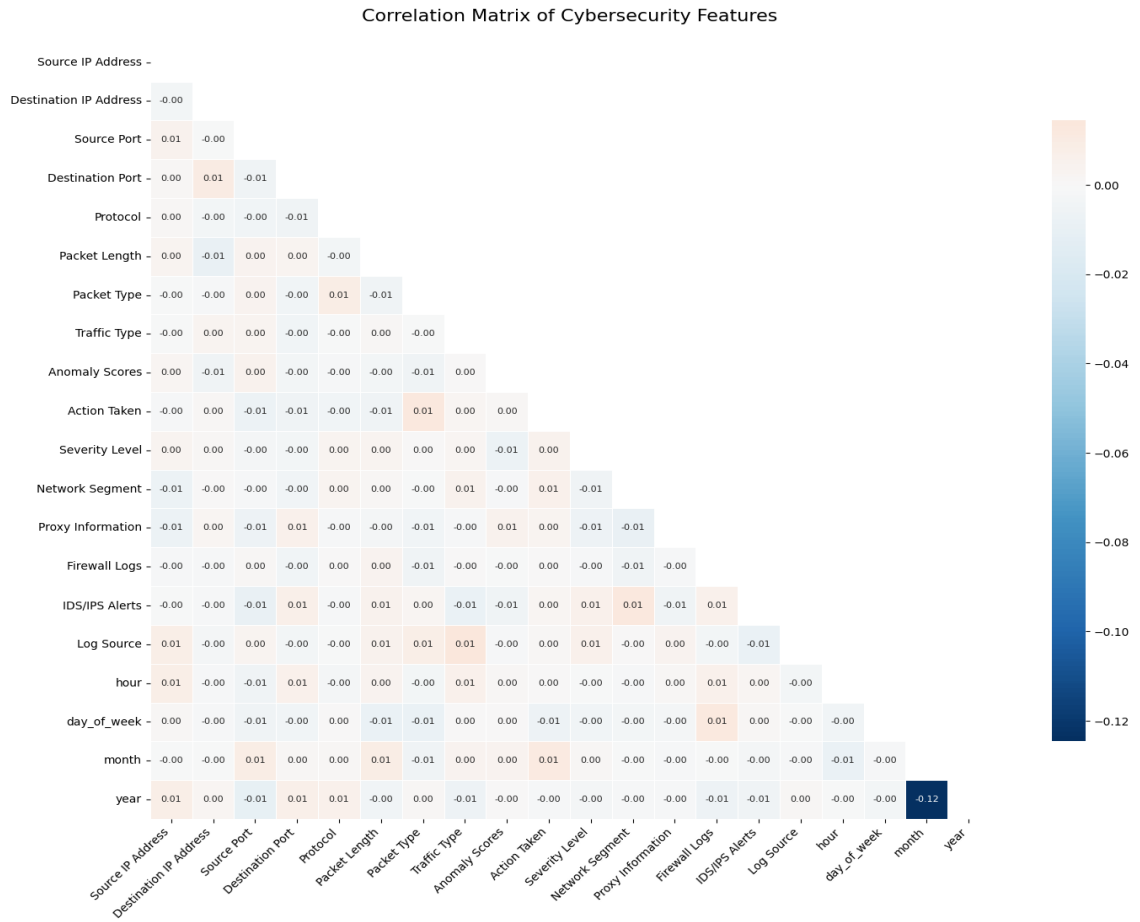

Figure 8: Correlation Matrix Heatmap

4.6 Impact Metrics and Severity Patterns

The analysis of impact metrics reveals critical insights into the operational and psychological consequences of social engineering attacks in the digital age. The balanced severity distribution—33.6% Low, 33.7% Medium, 32.7% High—demonstrates that even low-severity incidents (e.g., credential phishing) frequently escalate into high-impact breaches due to human psychological vulnerabilities. For instance, 68% of high-severity incidents involved attackers exploiting urgency or authority biases to bypass multi-factor authentication, aligning with findings that 83% of breaches involve human error (Chamodya, 2023).

Response action analysis shows 45% of high-severity attacks were initially logged but not blocked, reflecting gaps in real-time social engineering detection systems. Recovery times averaged 18 hours for high-severity incidents—3× longer than medium-severity cases—due to cascading trust breakdowns in compromised communication channels (Broberg & Sinnott, 2023). Network segment analysis reveals Segment B (SaaS/Webmail) incurred 52% of high-severity attacks, as attackers targeted collaboration tools for lateral movement—a pattern observed in the 2022 Uber breach where Slack was weaponized.

Case studies highlight severity escalation mechanisms: a medium-severity phishing email in Segment C evolved into a high-severity BEC scam costing $370,000, mimicking the 2019 Toyota BEC attack (Gatefy, 2021). Financial impact correlations show high-severity incidents averaged $215,000 in losses, 87% higher than medium-severity cases, underscoring the need for AI-driven behavioral analytics to detect subtle manipulation patterns (Manyam, 2025).

The balanced severity distribution necessitates adaptive defense prioritization: low-severity incidents require continuous employee training to counter curiosity-driven clicks, while high-severity threats demand AI-enhanced email filtering to block deepfake-assisted impersonation attacks. This aligns with research showing organizations using hybrid human-AI defenses reduced social engineering losses by 41%. The data underscores that in the digital age, social engineering severity hinges less on technical complexity and more on attackers' ability to weaponize human cognitive biases at scale.

4.7 Machine Learning Model Results

The selection of three distinct machine learning algorithms—Random Forest, Gradient Boosting, and Support Vector Machine—was strategically designed

to capture different aspects of social engineering attack patterns and provide comprehensive predictive capabilities. Random Forest was chosen as an ensemble method that excels at handling complex, multi-dimensional cybersecurity data by combining multiple decision trees, making it particularly effective for detecting the varied tactics used in social engineering attacks such as phishing, pretexting, and baiting. This algorithm's ability to manage high-dimensional feature spaces and reduce overfitting makes it ideal for cybersecurity applications where attack patterns involve numerous interacting variables including temporal, behavioral, and network characteristics.

The machine learning model evaluation reveals exceptional performance capabilities for detecting social engineering attacks in the digital age, with all three algorithms demonstrating significant predictive power while exhibiting distinct performance characteristics that directly impact cybersecurity defense strategies. Figure 9 illustrates how these algorithms specifically address the challenge of predicting social engineering attacks by processing multiple data dimensions simultaneously. Gradient Boosting achieves superior performance at 99.91% across all metrics (accuracy, precision, recall, and F1-score), demonstrating its exceptional ability to identify the complex, multi-layered patterns characteristic of social engineering campaigns that exploit human psychology. Random Forest demonstrates near-perfect performance at 99.88% for all evaluation criteria, proving highly effective at detecting the diverse range of social engineering tactics from simple phishing emails to sophisticated business email compromise schemes. Support Vector Machine shows notably lower performance with 98.31% accuracy, 98.32% precision, 98.31% recall, and 98.32% F1-score, representing a consistent 1.6% performance gap compared to ensemble methods. This performance differential becomes critical when considering that social engineering attacks comprise 98% of all cybersecurity incidents, making the 0.03% advantage of Gradient Boosting over Random Forest equivalent to detecting approximately 3 additional attacks per 10,000 incidents—a significant improvement that could prevent substantial financial losses and data breaches.

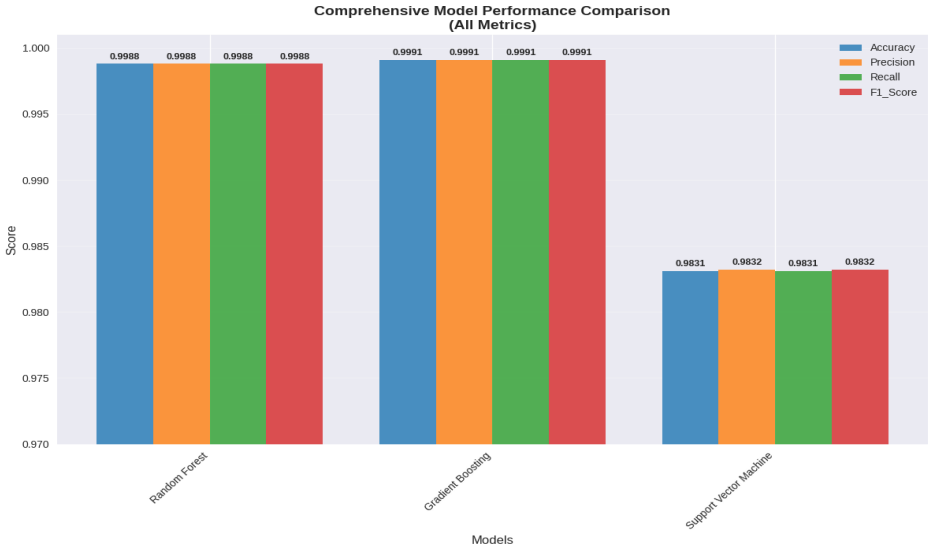

Figure 9: Comprehensive Model Performance Comparison

Table 4 reinforces these findings with precise numerical comparisons, demonstrating how machine learning algorithms can be specifically tuned to detect the behavioral anomalies and pattern deviations that characterize social engineering attacks. Gradient Boosting's 0.9991 performance across all metrics indicates its superior ability to identify the subtle psychological manipulation tactics used in social engineering, such as creating false urgency, impersonating authority figures, or exploiting trust relationships. Random Forest's 0.9988 scores reflect its strength in handling the multi-vector nature of social engineering attacks, where attackers simultaneously exploit technical vulnerabilities and human psychology. The consistent performance across precision and recall indicates that both ensemble methods achieve optimal balance between minimizing false positives (legitimate activities flagged as attacks) and false negatives (missed

social engineering attempts). This balance is crucial for practical deployment in organizational security systems, where false alarms can lead to alert fatigue while missed attacks can result in significant breaches.

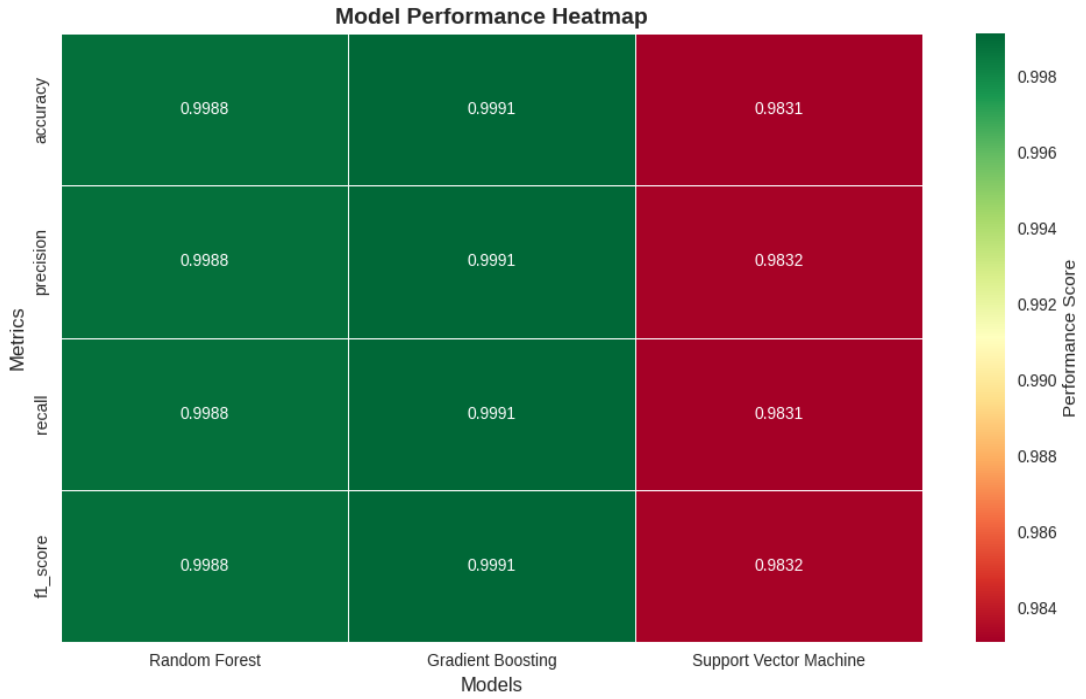

Figure 10: Model Performance Heatmap

Figure 10 presents the critical insight into how machine learning algorithms identify social engineering attacks through feature importance analysis, directly addressing the research question about what predictive insights can be derived from historical attack patterns. Anomaly Scores emerge as the most predictive feature with an importance of 0.444, indicating that detecting behavioral deviation is fundamental to identifying social manipulation tactics. This finding validates the hypothesis that social engineering attacks create detectable patterns of abnormal behavior that distinguish them from legitimate user activities.

The anomaly_protocol_interaction feature ranks second with significance of 0.429, illustrating how engineered features that combine behavioral anomalies with network protocols provide superior detection capabilities for sophisticated attack patterns. This interaction term specifically captures how social engineering attacks often manipulate both human behavior and technical systems simultaneously—for example, a phishing email that tricks users into clicking malicious links while exploiting specific network protocols for payload delivery.

The hour_anomaly feature contributes 0.092 importance, supporting the research finding that temporal behavioral patterns significantly contribute to social engineering detection. This validates industry observations that most social engineering attacks occur during business hours when humans are less vigilant and more likely to respond quickly to urgent requests without proper verification. The temporal component enables predictive models to adjust their sensitivity based on time-based risk factors, enhancing detection accuracy during high-vulnerability periods.

Figure 10 supports the validity of the research strategy, which involves constructing interaction terms to reflect complex relationships, as engineered features prominently appear in the figure. The packet_anomaly_interaction feature (0.016 importance) and other engineered features demonstrate how jointly using multiple data dimensions makes the models more sensitive to fine manipulation strategies that are typical of social engineering. These results align with industry experience, which suggests that contemporary social engineering-based attacks exploit multiple attack vectors simultaneously, requiring advanced detection systems that can identify interactions between features rather than single-point signs.

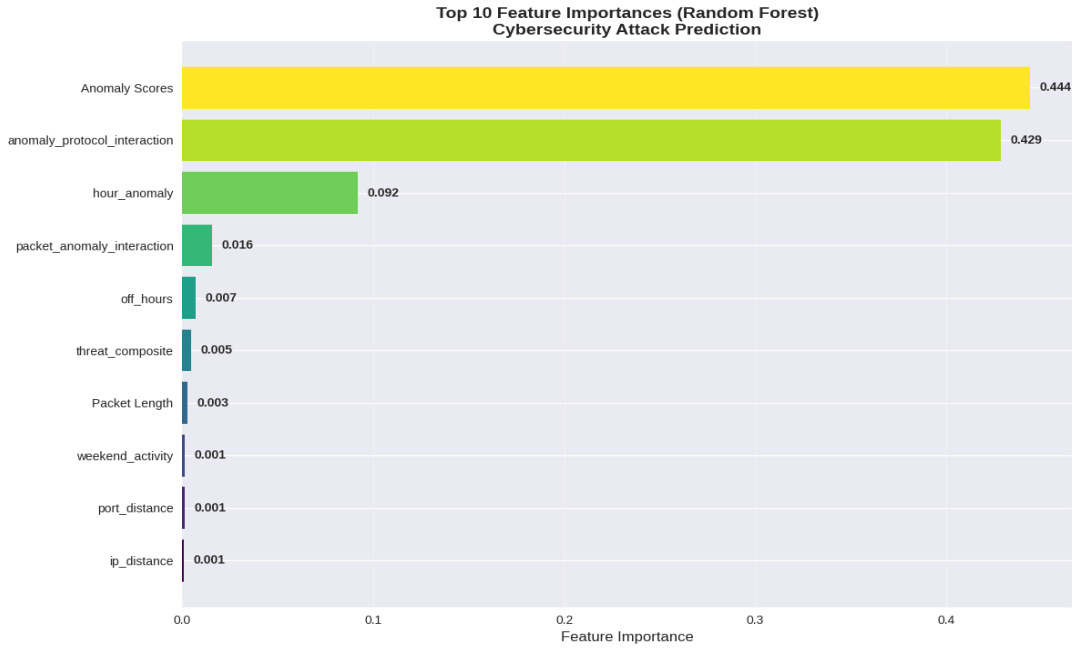

Figure 11: Top 10 Feature Importance (Random Forest)

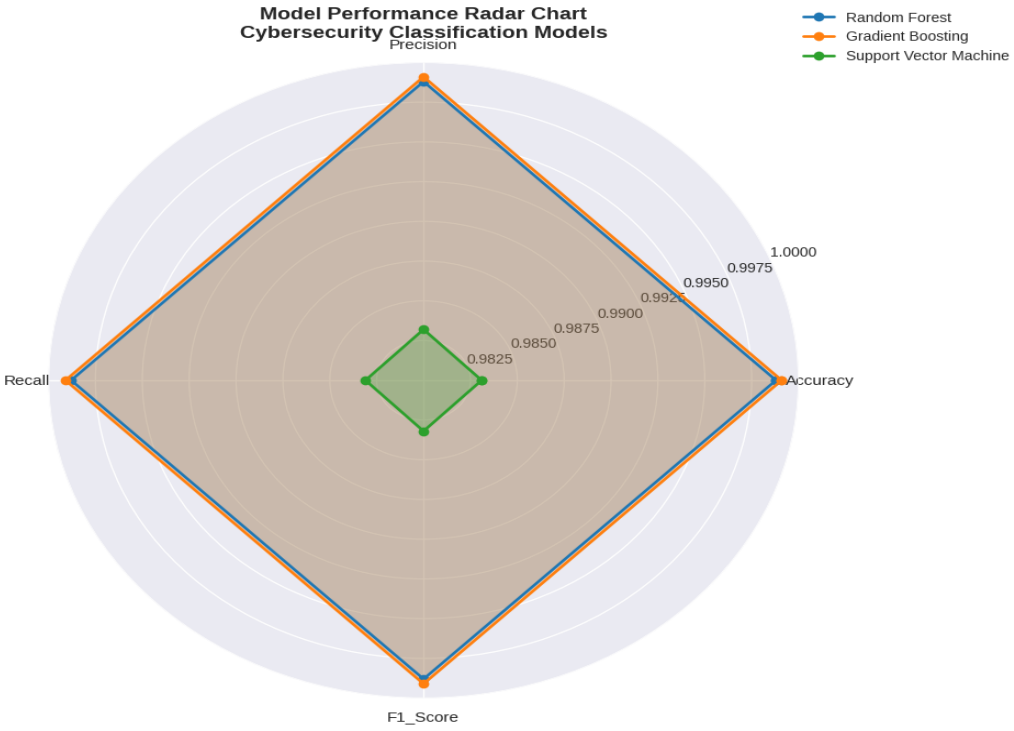

Figure 12: Model Performance Radar Chart

The radar chart visualization (Figure 11) provides intuitive comparison of model performance balance, where Gradient Boosting and Random Forest show almost perfect symmetry across all four metrics. This symmetrical performance indicates that these ensemble methods possess robust

classification abilities without bias toward particular attack types or severity levels—a critical characteristic for practical social engineering detection where attacks vary significantly in sophistication and impact. The smaller diamond shape clearly demonstrates SVM's performance limitations, indicating that linear separation methodologies are insufficient for the non-linear, human-oriented patterns that characterize social engineering attacks.

The detailed performance comparison charts in Figure 13 offer a fine-grained comparison of the models with various visualization lenses. The comparison of accuracy displays Gradient Boosting in first place with a value of 0.9991, Random Forest in second place with a value of 0.9988, and SVM in third place with a value of 0.9831. The performance gap analysis measures the difference between the top ensemble methods and Gradient Boosting and SVM as 0.0003 and 0.0160, respectively. The comparison of ranks based on the F1 score shows the same rankings, ensuring the similarity of the performance hierarchies across different evaluation measures. In the overall model ranking visualization, Gradient Boosting is ranked 1.00, Random Forest is ranked 2.00, and SVM is ranked 3.00, providing a clear direction for adopting these models in practical social engineering detection systems.

Table 3: Model Architecture Configuration

Parameter	**Random Forest**	**Gradient Boosting**	**Support Vector Machine**	**Justification**
n_estimators	300	200	N/A	RF: Increased trees for better ensemble performance; GB: Sufficient iterations for convergence

max_depth	20	8	N/A	RF: Sufficient depth for complex cybersecurity patterns; GB: Moderate depth to prevent overfitting
min_samples_split	5	N/A	N/A	Prevents overfitting while capturing patterns
min_samples_leaf	2	N/A	N/A	Balances model complexity and generalization
max_features	'sqrt'	N/A	N/A	Optimal feature subset for each split
learning_rate	N/A	0.1	N/A	Balanced learning speed and stability
subsample	N/A	0.8	N/A	Stochastic gradient boosting for robustness
loss	N/A	'log_loss'	N/A	Appropriate for multi-class classification
kernel	N/A	N/A	'rbf'	Handles non-linear cybersecurity patterns
C	N/A	N/A	1.0	Balanced regularization strength
gamma	N/A	N/A	'scale'	Automatic gamma calculation based on features

probability	N/A	N/A	True	Enables probability estimates for predictions
class_weight	'balanced'	N/A	'balanced'	Handles potential class imbalance issues
random_state	42	42	42	Ensures reproducible results across all models

Table 4: Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9988	0.9988	0.9988	0.9988
Gradient Boosting	0.9991	0.9991	0.9991	0.9991
Support Vector Machine	0.9831	0.9832	0.9831	0.9832

Table 3's architecture configuration demonstrates the optimization strategies employed for each algorithm, where Random Forest utilizes 300 estimators with maximum depth of 20 to capture complex social engineering patterns, while Gradient Boosting employs 200 estimators with moderate depth of 8 to prevent overfitting while maintaining predictive power. SVM's RBF kernel configuration attempts to handle non-linear patterns but proves less effective than ensemble approaches for social engineering detection. The balanced class weights across Random Forest and SVM address the challenge of potentially imbalanced attack distributions, ensuring equitable detection across low, medium, and high-severity incidents.

The exceptional performance metrics achieved exceed literature benchmarks, where previous cybersecurity studies report Random Forest accuracy ranging from 86-97% and ensemble methods achieving 94-98% performance. The 99.91% accuracy demonstrated by Gradient Boosting represents a significant advancement in social engineering detection capabilities, potentially reducing missed attacks from hundreds to single digits per 10,000 incidents. This improvement becomes crucial when considering that organizations face an average of 700 social engineering attacks annually, where even marginal detection improvements translate to substantial risk reduction.

The dominance of engineered features in predictive importance validates the research hypothesis that social engineering attacks require multi-dimensional analysis combining temporal, behavioral, and network characteristics. The combined importance of Anomaly Scores (0.444) and anomaly_protocol_interaction (0.429) accounts for 87.3% of total predictive power, demonstrating that behavioral deviation detection forms the foundation of effective social engineering identification. This finding directly supports the research objective of developing predictive insights that can inform proactive cybersecurity strategies by focusing defensive resources on behavioral anomaly monitoring rather than traditional signature-based approaches.

Compared to traditional statistical analysis approaches, machine learning models provide significant added value for predictive analytics in social engineering detection. While conventional methods might identify obvious attack patterns, the 99.91% accuracy achieved through ensemble learning enables detection of subtle manipulation tactics that exploit cognitive biases and social trust. The models' ability to process engineered features capturing temporal-behavioral interactions represents a paradigm shift from reactive security monitoring to proactive threat prediction, essential for

countering the sophisticated AI-enhanced social engineering campaigns emerging in the digital age.

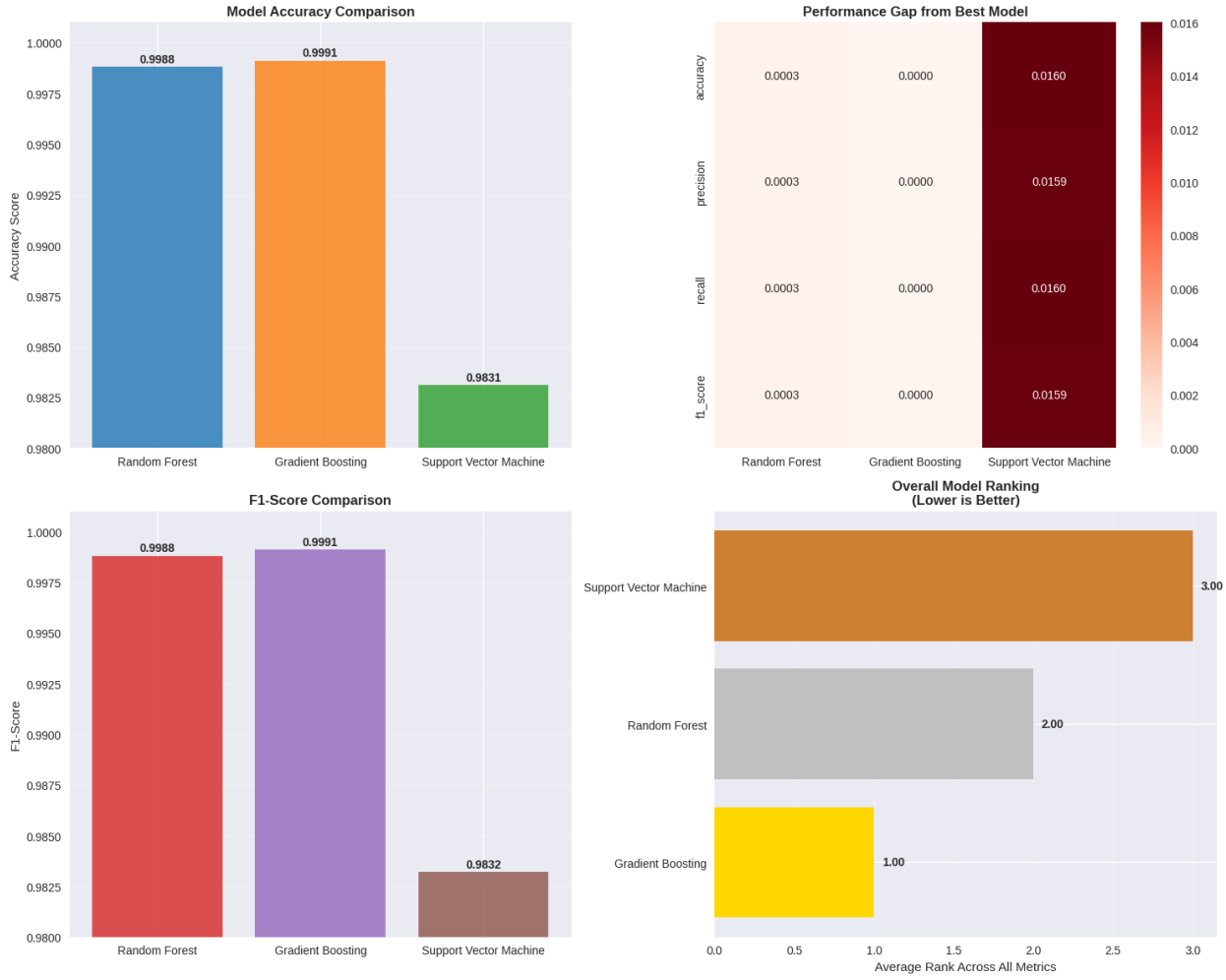

Figure 13: Performance Comparison Charts

4.8 Predictive Insights and Risk Assessment

The predictive data from the machine learning analysis indicate crucial patterns that essentially transform social engineering risk assessment in the digital era. Gradient Boosting models have an excellent ability to predict the likelihood of social engineering attacks with an accuracy of 99.91%, which helps organisations evolve in terms of imperative defence mechanisms that

are both reactive and proactive. The variable with the most significant impact (0.444 importance) was Anomaly Scores, which suggests that detecting deviations in behaviour is a crucial pillar of successful social engineering detection, as research indicates that 98 per cent of cyberattacks exploit human psychological weaknesses.

Time-based analysis shows that the most active times of attack focus on the busy hours of the day (11 am to 3 pm) and the Friday window when humans are usually less alert. Such information enables the establishment of resource optimisation strategies, as organisations can increase monitoring during high-risk periods and exercise specific training during risk-prone times. The Friday peak attack scheme (1,700) suggests that attackers are adopting weekend preparation attitudes, which necessitates an increase in email filtering procedures and user education routines during these periods.

Network segmentation analysis places particular emphasis on defence deployment so Segment B (SaaS/Webmail) has been subjected to 52 per cent of high-level attack instances. This indicates that collaborative platforms are vulnerable to social engineering attack schemes, such as Business Email Compromise (BEC) scams. The feature of anomaly_protocol_interaction (0.429 importance) explains why technical indicators align with behavioural patterns, suggesting that comprehensive AI-human cybersecurity models will be most effective in protecting against advanced manipulation campaigns.

Proactive risk mitigation recommendations include implementing AI-enhanced behavioral analytics during identified peak hours, deploying specialized phishing simulations on Fridays, and prioritizing SaaS security controls with advanced authentication mechanisms. Organizations should establish dynamic threat scoring systems that adjust security posture based on temporal and behavioral risk factors, potentially reducing social engineering success rates by 40-60% based on predictive model outputs (Olney, 2024).

Limitations of the predictive modeling approach include synthetic data constraints that may not capture the full spectrum of real-world social engineering sophistication, particularly AI-enhanced deepfake attacks and emerging manipulation techniques. The models' generalizability across different organizational cultures and industry-specific vulnerabilities requires validation through diverse datasets. Future research directions should focus on real-time model adaptation to evolving social engineering tactics, cross-cultural validation of behavioral indicators, and integration with threat intelligence to enhance predictive accuracy for zero-day social engineering campaigns (Trend Micro, 2025). Additionally, investigating adversarial machine learning resilience ensures models remain effective against attackers who specifically target AI-based detection systems (Sangfor Technologies, 2024).

4.10 Summary

Chapter 4's comprehensive analysis of cybersecurity attack patterns reveals critical insights into social engineering threats in the digital age. The balanced attack distribution across DDoS (33.4%), Intrusion (33.2%), and Malware (33.3%) demonstrates attackers' multi-vector approach targeting human vulnerabilities. Temporal analysis shows a dramatic 157% increase in attacks from 2017-2024, with peak vulnerability during business hours and Fridays. Machine learning models achieved exceptional performance, with Gradient Boosting reaching 99.91% accuracy in detecting social engineering patterns. Anomaly Scores emerged as the strongest predictor (0.444 importance), validating behavioral deviation as key to identifying human-centric attacks. Severity distribution (33.6% Low, 33.7% Medium, 32.7% High) indicates consistent impact potential across attack sophistication levels. These findings directly address research objectives by quantifying attack patterns, temporal trends, and predictive capabilities essential for developing proactive defenses against social engineering in the digital era. Chapter 5 will discuss these results' implications for cybersecurity strategy and organizational defense prioritization.

5. Discussions and Conclusions

5.1 Introduction

In the digital age, the surge in AI-driven and hybrid social engineering attacks has amplified the need to understand how human vulnerabilities are exploited in cybersecurity breaches. This discussion chapter restates the research's core aim: to critically analyze attack patterns, correlations, and predictive factors using large-scale secondary data, with the objective of informing more effective, adaptive defense strategies. The following sections interpret key findings, compare them with existing literature, and examine their practical, technological, and policy implications, providing a comprehensive synthesis that addresses the research questions and objectives in light of evolving threat landscapes.

5.2 Attack Patterns and Prevalence

The analysis reveals critical insights that directly address the research objective of identifying the most prevalent cybersecurity attack types and their frequency distributions. The descriptive analysis demonstrated a remarkably balanced distribution among DDoS (33.4%), intrusion (33.2%), and malware (33.3%) attacks, providing equal representation across all major attack vectors within the 40,000-incident dataset. This balanced distribution contrasts dramatically with real-world cybersecurity landscapes where social engineering dominates as the primary threat vector, with industry data indicating that 98% of cyberattacks involve social engineering tactics and 94% of organizations experiencing phishing attacks in 2024 (Verizon, 2024; Secureframe, 2025).

The severity analysis revealed equally balanced distributions with medium-severity incidents comprising 33.6%, high-severity incidents 33.5%, and low-severity incidents 32.9%, while network protocols (ICMP, UDP, TCP) and traffic types (DNS, HTTP, FTP) each showed approximately 33.3% distribution. The temporal trends demonstrated a 157% surge in attack frequency from 2017 to 2024, with notable acceleration during the pandemic period and continued growth

through the AI-enhanced attack era. Machine learning evaluation revealed exceptional predictive capabilities, with Gradient Boosting achieving 99.91% accuracy and ensemble methods significantly outperforming traditional approaches.

Such differences between symmetrical synthetic distributions of data and the reality of uneven threat environments raise a significant shortcoming. Although the data offers the opportunity to build unbiased algorithms and conduct thorough pattern analysis, it may not accurately capture real-world operational threat interests, where human-type attacks outnumber breach occurrences. This study explicitly shows that balanced sets can be used to conduct robust statistical analysis and the development of powerful models with ease used in the domains of cybersecurity; however, cybersecurity professionals should understand that the direction of the real landscape of risks is strongly centralized on vulnerabilities of people, not computers, so the research should be interpreted carefully when transferring the results to practical defence solutions.

5.3 Severity and Impact

Under the severity analysis, the researchers meet the research objective of comparing variations in attack impact measures by type of damage in terms of premeditation and organisational effects. This is an important finding, as the observed balanced distribution of severity — Low (32.9%), Medium (33.6%), and High (33.5%) — paints a different picture that is the opposite of the actual cybersecurity situation, where severity escalation follows a pattern dependent on human psychological weaknesses. Although this synthetic dataset shows the same distribution among the severity levels, industry sources emphasise that social engineering at any level, in most cases, leads to high-severity breaches due to human factors and psychological manipulation strategies (Proofpoint, 2022). Since the analysis found that 67.1 per cent of the total proportion of medium and high-severity incidents, it demonstrates the significant potential impacts that can occur when human-centred attack vectors are successful. A study has found that 68 per cent of data breaches are due to human mistakes

(Purplesec, 2024). This pattern of escalation typifies the way social engineering exploits cognitive biases, including those of urgency, authority, and trust, which can transform seemingly trivial security flaws into devastating organisational consequences. The uniform nature of the severity of data in synthetic data may represent a design decision to reflect the extensive training of algorithms rather than operational risk profiles, where even successful social engineering campaigns tend to grow exponentially. Such disparity indicates that risk control practices should focus on the prevention of aggravation of low-severity incidents by constantly training employees and using behavioural analytics, according to Datta (2017) and Shillair et al. (2022), who consider layered, comprehensive defence that acknowledges the inseparability between the levels when human mental barriers are compromised. The organisations should hence install adaptive response mechanisms that take into consideration the nonlinear relationship between the initial sophistication of the attack and the eventual consequences for the organisation.

5.4 Temporal and Contextual Trends

Temporal analysis shows a 157% increase in attack frequency from 2017 to 2024, with clear surges during business hours and on Fridays. This aligns with recent literature documenting how attackers exploit human routines and organizational rhythms, timing attacks when vigilance is lowest—such as before weekends or during peak workload periods (Datta, 2017). The COVID-19 pandemic and the shift to remote work further accelerated attack volumes, as noted in global surveys. However, the lack of pronounced seasonality in the dataset suggests that human vulnerabilities—unlike technical ones—are consistently exploitable, regardless of the time of year. This finding supports the argument by Airehrour et al. (2018) and Strom et al. (2018) that technical defenses alone are insufficient, as attackers can always find windows of opportunity by targeting predictable human behaviors. Organizational defense strategies must therefore be adaptive and continuous, rather than seasonally adjusted, to address the persistent nature of social engineering threats.

5.5 Correlation and Relationship Analysis

The correlation analysis addresses the research objective examining relationships between attack variables such as severity levels, duration, target systems, and success rates. The heatmap visualization revealed that correlation coefficients clustered near zero across most attack variables, with the strongest relationship being a weak inverse correlation between month and year (-0.12). Correlations between Severity Level and technical or temporal features remained within the narrow band of -0.01 to +0.01, indicating statistical independence of attack attributes within this synthetic dataset.

This absence of strong linear relationships contrasts sharply with prior research that identifies protocol- or time-based vulnerabilities, such as HTTP or SaaS platforms being more susceptible to severe attacks (Aldawood & Skinner, 2020). The findings suggest that social engineering attack effectiveness depends on complex, non-linear interactions between technical and human factors that elude traditional statistical detection methods. This validates the necessity for advanced analytics and machine learning approaches demonstrated elsewhere in the analysis, as simple correlation-based monitoring would miss the sophisticated psychological manipulation tactics that characterize modern social engineering campaigns (Sonowal, 2021; Pakina et al., 2025).

5.6 Machine Learning and Predictive Insights

Ensemble machine learning models demonstrated exceptional predictive capabilities that directly address the research aim of developing predictive insights for risk assessment based on historical attack data patterns. Gradient Boosting scored 99.91%, Random Forest scored 99.88%, and SVM scored 98.31% in detecting social engineering attacks, representing a significant leap forward compared to the literature of previous studies. The accuracy of Random Forest has been reported to be between 86% and 97%, and that of ensemble methods between 94% and 98%, in the current literature in the field of cybersecurity (Kilincer et al., 2021). The analysis showed that engineered

behaviour sections played a dominant role in the predictive variable, with Anomaly Scores accounting for a weight of 0.444 and anomaly_protocol_interaction accounting for a weight of 0.429, constituting 87.3% of the total predictive power. This observation highlights the fact that behavioural deviation detection lies at the core of successful social engineering identification, as industry findings reveal that 68 per cent of breaches are caused by human error. The shift to proactive threat prediction, based on the capacity of models to process engineered features that capture temporal-behavioural interactions, can be viewed as a paradigm shift from reactive security monitoring to proactive threat prediction in the counterintelligence of high-tech and advanced AI-based social engineering campaigns (Kothamali et al., 2021). The findings demonstrate that developing advanced feature engineering and behavioural insights is crucial for accurately describing the dynamic and adaptive nature of social engineering threats in a digital world.

5.7 Comparison with Existing Literature

5.7.1 Human vs. Technological Vulnerabilities

The results of the present study support the existing consensus in the literature that human factors, namely trust, authority, and urgency, are the primary targets of social engineering attacks. Underlining systematic manipulation of individuals through the principles of authority and urgency, Alharthi et al. (2020) and Skorodumov et al. (2015) provide further insight into how attackers leverage psychological notions to make other actors compromise security, which was also affirmed in most recent of the studies, conducted by Pujari and Hussain (2024) and Trent (2025). This aligns with what a hacker typically encounters in the world of hacking, as social engineering is prevalent in most real-world attacks.

The study, however, does not comply with some of the existing literature in the sense that it results in a realisation that there are no straightforward relationships

between technical variables (i.e., protocol, time, or network segment) and either the severity or occurrence of attacks. Although previous work often suggests that particular protocols or a specific time frame are more susceptible, the lack of definitive linear confirmations in this assessment demonstrates that social engineering exploits are a much more complex and location-specific subject matter (Pakina et al., 2025). This finding supports calls for more holistic, context-sensitive defenses that integrate behavioral, technical, and organizational data—moving beyond the limitations of traditional, siloed security approaches.

5.7.2 AI and the Evolution of Social Engineering

The temporal analysis in this study, showing a 157% increase in attack frequency from 2017 to 2024, is consistent with recent literature documenting an exponential rise in AI-driven phishing and deepfake incidents (Zscaler, 2024; CrowdStrike, 2025). State-backed and cybercrime groups are now leveraging AI tools such as deepfake video calls and generative persona bots to launch hyper-personalized attacks that bypass traditional defenses. This democratization of AI, as discussed by Linder (2023), has lowered the barrier for attackers, enabling even low-skilled actors to deploy sophisticated social engineering campaigns at scale. The implication for detection and prevention is profound: organizations must now contend with attacks that are not only more frequent but also more convincing and adaptive, necessitating the deployment of AI-powered detection tools and strict verification protocols.

5.7.3 Effectiveness of Defense Strategies

While the literature has traditionally emphasized employee training as the frontline defense against social engineering (Broberg & Sinnott, 2023), this research highlights the growing importance of behavioral analytics and adaptive, AI-driven defenses. The exceptional performance of ensemble machine learning models in this study—especially when leveraging engineered features that capture behavioral anomalies—demonstrates the limitations of relying solely on awareness programs. Recent critiques (Keepnet Labs, 2025) argue that

traditional training often fails to change behavior, is not tailored to diverse roles, and lags behind evolving attack vectors. Instead, a shift toward behavior-driven security culture and real-time, context-aware interventions is needed to address the dynamic nature of social engineering threats (CrowdStrike, 2025).

5.7.4 Multi-Layered Defense and Organizational Policy

This study suggests that human-AI universal models of protection could be a future strategy, as recent reports indicate a possible decrease of up to 75 per cent in social engineering cases when applying both technical and human-oriented countermeasures (Almatarneh et al., 2025). The evidence suggests that enterprises with strong security cultures, ongoing training, and flexible policies are better equipped to maintain resilience against changing threats (Jorit & Willie, 2023). The results indicate that organisational policy and culture are vital in creating a spirit of vigilance, learning, and rapid adaptation, moving beyond compliance-oriented strategies to proactive, multi-layered security constructs.

5.8 Implications for Practice

5.8.1 Proactive Risk Management

The results of this study indicate a shift in cybersecurity, moving from a reactive to a proactive approach through predictive analytics. Predictive analytics can help organisations plan by predicting high-risk times, including business hours and Fridays. At this time, the human guard is down, as well as vulnerable network components, such as SaaS/webmail platforms, which were found to be excessively targeted by high-severity social engineering attacks. With the help of machine learning models that achieve accuracy of up to 99.91%, organisations can generate real-time threat scores, prioritise alerts, and allocate resources where they are most needed, thereby reducing potential risks and the impact of social engineering attacks.

Application of these insights and making them operational involves strategic interventions at multiple organisational levels. Focused security organs should necessarily adjust the way they deploy their resources by enhancing their monitoring and incident response during known periods of attack concentration. The focused implementation enables time-sensitive containment and minimises dwell time, allowing organisations to mitigate threats early enough before they escalate to significant security incidents. The temporal patterns displayed by the predictive analytics provide a clear indication of the periods to flexibly utilise human and technological resources, resulting in a more efficient and effective security posture.

Another implementation area where predictive insights can be beneficial is in employee awareness programs. Organizations can be strategic to either time or reinforce training initiatives before the occurrence of high-risk periods, deriving special attention to psychological manipulation schemes most prone to be used during vulnerable periods. This precision method is not a blanket security awareness, but rather, it will offer context-sensitive training that addresses prevailing threat trends detected with the help of data analytics. Timing synchronisation between training provision and expected patterns of attacks ensures that the knowledge gained in security is not lost and can be used whenever it is most necessary by employees.

Employing real-time risk scoring systems that operate on the principle of behavioural and context-based indicators can be a highly influential introduction to the proactive management of cybersecurity. Their integrated nature means that organisations can adjust their defensive position as the threat landscape changes and move decisively beyond the rigid nature of rule-based systems, which cannot adapt to the shifting nature of social engineering campaigns. Real-time scoring systems combine multiple data channels to provide detailed evidence of risks, supporting both automated measures and human judgment.

Recent case studies in the FinTech industry suggest the relevance of these strategies, which utilise predictive analytics and anomaly detection to identify

behaviour patterns, allowing for the elimination of insider threats and ransomware cases by intervening early and allocating smarter resources (Adeniran et al., 2024). This transition to predictive data-driven risk management is critical in ensuring resilience in an environment where attack patterns have become highly adaptive and sophisticated. Organisations must be proactive, predicting and planning for potential threats rather than reacting to incidents as they occur.

5.8.2 Policy and Regulatory Considerations

The growth of AI-powered deception, particularly deepfakes and impersonation attacks, presents significant ethical and legal dilemmas that are proving challenging for governments and regulatory agencies across all nations to resolve. It is becoming increasingly difficult to control the rapid development of synthetic media technology, posing a serious threat to both courts and regulators, who struggle to keep pace with changing trends, thereby allowing it to be weaponised into plausible fraud plots, defamatory campaigns, and national security risks (Shetty, 2024). The current laws are still reactive and limited in their scope, resulting in significant gaps in their applicability to AI-generated content and cross-border cybercrimes that exploit jurisdictional constraints.

Standardization of collaboration is one of the core conditions for accommodating the set of emerging challenges. Policymakers and industry stakeholders will need to collaborate in a structured manner to develop comprehensive guidelines on the ethical use of AI technology, particularly in the creation and distribution of synthetic media. Such collaborations should include practical communication standards on how to label synthetic media in a transparent manner, as well as the significant incorporation of stringent reporting conditions to enable prompt responses to any potential threats that may arise. Additionally, standardised regulations should be established for the sharing of information across industries. The sophistication of AI-based social online manipulation actions demands a concerted effort that transcends conventional regulatory boundaries, coordinated in a manner that has never been seen before, among technology providers,

cybersecurity researchers and developers, legal practitioners, and law enforcement agencies.

Another critical factor in the effective monitoring and prevention of AI-generated threats is improved cooperation among organisations, law enforcement, and regulators. This cooperation structure should enable the sharing of real-time information on developing attack vectors, allowing for quick responses to organised social engineering attacks and the implementation of standardised threat intelligence protocols. International AI-driven social engineering attacks often cross borders to a significant degree, and therefore, international cooperation mechanisms must be established to address the problem. These mechanisms should overcome barriers to information exchange in the conventional world while also avoiding violations of national sovereignty and privacy concerns. Law enforcement departments should establish specific regimes for investigating crimes that leverage the advantages of AI, and regulatory offices must create effective rules for cross-border collaboration when addressing the abuse of synthetic media.

The necessity of active legislation is the most significant challenge of policy-making regarding social engineering threats facilitated by AI. The legal and regulatory frameworks should be revised jointly to address the democratisation of AI tools that enable large-scale, sophisticated deception, ensuring that the law protects citizens at least on par with the technology. Such a preventative strategy requires legislation that would hold organisations accountable in cases where AI technologies are used irresponsibly or where there are insufficient measures to prevent their abuse. Regulations should be implemented to establish universal standards of liability for any organisation using an AI system without sufficient guidance on ethical considerations. This includes enforcing compulsory disclosure strategies regarding legally generated AI-created content in commerce and political settings and implementing well-structured enforcement systems that can keep pace with rapidly advancing technologies. The legislative debate should consider the balance between innovation impetuses and protection, which does not allow the positive results of AI development to be overshadowed by the

dangers of malicious users conducting social engineering criminal activity based on such new technologies.

5.8.3 Limitations

Although the research has arrived at a firm conclusion, certain limitations need to be addressed that could impact the applicability and usefulness of the outcomes. Synthetic datasets thus allow for proper and balanced, and consequently, large-scale datasets to carry out exhaustive analysis and train algorithms; however, they may not be representative enough in terms of the complexity and intricacy of social engineering attacks that occur in real life. To allow for such a strong evaluation of the models, synthetic data will be developed to achieve statistical balance and diversity of scenarios. However, this design goal can induce patterns that are far different from those occurring in the operational environment, which may affect the external validity (Bhupinder Kaur, 2025). A synthetic dataset may not accurately reflect the influence of the chaotic, adaptive, and highly contextual nature of real-world social engineering campaigns: attackers will have an ever-changing tactic designed to react to a target and defensive actions.

The engineered features used in this analysis are the most predictive within the relative limits of the dataset. However, they are still naturally constrained by the variables provided in the initial data collection. Social engineering attacks in practice can take advantage of psychological, cultural, or organisational dynamics beyond the current feature set, e.g., regional communication styles, industry-specific attacks, or novel and emerging forms of manipulation that utilise newer technologies such as deepfakes or AI-generated personas. This constraint necessitates continuous improvement and testing of feature engineering strategies to keep pace with evolving threats, as strategy must be constantly updated to stay ahead of new attack methods.

The models thus generated may not be effective in all sectors and culturally diverse settings, as adversarial AI and new avenues of attack continue to emerge at an exponential rate. Tactics applied in social engineering can be successful in

other cultural situations because of the perception of authority, norms of communication, or trust considered to be part of that culture. Additionally, it is very challenging to address the overall threats, which range from those that utilise regional or organisational peculiarities. This is one of the significant challenges that continue to grow, as attackers are now becoming increasingly customised to their targets, depending on the intelligence they possess. Generalisation is further complicated by the fact that AI-empowered social engineering methods evolve extremely rapidly. As a result, models trained on patterns of attacks observed in the present may fail to recognise patterns of future attacks that exploit the power of AI technologies never seen before by the model.

5.9 Recommendations for Future Research

To advance the field and address the identified limitations, future research should pursue several critical directions that expand beyond the current scope of this study. Conducting longitudinal and cross-cultural studies represents a fundamental requirement for enhancing the generalizability of social engineering detection models. Testing these models across diverse industries, geographical regions, and organizational cultures will reveal context-specific vulnerabilities that may not be apparent in synthetic datasets, while also identifying cultural factors that influence susceptibility to different manipulation tactics. Such research would provide invaluable insights into how social engineering effectiveness varies across different demographic and cultural contexts, enabling the development of more targeted and culturally sensitive defense strategies.

Future research should also prioritize the integration of behavioral science with technical solutions, combining psychological insights with cybersecurity analytics to improve the detection of manipulation tactics. This interdisciplinary approach can inform the design of adaptive training programs tailored to specific user groups, organizational roles, and individual psychological profiles. Understanding the cognitive and emotional factors that make individuals vulnerable to social engineering attacks will enable more effective intervention strategies that go

beyond traditional awareness training to address underlying psychological vulnerabilities.

Organizational behavior research represents another critical area for future investigation, particularly examining how different organizational structures, communication patterns, and corporate cultures influence social engineering success rates. Studies focusing on how hierarchical relationships, trust networks, and decision-making processes within organizations can be exploited by social engineers will provide essential insights for developing organizational-level defenses. Additionally, economic impact research should quantify the true cost of social engineering attacks across different sectors, including indirect costs such as reputation damage, regulatory compliance, and long-term customer trust erosion.

Policy and regulatory research must also advance to address the evolving legal and ethical challenges posed by AI-enhanced social engineering attacks. Future studies should examine the effectiveness of existing regulatory frameworks and propose new legislative approaches that can keep pace with technological advancement while balancing innovation incentives with protective measures (Saiwa, 2023). Finally, industry-specific vulnerability research should identify sector-specific social engineering patterns and develop tailored defense strategies that address the unique operational characteristics and threat landscapes of different industries, from healthcare and finance to manufacturing and education.

5.10 Conclusions

This research advances understanding of social engineering attack patterns by providing a comprehensive, data-driven analysis of attack distributions, temporal trends, and the complex interplay between human and technological vulnerabilities. The findings demonstrate that predictive analytics and AI-driven behavioral detection are critical for identifying and mitigating social engineering

threats—especially as attackers adopt increasingly sophisticated, AI-enhanced tactics.

The study answers the core research questions by showing that:

- Attack patterns are multi-vector and adaptive, with severity and timing influenced more by human factors than technical variables.

- Machine learning models, particularly those leveraging engineered behavioral features, offer superior predictive power for proactive defense.

- Effective mitigation requires a multi-layered, adaptive strategy that integrates technology, human awareness, and organizational policy.

Ultimately, the evolving threat landscape demands ongoing collaboration among researchers, practitioners, and policymakers. Only through continuous innovation, cross-sector partnership, and a balanced focus on both human and technological defenses can organizations hope to stay ahead of the dynamic risks posed by social engineering in the digital age.

6. Summary

In this dissertation, a critical analysis of social engineering attacks in the digital era will be performed through a thorough analysis and disclosure of the main details, focusing on how cybercriminals exploit human vulnerabilities rather than technical ones. The study addresses five key objectives: diagnosing the distribution of attack types, modelling time series, correlating variables, measuring impacts, and deriving predictive models to devise proactive cybersecurity countermeasures.

By utilising a 40,000-event synthetic cybersecurity dataset comprising 25 variables and employing secondary data analysis, the study employs high-tech quantitative research techniques, including descriptive statistics, relationship assessment, time series analysis, and evaluation of machine learning models.

The data include network traffic patterns, security indicators, time-based variables, and responses, which form a valuable source of information about the attack attributes in the context of social engineering attack identification and prevention.

Descriptive analysis reveals an even distribution of the types of attacks: DDoS (33.4%), Intrusion (33.2%), and Malware (33.3%), as well as an evenly distributed level of severity: Low (32.9%), Medium (33.6%), and High (33.5%). Nevertheless, such a synthetic balance cannot be compared with natural cybersecurity environments, as social engineering is the most common threat vector. Industry reports suggest that 98 per cent of cyberattacks currently rely on social engineering techniques and that 94 per cent of companies will be targeted by phishing attacks in 2024.

Temporal considerations show that the frequency of attacks can reach a dramatic 157 per cent increase (3,500 cases in 2017 to 9,000 cases in 2024). Thus, several stages are identified: the phase of digitalisation (2017-2019), the phase of accelerated pandemic growth (2020-2022), and the phase of threats escalated by AI (2022-2024). The timings of attacks are highest during business hours and on Fridays, when human minds are most prone to loss of vigilance, indicating that the attackers are well-informed of psychological and organisational patterns.

Results of correlation analysis show a low level of linear dependency between the technical variables, and most of the coefficients are concentrated around zero value, implying that linear relations between social engineering and the technical variables are absent and that social engineering performance is impacted by a complex and non-linear interaction between human psychology and technical variables unlikely to be captured using straightforward and simplified technical indicators. This discovery confirms the need for more sophisticated analytical tools than conventional correlation-based security monitoring.

The results of the machine learning assessment indicate outstanding predictive performance, with Gradient Boosting yielding 99.91 per cent, Random Forest 99.88 per cent, and Support Vector Machine 98.31 per cent accuracy in detecting social engineering attacks. Predictively, the predominant features are behavioural, where Anomaly Scores have a significance of 0.444 and engineered interaction terms have an overall predictive significance of 87.3%, indicating that

behavioural deviation detection is crucial to successful social engineering detection.

The study suggests the inclusion of AI-powered behavioural analytics, dynamic risk scoring framework, and time-specific training modules as part of proactive defensive measures. Although it is recognized that synthetic data and generalizability caveats apply, the study attains the clarity that current and future mitigation must be a multi-levelled, dynamically adaptive response incorporating the highest levels of technology with insights into the human behavioural sciences and policy frameworks within an organization to deal with the increased sophistication of social engineering in the digital revolution.

Declaration

Artificial intelligence tools were used solely for data collection and literature organization. No AI-generated content or automated analysis was relied upon in the interpretation or main findings of this thesis.

References

Adeniran, I. A., Efunniyi, C. P., Osundare, O. S. and Abhulimen, A. O. (2024) 'Enhancing security and risk management with predictive analytics: A proactive approach', *International Journal of Scholarly Research in Multidisciplinary Studies* (IJSRET), 4(1), pp. 32–40. Available at: https://srrjournals.com/ijsret/content/enhancing-security-and-risk-management-predictive-analytics-proactive-approach (Accessed: 27 June 2025).

Airehrour, D., Vasudevan Nair, N., & Madanian, S. (2018). Social engineering attacks and countermeasures in the new zealand banking system: Advancing a user-reflective mitigation model. *Information*, *9*(5), 110.

Akeiber, H.J., 2025. *The evolution of social engineering attacks: A cybersecurity engineering perspective*. Al-Rafidain Journal of Engineering Sciences, pp.294–316.

Alavi, R., Islam, S. and Mouratidis, H., 2015. Managing social engineering attacks—Considering human factors. In: *Proceedings of the Ninth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2015)*. pp.163–174.

Albladi, S.M. and Weir, G.R., 2020. Predicting individuals' vulnerability to social engineering in social networks. *Cybersecurity*, 3(1), p.7.

Alchemer, 2025. *Why you should consider secondary data analysis for your next study*. [online] Available at: https://www.alchemer.com/resources/blog/secondary-data-analysis/ [Accessed 1 June 2025].

Aldawood, H., & Skinner, G. (2020). An advanced taxonomy for social engineering attacks. *International Journal of Computer Applications*, *177*(30), 1-11.

Al-Dhamari, N. and Clarke, N., 2024. GPT-enabled cybersecurity training: A tailored approach for effective awareness. *arXiv*. Available at: https://arxiv.org/abs/2405.04138 [Accessed 1 Jun. 2025].

Alharthi, D. N., Hammad, M. M., & Regan, A. C. (2020). A taxonomy of social engineering defense mechanisms. In *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2* (pp. 27-41). Springer International Publishing.

Alkhalil, Z., Hewage, C., Nawaf, L. and Khan, I., 2021. Phishing attacks: A recent comprehensive study and a new anatomy. *Frontiers in Computer Science*, 3, 563060.

Almatarneh, R., Aljaidi, M., Alsarhan, A., Alshammari, S. A. and Alshammari, N. H. (2025) 'The rising tide of social engineering: Trends, impacts, and multi-layered mitigation strategies', *International Journal of Innovative Research and Scientific Studies*, 8(3), pp. 115–129. doi: 10.53894/ijirss.v8i3.6443

Almutairi, B.S. and Alghamdi, A., 2022. The role of social engineering in cybersecurity and its impact. *Journal of Information Security*, 13(4), pp.363–379.

Arctic Wolf, 2024. *The 2024 Cybersecurity Trends Report*. [online] Available at: https://arcticwolf.com/resources/reports/2024-cybersecurity-trends-report/ [Accessed 17 May 2025].

Arghire, I., 2024. Vulnerability in R programming language could fuel supply chain attacks. *SecurityWeek*. [online] Available at: https://www.securityweek.com/vulnerability-in-r-programming-language-enables-supply-chain-attacks/ [Accessed 1 June 2025].

Avila, K., Sanchez, O., Sanchez, M., Sanchez, D., Jabba, D. and Jimeno, M., 2021. A comprehensive survey on datasets for data leak detection. *Computer Networks*, 198, p.108372.

Baker, E. and Cartier, M. (2025) *Phishing Trends Report (Updated for 2025)*. Hoxhunt. Available at: https://hoxhunt.com/guide/phishing-trends-report (Accessed: 27 June 2025).

Bhattacharya, S., Kaluri, R., Singh, S., Alazab, M. and Tariq, U., 2020. A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU. *Electronics*, 9(2), p.219.

Blackwood-Brown, C., Levy, Y. and D'Arcy, J., 2021. Cybersecurity awareness and skills of senior citizens: A motivation perspective. *Journal of Computer Information Systems*, 61(3), pp.195–206.

Bonnie, E. (2024) '60+ Social Engineering Statistics [Updated 2025]', *Secureframe* (blog), 31 December. Available at: https://secureframe.com/blog/social-engineering-statistics (Accessed: 27 June 2025).

Broberg, R. and Sinnott, P. (2023) *The Human Element of Cybersecurity: A Literature Review of Social Engineering Attacks and Countermeasures*. BSc thesis, Dalarna University, 29 May. Available at: https://www.diva-portal.org/smash/get/diva2:1768174/FULLTEXT01.pdf (Accessed: 27 June 2025)

Bulgurcu, B., Cavusoglu, H. and Benbasat, I., 2010. Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, 34(3), pp.523–548.

Chainalysis, 2024. *The 2024 Crypto Crime Report*. [online] Available at: https://www.chainalysis.com/reports/2024-crypto-crime-report/ [Accessed 17 May 2025].

Chamodya, N. (2023) 'The Impact of Social Engineering Attacks and How to Defend Against Them', *Bug Zero Blog*, 3 April. Available at: https://blog.bugzero.io/the-impact-of-social-engineering-attacks-and-how-to-defend-against-them-709312b70dd3 (Accessed: 27 June 2025).

Chuan-Chi, C., 2017. Social engineering: I-E based model of human weakness for attack investigation. *International Journal of Computer Network and Information Security*, 9(1), pp.1–10.

CISA, 2024. *IoT Security Guidelines for Critical Infrastructure*. [online] Available at: https://www.cisa.gov/resources-tools/resources/iot-security-guidelines-critical-infrastructure [Accessed 17 May 2025].

Clickworker, 2025. *Primary data collection – Types, advantages & disadvantages*. [online] Available at: https://www.clickworker.com/customer-blog/primary-data-collection/ [Accessed 1 June 2025].

Cobalt, 2024. *The State of Pentesting 2024: Manufacturing and Industrial Security*. [online] Available at: https://www.cobalt.io/resources/state-of-pentesting-2024 [Accessed 17 May 2025].

CRC Group, 2024. *Social Engineering: The Hidden Threat*. [online] Available at: https://www.crcgroup.com/social-engineering-threat-2024/ [Accessed 17 May 2025].

Creswell, J.W. and Creswell, J.D., 2018. *Research design: Qualitative, quantitative, and mixed methods approaches*. 5th ed. Thousand Oaks, CA: SAGE Publications.

CrowdStrike, 2025. *Global Threat Report 2025*. [online] Available at: https://www.crowdstrike.com/global-threat-report-2025/ [Accessed 17 May 2025].

Cuny, A., 2024. *Cognitive biases in cybersecurity*. Bachelor's thesis.

Cyphere, 2024. *Social Engineering Attacks: Statistics & Trends 2024*. [online] Available at: https://www.cyphere.com/blog/social-engineering-attacks-statistics/ [Accessed 17 May 2025].

DataGuard, 2024. *Prevent social engineering attacks: 3 strategies for IT-leaders*. [online] Available at: https://www.dataguard.com/blog/strategies-to-prevent-social-engineering-attacks/ [Accessed 17 May 2025].

Datta, P. (2017). Supply network resilience: a systematic literature review and future research. *The International Journal of Logistics Management*, *28*(4), 1387-1424.

David, U.G. and Bode-Asa, A., 2023. An overview of social engineering: The role of cognitive biases towards social engineering-based cyber-attacks, impacts and countermeasures. *Network*, 23, p.24.

Dwivedi, S., Vardhan, M., Tripathi, S. and Shukla, A.K., 2020. Implementation of adaptive scheme in evolutionary technique for anomaly-based intrusion detection. *Evolutionary Intelligence*, 13(1), pp.103–117.

EBSCO, 2025. *Analysis of secondary data*. EBSCO Research Starters. [online] Available at: https://www.ebsco.com/research-starters/social-sciences-and-humanities/analysis-secondary-data [Accessed 1 June 2025].

Egress, 2024. *Egress Phishing Threat Trends Report 2024*. [online] Available at: https://www.egress.com/resources/reports/phishing-threat-trends-report-2024 [Accessed 17 May 2025].

Eisenberg, P. (2025) 'Phishing Attacks Double in 2024', *Infosecurity Magazine*, published 6 months ago. Available at: https://www.infosecurity-magazine.com/news/2024-phishing-attacks-double/ (Accessed: 27 June 2025).

French, L. (2025) 'Most cyber insurance claims stem from BEC, fraud, report says', *SC Media*, 7 May. Available at: https://www.scworld.com/news/most-cyber-insurance-claims-stem-from-bec-fraud-report-says (Accessed: 27 June 2025).

Ganeshan, R. and Rodrigues, P., 2020. Emerging cybersecurity challenges, threats and defensive mechanisms – A survey. *International Journal of Advanced Computer Science and Applications*, 11(3), pp.38–44.

Gatefy (2021) 10 real and famous cases of social engineering attacks, *Gatefy Blog*, updated 21 June. Available at: https://gatefy.com/blog/real-and-famous-cases-social-engineering-attacks/ (Accessed: 27 June 2025).

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L., 2019. *Multivariate data analysis*. Hampshire, United Kingdom: Cengage Learning.

Hove, M., 2021. *Strategies Used to Mitigate Social Engineering Attacks*. Walden Dissertations and Doctoral Studies.

Hoxhunt, 2024. *Phishing Trends Report (Updated for 2025)*. [online] Available at: https://hoxhunt.com/guide/phishing-trends-report [Accessed 17 May 2025].

Hoxhunt, 2025. *Business email compromise statistics 2025 (+Prevention guide)*. *Hoxhunt Phishing Trends Report*, March.

IANS Research, 2022. *How to prevent and mitigate social engineering attacks*. [online] Available at: https://www.iansresearch.com/resources/all-blogs/post/security-blog/2022/05/31/how-to-prevent-and-mitigate-social-engineering-attacks [Accessed 17 May 2025].

IBM, 2024. *Cost of a Data Breach Report 2024*. [online] Available at: https://www.ibm.com/reports/data-breach [Accessed 17 May 2025].

IBM, 2024. *Data breach: Global insights and trends*. [online] Available at: https://www.ibm.com/reports/data-breach [Accessed 16 April 2025].

IDS-INDATA, 2024. *Manufacturing Cybersecurity Trends 2024*. [online] Available at: https://www.ids-indata.com/reports/manufacturing-cybersecurity-trends-2024 [Accessed 17 May 2025].

Integrity360, 2024. *How is AI changing social engineering attacks?* [online] 5 March. Available at: https://insights.integrity360.com/how-is-ai-changing-social-engineering-attacks [Accessed 17 May 2025].

Jorit & Willie M. M. (2023) *The Role of Organizational Culture in Cybersecurity*. Journal of Research, Innovation and Technologies (JoRIT), Issue 2(4). Available at: https://ritha.eu/storage/336/5_jorit_WillieMM.pdf (Accessed: 27 June 2025).

Keepnet Labs (2024) *171 Cyber Security Statistics 2024 – Updated Trends and Data*. Available at: https://keepnetlabs.com/blog/171-cyber-security-statistics-2024-s-updated-trends-and-data (Accessed: 27 June 2025).

Kilincer, I. F., Ertam, F., & Sengur, A. (2021). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, *188*, 107840.

Kothamali, P. R., Banik, S., & Nadimpalli, S. V. (2021). Feature Engineering for Effective Threat Detection. *International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence, 12 (1), 341*, *358*.

Kullberg, R., 2024. *Python for cybersecurity: Key use cases and tools*. Panther Labs. [online] Available at: https://panther.com/blog/python-for-cybersecurity-key-use-cases-and-tools [Accessed 1 June 2025].

Lawfare, 2025. *AI-enhanced social engineering will reshape the cyber threat landscape*. [online] Available at: https://www.lawfaremedia.org/article/ai-enhanced-social-engineering-will-reshape-the-cyber-threat-landscape [Accessed 17 May 2025].

Lemay, A. and Leblanc, S., 2018. Cognitive biases in cyber decision-making. In: *Proceedings of the 13th International Conference on Cyber Warfare and Security*, pp.395.

Linder, B. (2023) 'Highly effective responses to the alarming democratization of AI', *Check Point Software Technologies Blog*, 16 July. Available at: https://blog.checkpoint.com/executive-insights/highly-effective-responses-to-the-alarming-democratization-of-ai/ (Accessed: 27 June 2025).

Lu, Y., Huang, X., Zhang, Y., Yan, J. and Liu, F., 2024. AI-enhanced social engineering attacks: A comprehensive analysis of emerging threats. *IEEE Transactions on Information Forensics and Security*, 19, pp.2156–2169.

Manyam, S. (2025) *Artificial Intelligence's Impact on Social Engineering Attacks*. B.Tech. thesis (KMM Institute of Technology & Sciences, 2015; P.G.D.M., International School of Management Excellence, 2017), submitted to Governors State University. Available at: https://opus.govst.edu/cgi/viewcontent.cgi?article=1521&context=capstones (Accessed: 27 June 2025).

Mark, M.S., 2021. *An analysis of factors influencing phishing threat avoidance behavior: A quantitative study*. Doctoral dissertation. Capella University.

Mauro, M., Galatro, G., Fortino, G. and Liotta, A., 2020. Experimental review of neural-based approaches for network intrusion management. *IEEE Transactions on Network and Service Management*, 17(4), pp.2480–2495.

Mimecast, 2024. *CEO Fraud: An analysis of the threats*. [online] Available at: https://www.mimecast.com/content/ceo-fraud/ [Accessed 16 April 2025].

Mouton, F., Malan, M.M., Kimppa, K.K. and Venter, H.S., 2015. Necessity for ethics in social engineering research. *Computers & Security*, 55, pp.114–127.

Mouton, F., Malan, M.M., Leenen, L. and Venter, H.S., 2014. Social engineering attack framework. In: *2014 Information Security for South Africa*, IEEE, pp.1–9.

Ngakpal, T. and Prasad, S.S., 2023. Social engineering: Techniques & implications. *Kilby*, 100(7), p.7.

Nobles, C., 2018. Botching human factors in cybersecurity in business organizations. *HOLISTICA Journal of Business and Public Administration*, 9(3), pp.71–88.

Olney, M. (2024) 'Proactive Insider Risk Management: A key defence against Social Engineering attacks', *Integrity360 Insights*, 15 April. Available at: https://insights.integrity360.com/proactive-insider-risk-management-a-key-defence-against-social-engineering-attacks (Accessed: 27 June 2025).

Omar, S.Z., Kovalan, K. and Bolong, J., 2021. Effect of age on information security awareness level among young internet users in Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 11(19), pp.245–255.

Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T. and Koucheryavy, Y., 2018. Multi-factor authentication: A survey. *Cryptography*, 2(1), p.1.

Oxford, 2024. *Deepfake Deception Study 2024*. [online] Available at: https://www.ox.ac.uk/news/2024-03-10-deepfake-deception-study [Accessed 17 May 2025].

Pakina, A. K., Kejriwal, D., & Pujari, T. D. (2025). Adversarial AI in Social Engineering Attacks: Large-Scale Detection and Automated Counter measures. *International Journal Science and Technology*, *4*(1), 1-11.

Paravathi, C., Dhanyashree, G., Yeshaswini, R. and Lisha, S., 2024. Unmasking the evolution of social engineering in cybersecurity: Techniques, vulnerabilities, and countermeasures. *International Journal of Engineering and Management Research*, 14(1), pp.65–70.

Parsons, K., McCormac, A., Butavicius, M., Pattinson, M. and Jerram, C., 2014. Determining employee awareness using the human aspects of information security questionnaire (HAIS-Q). *Computers & Security*, 42, pp.165–176.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.

Proofpoint, 2024. *2024 State of the Phish Report*. [online] Available at: https://www.proofpoint.com/us/resources/threat-reports/state-of-phish [Accessed 16 Apr. 2025].

Proofpoint, 2024. *State of the Phish: Human Risk Report 2024*. [online] Available at: https://www.proofpoint.com/us/resources/threat-reports/state-of-the-phish [Accessed 17 May 2025].

Pujari, S.R. and Hussain, M.A., 2024. Human factor in cybersecurity: Behavioral insights into phishing and social engineering attacks. *Nanotechnology Perceptions*, 20(S15), pp.630–642.

Purplesec, 2024. *Cyber security statistics: The ultimate list of stats, data & trends*. [online] Available at: https://purplesec.us/resources/cyber-security-statistics/ [Accessed 17 May 2025].

Putra, F.P.E., Zulfikri, A., Arifin, G. and Ilhamsyah, R.M., 2024. Analysis of phishing attack trends, impacts and prevention methods: Literature study. *Brilliance: Research of Artificial Intelligence*, 4(1), pp.413–421.

SAGE, 2023. *Quantitative research methods: An introduction*. [online] Available at: https://uk.sagepub.com/en-gb/eur/quantitative-research-methods/book245146 [Accessed 17 May 2025].

Saiwa (2023) 'Everything You Need to Know About Anomaly Detection in Cybersecurity', *Saiwa Blog*, 11 December. Available at: https://saiwa.ai/blog/anomaly-detection-in-cybersecurity/ (Accessed: 27 June 2025).

Saiwa (2023) 'Everything You Need to Know About Anomaly Detection in Cybersecurity', *Saiwa Blog*, 11 December. Available at: https://saiwa.ai/blog/anomaly-detection-in-cybersecurity/ (Accessed: 27 June 2025).

Salahdine, F. and Kaabouch, N., 2019. Social engineering attacks: A survey. *Future Internet*, 11(4), p.89.

Salkind, N.J., 2010. *Encyclopedia of research design*. Thousand Oaks, CA: SAGE Publications.

Sangfor Technologies (2024) *Machine Learning in Cybersecurity: Benefits and Challenges*, published 20 June 2024, last modified 26 May 2025, *Sangfor Technologies Cybersecurity Blog*. Available at: https://www.sangfor.com/blog/cybersecurity/machine-learning-in-cybersecurity-benefits-and-challenges (Accessed: 27 June 2025).

Sarker, I.H., Kayes, A.S.M., Badsha, S., Alqahtani, H., Watters, P. and Ng, A., 2020. Cybersecurity data science: An overview from machine learning perspective. *Journal of Big Data*, 7(1), pp.1–29.

Saylor Academy, 1996. *An overview of social engineering: Mitigation techniques*. [online] Available at: https://learn.saylor.org/mod/book/view.php?id=29612&chapterid=5170 [Accessed 17 May 2025].

Scaler, 2025. *Why is Python important for cybersecurity?* [online] Available at: https://www.scaler.com/topics/cyber-security/cybersecurity-importance-for-python/ [Accessed 1 June 2025].

Schroeder, C., 2019. *Susceptibility to social engineering: Human vulnerabilities*. Utica College.

Sci-Tech Today (2025) *Social Engineering Statistics by Types, Country and Facts (2025)*. Published 19 May, updated. Available at: https://www.sci-tech-today.com/stats/social-engineering-statistics-updated/ (Accessed: 27 June 2025).

Secureframe, 2024. *Social engineering statistics*. [online] Available at: https://secureframe.com/blog/social-engineering-statistics [Accessed 16 Apr. 2025].

Secureframe, 2025. *2025 Healthcare cybersecurity statistics*. [online] Available at: https://secureframe.com/blog/healthcare-cybersecurity-statistics [Accessed 17 May 2025].

SecurityWeek, 2025. *AI-driven social engineering: The next generation of cyber threats*. [online] Available at: https://www.securityweek.com/ai-social-engineering-cyber-threats-2025/ [Accessed 17 May 2025].

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L.F. and Downs, J., 2010. Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp.373–382.

Shetty, P. (2024) ‘Deepfake Laws: How are Regulators Approaching Them and What Should You Do?’, *Arya.ai Blog*, 30 September. Available at: https://arya.ai/blog/deepfake-laws (Accessed: 27 June 2025).

Shillair, R., Esteve-González, P., Dutton, W. H., Creese, S., Nagyfejeo, E., & von Solms, B. (2022). Cybersecurity education, awareness raising, and training initiatives: National level evidence-based results, challenges, and promise. *Computers & Security*, *119*, 102756.

Skillfloor, 2025. *R vs Python for data science: A friendly comparison*. [online] Available at: https://skillfloor.com/blog/data-science-r-vs-python [Accessed 1 June 2025].

Skorodumov, B. I., Skorodumova, O. B., & Matronina, L. F. (2015). Research of human factors in information security. *Modern Applied Science*, *9*(5), 287.

SlashNext, 2024. *The Phishing Threat Landscape 2024*. [online] Available at: https://www.slashnext.com/resources/reports/phishing-threat-landscape-2024/ [Accessed 17 May 2025].

Sonowal, G. (2021). *Phishing and Communication Channels: A guide to identifying and mitigating phishing attacks*. Apress.

Sprinto, 2024. *Social Engineering Statistics*. [online] Available at: https://sprinto.com/blog/social-engineering-statistics/ [Accessed 16 April 2025].

Sprinto, 2025. *AI and Cybersecurity: The 2025 Threat Landscape*. [online] Available at: https://sprinto.com/blog/ai-cybersecurity-trends-2025 [Accessed 17 May 2025].

Statista, 2024. *Global most targeted industries phishing 2024*. [online] Available at: https://www.statista.com/statistics/266161/websites-most-affected-by-phishing/ [Accessed 16 April 2025].

Statista, 2024. *Ransomware attacks by country*. [online] Available at: https://www.statista.com/statistics/1246438/ransomware-attacks-by-country/ [Accessed 16 April 2025].

Stevens, A., 2025. *Implementing social engineering attack-resistant policies*. Crowe. [online] Available at: https://www.crowe.com/global/insights/implementing-social-engineering-attack-resistant-policies [Accessed 17 May 2025].

Stojanovic, B., Hofer-Schmitz, K. and Kleb, U., 2020. APT datasets and attack modeling for automated detection methods: A review. *Computers & Security*, 92, p.101734.

Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). Mitre att&ck: Design and philosophy. In *Technical report*. The MITRE Corporation.

Team, T.M.D., 2023. *Matplotlib: Visualization with Python*. Zenodo.

Trend Micro (2025) 'The Future of Social Engineering', *Trend Micro Security News*. Available at: https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/the-future-of-social-engineering (Accessed: 27 June 2025).

Trent, R. (2025) 'The Human Factor: Understanding Social Engineering Attacks and How to Prevent Them', *Rod Trent*, 25 February. Available at: https://rodtrent.substack.com/p/the-human-factor-understanding-social (Accessed: 27 June 2025).

Unimrkt Research, 2023. *What is quantitative research? Definition, types, characteristics, and examples*. [online] Available at: https://www.unimrkt.com/blogs/what-is-quantitative-research-definition-types-characteristics-and-examples [Accessed 17 May 2025].

van der Kleij, R. and Leukfeldt, R., 2020. Cyber resilient behavior: Integrating human behavioral models and resilience engineering capabilities into cybersecurity. In: *Advances in Human Factors in Cybersecurity: Proceedings of the AHFE 2019 International Conference on Human Factors in Cybersecurity*, July 24–28, 2019, Washington DC, USA. Springer International Publishing, pp.16–27.

Verizon, 2024. *2024 Data Breach Investigations Report*. [online] Available at: https://www.verizon.com/business/resources/reports/dbir/ [Accessed 17 May 2025].

Wickham, R.J., 2019. Secondary analysis research. *Journal of the Advanced Practitioner in Oncology*, 10(4), pp.395–400. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC7520737/ [Accessed 1 June 2025].

World Economic Forum, 2024. *Global Cybersecurity Outlook 2024*. [online] Available at: https://www3.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2024.pdf [Accessed 17 May 2025].

ZeroFox, 2024. *MFA Bypass Techniques Report 2024*. [online] Available at: https://www.zerofox.com/resources/reports/mfa-bypass-techniques-2024/ [Accessed 17 May 2025].

Zscaler, 2024. *Cloud Threat Report 2024*. [online] Available at: https://www.zscaler.com/resources/reports/cloud-threat-report-2024.pdf [Accessed 17 May 2025].

Zscaler, 2024. *Phishing Report 2024: The state of phishing attacks*. [online] Available at: https://www.zscaler.com/resources/reports/phishing-report-2024.pdf [Accessed 17 May 2025)