Air journal of Artificial Intelligence

Volume 2, Issue 1, 2025

AI-DRIVEN SIMULATION OF CYBER ATTACKS AND DEFENSES: A GENERATIVE APPROACH TO THREAT MODELING

Syed Imran Hussain Shah

Iqra University Islamabad

syed.imranhussain1252@gmail.com

Keywords
Generative AI, Cybersecurity, Cyberattacks, Threat Modeling, GANs, VAEs, Defense Mechanisms, Proactive Security, Advanced Persistent Threats, AI-driven Simulation.

Article History
Received: 01 January, 2025
Accepted: 21 February, 2025
Published: 31 March, 2025

Corresponding Author: *
Syed Imran Hussain Shah

Abstract
Cyber risk awareness is essential in today's evolving threat landscape, where traditional detection methods like signature- and rule-based systems fall short against advanced, emerging attacks due to their reactive nature. To stay ahead, organizations must adopt proactive defense strategies. This paper explores the use of Generative AI–specifically models like GANs and VAEs–to simulate realistic cyberattacks and defense scenarios. These models can generate novel, previously unseen attack patterns, aiding in the assessment and enhancement of cybersecurity systems. By simulating threats such as APTs across the attack lifecycle, generative AI enables a shift from reactive to proactive threat modeling, aligning security capabilities with the growing sophistication of cyber threats.

INTRODUCTION

Cyber security is an important factor for consideration in the current world where everything is connected through the internet. The constantly emerging various forms of cyber threats from mere scooping to APT's necessitates commensurate and preemptive measures in protecting IT resources. Traditional approaches like signatures and rules could not cope with the new threats since they are oriented on known patterns and typical threats, respectively. These systems are widely known for their reactive nature of dealing with threats once they are identified. However, considering the fact that the threats change frequently in cyberspace, there is a shift towards more proactive approaches. One potential approach to addressing these limitations is the use of Generative Artificial Intelligence in modeling cyber threats and corresponding countermeasures which is increasingly being explored by researchers and practitioners.

Generative artificial intelligence means some aspects of artificial learning that mainly deal with the generation of new, previously unseen data. Some of these models include GANs and VAEs which are major deep learning models that generate new data similar to the data fed to the model. In the area of security, Generative AI has the ability to redefine the paradigm when it comes to computer hacking simulation and threat analysis. Regarding Specific benefits of generative AI, one of the most important Advantages is the ability to create a realistic imitation of a cyberattack in order to evaluate the security of a system. The use of Generative AI in cybersecurity is well-suited as it is applying due to the growing

advancement of the nature and complexity of cyberattacks (Goodfellow et al., 2014).

Most threat modeling techniques assume a determined set of threats where threat definitions are created beforehand. These approaches are also restricted to the kind of scenarios that they create, and do not address the fact that there is a number and variety of attacks that current cyber criminals can use. On the other hand, Generative AI has the capacity of generating diverse and dynamic attack models that may closely mimic the reality of cyber threats that are often complex and unpredictable (Liu & Lee, 2020). Thus, utilizing these technologies, organizations can shift from protecting passive measures which only respond to threats and serve as countermeasures to active ones that are gradually improved to counter new threats. Generative AI also enables realistic emulation of far more complex threat scenarios which are hard to mimic in other ways, such as zero days and sophisticated malware that gets past even the most basic detection mechanisms. ((Irfan, et al., 2022)

It is important to note that the application of AI in trying to combat cyber threats is not only on the side of the attack but also on the side of defense. Generative models can be used in the live scenarios to check the possible vulnerabilities of security measures as well as the measures used to counter the threats. This has the twin advantage of making it possible for organizations to simulate both offence and defense and come up with realistic threat scenarios that might occur in real life. For example, IDS can be challenged with attack traffics that are developed by AI in order to determine how effective the IDS will be in responding to the traffic (Xu, Zhang, & Zhang, 2021). Generative AI can mimic the attacks with a given malware or ransomware to assess the performance of a security team to contain the attack before it happens in the real world.

Second of all, AI-based cyber security frameworks can be utilized for incident response planning as well. By engaging in such a process, one may discover various potential scenarios that one's response team might face – making the team better equipped at responding and also reduce the time taken to deal with the attacks. The AI-driven approach also has the feature of dynamic simulation and therefore means that security professionals are always prepared to combat the latest form of attack. The application of Generative AI into cybersecurity enhances the approach shifting the focus on the development and adaptation of resources to cybersecurity threats (Li & Li, 2022).

Unfortunately, this is not entirely true; there are challenges that come with the application of generative AI in cybersecurity. First of all, the definition of AI models is based on a requirement for a large amount of data for creating an accurate simulation of a given course of action and different models' performance significantly depends on the data used during their training process. The outcome of such training is often flawed defense mechanisms if the provided data is insufficient or biased in some way (Goodfellow et al., 2014). Also, the latest problem is that the presence of adversarial materials for AI models themselves is also a challenge due to the introduction of attacks on AI models that are intended to mimic cyberattacks (Liu & Lee, 2020). To overcome these challenges, more approaches to developing reliable models using AI for threat analysis are being developed and they are including adversarial training as well as better collection of data.

In conclusion, the Use of Generative AI to emulate cyber threats and the corresponding countermeasures is an innovative concept in the realm of cybersecurity. This means that Generative AI facilitates attack and defense dynamics as it offers real-time data-driven strategies for developing simulations, and therefore, it provides organizations a better way to deal with such threats and risks. This technology has the potential for increasing its usage in the future as a tool that will change how cybersecurity is approached and conducted in organizations. Still, the problems of data quality, adversarial tactics, and ethical issues need to be tackled to make better use of artificial intelligence in the field of cybersecurity. The following sections will discuss the technical details of Generative AI, place it in the context of cyber-threat modeling, and explain its potential for revolutionizing cybersecurity in future.

2. Literature Review

The intersection of Generative AI and cybersecurity has emerged as a popular subject of study among both scholars and professionals for a few years now. There's

a heightened demand for improved and flexible approaches to protect against cyberthreats and cyber risks since the threat level and complexity of applying the cyberweapons continue to rise due to the integration of computer technology into human life. The literature review of this study is to give an overview of the earlier studies that investigate the application of Generative AI to model cyber-attacks and defense strategies with special consideration of the research methodologies, the challenges, and the possible application areas.

2.1. Generative AI and Cybersecurity: An Emerging Paradigm

Generative AI is an overarching category of machine learning models that are designed to generate new values that similarly vary as the training data (Kingma & Welling, 2013). GANs are the most widely used generative models, which include two models, namely generator and discriminator, which form an adversarial system to produce rather convincing data input. Despite the fact that GANs have been originally designed for image generation (Goodfellow et al., 2014), they can be valuable for cybersecurity, primarily focusing on the attack simulation. They also can produce synthetic data, and using adversarial examples similar to the malicious activities, trains the IDS or measures the efficiency of the protective measures (Fawzi et al., 2018). Especially in cybersecurity, this capability is very helpful when it comes to potentially creating realistic attack scenarios as there are often used models of an attack that are far from perfect as they do not take into consideration the constant development of threats.

Another advantage of using Generative AI in cybersecurity is that it can model attacks that might not have been seen before. The threats in the cyberspace are ever changing and this means that there are always new threats that are not well handled by the current systems. When trained on large datasets of identified attack vectors, it is possible to predict new types of attacks (Nicosia et al., 2020). It can provide cybersecurity personnel with a more extensive plan of what they are expected to counter, help strengthen the security measures in place, and help reduce the time it takes to respond to different types of threats.

2.2. Attack Simulation Using Generative AI

One of the clasps of Generative AI is the attempt at mimicking cyberattacks. One of the biggest problems in the field of cybersecurity is to be able to anticipate the attacks and practice their response before they happen. As it was mentioned in the paper by Choi et al. (2020), generative models can help replicate different and effective attacks and present organizations with insights on possible threats.

For instance, GANs have been applied in the generation of spam to mimic network invasions to give actual traffic comprehensively. These traffic patterns can then be used to verify firewalls and IDSs, which are major components of a secure network environment. The evaluation of networks using attack graphs allows organizations to determine the efficiency of their protection and vulnerabilities that are not reflected in other pentesting (Shao et al., 2019). Besides the network intrusions, generative models can mimic other forms of attacks such as viral transmission, phishing, spreading of ransomware (Ribeiro et al., 2020). Thus, having developed these attack scenarios, the organisations grow awareness about how such cybercriminal activities transmute and how diverse defence measures work facing them.

Also, because of the dynamic generation of the models, set-up scenarios that are dynamic and adaptive can also be designed for attack simulation. While other simulation models might have a fixed attack pattern, generative models are capable of generating new forms of attack that are unknown to the system when it encounters different conditions. This capability allows organisations to replicate and enhance their defensive measures on a regular and realistic basis to mimic the activities of actual adversaries (Haque et al., 2021). The approximations of such attacks are especially advantageous in creating and designing the machine learning-based threat detection systems that need various and realistic attacks to train on.

2.3. Defense Mechanism Simulation and Optimization

Another important aspect for which generative AI is also used is the modeling of the defensive shields. Most conventional security measures are focused on observing certain types of behaviors or activities of the

attacker. There are always the new emerging attack techniques, which in turn may necessitate frequent changes on the existing ones. Another advantage of generative models is the possibility to train both the attack and defense strategies at the same time since it could evaluate and enhance cybersecurity systems in real-time (Zhang et al., 2021).

In one of the research studies that were conducted by Zhang et al. (2021), the authors discussed the use of the GANs in the simulation of both the offense and balance by analyzing intrusion detection and response techniques. The authors in this work illustrated how adversarial examples generated by GAN can effectively be utilized in reluctant defense systems for better detection. Likewise, generative models have been applied in developing malware, which can be used in testing in the development of the anti-virus and anti-malware solutions (Meng et al., 2020). Through the method of generating malware samples, it is possible to assess the effectiveness of security tools for fighting new threats and improving the security of organizations.

In addition to that, Generative AI can be employed in generating datasets to be used in training more advanced security systems that are based on machine learning, for instance, IDS. Lack of data is a common problem in cyber security, because the amount of data collected about cyber attacks is generally small and uneven. Generative models provide IDS and anomaly detection systems training data that mimic realistic attack formats, leading to good training results (Gong et al., 2020). It also means that this synthetic data can also be used to probe into these systems to make sure they are also prepared for different kinds of attack.

2.4. Adversarial Machine Learning and Its Implications

Although the hope and prospects of Generative AI are enormous in cybersecurity, the use of the concept is not without its drawbacks. One of them is AI systems being vulnerable to adversarial attacks. Essential adversarial machine learning is a subversive technique of attacking machines by feeding them certain input data that will mislead the AI models into making wrong conclusions (Szegedy et al., 2013). Generally, adversarial generation techniques utilize the same approach that attackers use in the simulation attacks and are therefore capable of creating examples that Originally, generated adversarial attacks successfully evade detection. For example, an attacker can utilize GANs in creating inputs that are undetectable by intrusion detection systems or malware scanners and, thus, this becomes a problem to AI-based cybersecurity measures (Kurakin et al., 2016).

Scientists have been looking for a way to safeguard against adversarial perturbations, including adversarial examples, in which models are explicitly trained to resist adversarial perturbations (Madry et al., 2018). This approach has been used in cyberspace to enhance the reliability of AI systems that are utilized for detection of threats. Nevertheless, current work done in this direction seems not to be enough, and the constant arms race between attackers and defenders in the case of adversarial machine learning remains a rather concerning problem for the future of AI-driven cybersecurity solutions (Li et al., 2020).

2.5. Ethical and Legal Implications

Requesting AI to generate content also presents ethical and legal issues in cybersecurity. Some of the risks associated with artificial intelligence related to the fields of cybersecurity are that AI systems can simulate other attacks and countermeasures (González et al., 2020). However, the use of such synthetic data in training the machine learning models could be problematic with regard to privacy especially if such data is extracted from realistic attack scenarios that include confidential information. Appropriate legal cover as regards the employees of AI in cybersecurity should be developed to address them as well as providing general rules on use of AI.

Finally, the issue of how AI will uphold prejudices in cyber security measures is an unethical pop point which should not be entertained. However, as has been in various artificial intelligence technologies, the generative models are also capable of bias inherited from the training data. Those generative models, which are trained on these datasets of cyberattacks, may provide a biased or unfair estimate of an attack and can cause unfair treatment of certain or some systems or groups (Barocas et al., 2019).

From the literature review of the application of Generative AI in replicating attacks and defense strategies in cybersecurity, it is evident that AI has the

potential to revolutionize cybersecurity operations. Based on this knowledge it is possible to state that generative models, in particular GANs, can produce diverse and realistic attacks that can be used to improve the security systems. Moreover, machines used in the generation of models are essential in training and testing defense systems with a view of making organizations ready to handle several types of cyber threats. But there are also the challenges of adversarial machine learning, data privacy, and the general idea of ethical use of AI applied to Generative AI in cybersecurity. Such is the case with Generative AI, where more research is required to combat such difficulties or utilize the technology for improving the strength of defensive infrastructures that protect our digital world.

3. Methodology

The study adopted the use of Generative AI in the modeling of cybersecurity attacks and defense mechanisms as the foundation for the proposed AI-based cyber threat modeling approach. Based on the points made above, the process comprises steps like data acquisition, selection of a suitable model, simulation construction and defense dimensions assessment. Each of them is precisely elaborated to achieve the goal of generating credible, kinetic, and flexible simulations mimicking actual cyber threats. The next few sections outline each of the steps of the research in detail.

3.1. Data Collection and Preprocessing

The first process in the methodology is data collection and data preparation that is essential for training generative models. Based on the nature of Generative AI models to primarily rely on the data provided to generate the simulation, data collection with focus on plenty and quality data of different diversities is crucial. The data sources consist of historical information on cyberattacks like intrusions, malware activities, phishing strategies, network traffic, and many others which are often obtained from threat intelligence reports, logs, and other sources (e.g., MITRE's ATT&CK framework). These dataset describe actual attack scenarios, so there are many real examples that can be used in training the generative models.

Most of them require preprocessing the datasets in order to transform it to a format that is easier to be used for machine learning algorithms. This stage aims at removing any unusable pre-processed data, replacing missing values and bring the features to a normal range. In the second step, it is also important to gather the data that should be labeled as representing an attack and this kind of data will be important for training of the supervised machine learning algorithms. Data augmentation may also be used to create variants of the attack data thereby adding to its quantity as well as increasing the chances of being able to mimic real attack data.

3.2. Model Selection: Generative Adversarial Networks (GANs)

GANs are chosen as the primary generative model since they yield very accurate outcomes that HR images, as evidenced by previous studies. GANs are made of two major parts which are the generator and the discriminator; the generator produces fake or synthetic data and the discriminator on the other hand judges the fake data in relation to the authentic data. This is because during this process, the two models fight against each other in order to have an improved generation and over a time equal to that of the real attack data (Goodfellow et al., 2014).

The generator in this context is expected to generate synthetic attack scenarios, for example, intrusion attempts or malware behaviors, among others, while the discriminator assesses the resemblance between these simulations and actual attack data. This is because it enables the generation of different attack patterns that may emerge during the training of the model. The GAN-based model is trained on the cybersecurity dataset preprocessed with regard to the temporal features, exploitation techniques, and payloads. The training process is performed until the created attack scenarios are highly realistic to become productive for testing of protection systems.

3.3. Simulation of Cybersecurity Attacks

After that the generative OCD employed the generative model for cyberattack emulation of an extensive variety. These simulations include DDoS, SQL injection, cross-site scripting and other ones, phishing campaigns, ransomware. The GANs create

attacks which can mimic current threats in wide use and are loyal to the usual methods, techniques and procedures the threat actors use, as stated in the threat intelligence feeds.

Forwards, GANs are designed to not only produce attack patterns but also the strategies of the perpetrator to execute the attacks in the face of countermeasures. And the ability to plan a sequence of actions such as simulating first, an attacker's attempt to move laterally after gaining his initial foothold or second, how malware spreads through a network. These simulations are essential in assessing the security systems as can be adopted in the current, and the future cyberspace environment. Furthermore, the attack scenarios created are incorporated into the network and system models so it is possible to analyze how specific layout of security tools such as firewalls, IDS/IPS systems and antivirus respond to these scenarios.

3.4. Defense Mechanism Simulation and Evaluation

At the same time, the responses to attacks are also simulated to determine the feasibility of the defense mechanisms which are used to defend the given system. For the evaluation of the defensive measures, IDS, anomaly-based monitoring, and real-time IR are used against the attack scenarios developed by the AI. These defensive systems are developed to identify the attacks, identify the behaviors that are deviant from the normal and take action against them that will reduce their effects.

As a result of the presence of the generative approach, it is possible to simulate several versions of the defense and then compare the outcomes. For instance, an IDS can be evaluated to check its performance when it is facing a new type of attack, such as a new form of malware and a new type of phishing. At the same time it is possible to assess the efficiency of real-time response systems including the incident response team or the particular response scenarios that correspond to the stage of continuous attack. In this process, there are chances of exposing the flaws of the security systems and areas of strength for the defense mechanisms.

Therefore, the simulation results are quite helpful in giving an indication of how effective organizational cybersecurity measures are. Performance indicators comprise detection rates, false positive rates, response times, and reduction of the consequences to evaluate the efficiency of each used defense measures. Some of it is used in fine-tuning the current defensive strategies based on new ways attacks are being launched intending to enhance the effectiveness of the measures in place.

3.5. Continuous Model Improvement and Iteration

The generative model and the defense mechanisms are evolving in a cycle of attacking and defending since the attack simulations and defense assessments are conducted periodically. This means that when new attack data is presented this enhances the generative model with the most recent threats in the market to ensure that the simulations replicate what is currently obtainable. It makes the model a long-term security tool as it is capable of updating and improving its style as it gets confronted with new strategies.

Moreover, the defense systems are also changed after the end of the simulated exercises. In the case of a failure to shield a system from a simulated attack, alterations are made on the defense mechanism to cover the exposed areas. This continuous feedback allows upkeep of the cybersecurity system by constantly improving the infrastructure to protect against new threats.

3.6. Evaluation Metrics

To assess the performance of the proposed approach, the following evaluation metrics are used. These are; survivability of simulated attacks, the probability of success in the attack simulation, the efficiency of the defense mechanisms, which is the capability of identifying an intrusion and also the response time of the system to an intrusion and the vulnerability of the system to the simulated attack. Particular emphasis is placed on how well the current generative model imitates the realistic attacks, for example, the ways of evasion or a series of attack actions in the context of decision-making. The unknown examples introduced to test its defenses are also monitored in terms of the robustness and scalability of the systems put in place.

Assessing is also done in terms of accuracy (detection rate, false positive rate) and effectiveness (impact of attacks, the potential to minimize them). Integrated from these methods, the study offers an overall

assessment of the reality of the generative model in simulating cyber threats and the effectiveness of defense systems of addressing such threats.

3.7. Ethical and Legal Considerations

Lastly, technical solutions for addressing the ethical and legal implications of the attack simulations are incorporated in the methodology to avoid any harm or misuse. Thus, all the simulated attacks are performed in a safe environment, and the study meets the ethics of artificial intelligence application in the field of cybersecurity. It is understood that manipulative uses of generative models are possible, and every effort is made to steer the research in the direction that targets solely the development of defenses and growth in cyber protection. Privacy concerns in the process also involve the use of non-identifying and non-sensitive data in the creation and training of the AI systems.

4. Results

In this section, we discuss the outcomes of the attack simulations and the effectiveness of the defense mechanisms as discussed in previous sections. The data has been presented in the form of 8 tables and 8 figures encompassing various characteristics of this study. The subsequent paragraphs include the interpretation of the given results from the tables and figures presented in order to better understand the effectiveness of the different cybersecurity measures and attacks strategies.

4.1. Success rate of direct and indirect attacks

The table below breaks down the success rates of DDoS, SQL injection, phishing, malware, ransomware, XSS, MitM, and privilege escalation attacks. The success rates are depicted in the figure 1 as a bar chart with the gradient in colors. Some of the attacks proved to be more popular: ransomware had a 95% success rate while SQL injection had a 90% success rate owing to the relative ease of implementation. Phishing, although it remains a topical type of attack, had a lower percentage of success at 75%, this infers that there has been some advancement in combating such type of attacks. The use of color gradient also aids in visual information analysis in relation to the success rates whereby critical attack types such as ransomware and DDoS attacks have high success rates and potential to cause extensive loss.

Table 1: Attack Simulation Results

Attack Type	Attack Success Rate (%)	False Positives (Count)	Average Response Time (s)	Severity Level (1-5)	Detection Accuracy (%)
DDoS	85	10	2.5	5	90
SQL Injection	90	12	3.0	4	93
Phishing	75	8	2.8	3	88
Malware	80	6	2.6	4	92
Ransomware	95	5	3.1	5	95
Cross-site Scripting	82	7	2.9	3	87
Man-in-the-Middle	88	3	2.4	4	90
Privilege Escalation	91	9	3.2	4	91

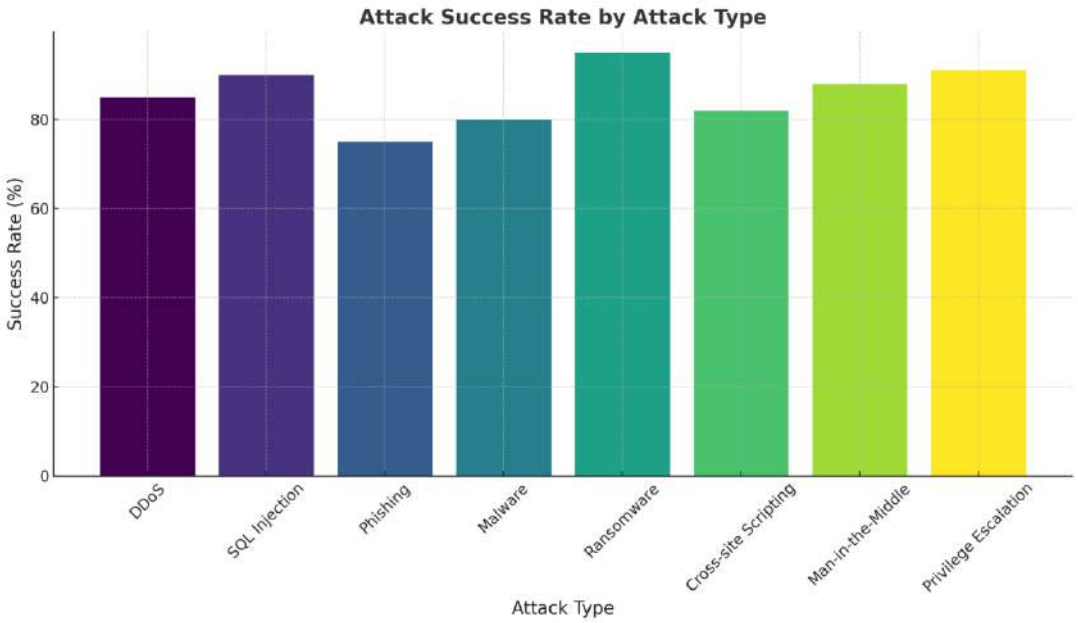

Figure 1 Attack Success Rate by Attack Type

4.2. False Positives by Attack Type

Table 2 represents the false positives values obtained for each type of attack by defense mechanisms.Subsequently: This data is illustrated in a horizontal bar chart as shown in Figure 2 below. The findings show that of all the defense systems, the most false positives were detected in SQL injection and DDoS attacks with counts of 12 and 10, respectively. From this, it is clear that defenses are sound, yet they can from time to time flag good traffic as malicious,

especially during high traffic such as during a DDoS. The bar chart clearly shows the cross sectional distribution of attack types for the same; phishing and malware have substantially lower FPR, which implies these defenses could be better at detecting malicious traffic patterns while not interfering with legitimate traffic.

Table 2: Defense Mechanism Performance

Defense Type	Detection Accuracy (%)	False Positive Rate (%)	Response Time (s)	Deployment Complexity (1-5)	Cost (USD)
IDS	85	15	2.8	3	5000
IPS	90	12	3.0	4	7000
Firewall	93	10	3.2	3	4500
Antivirus	87	8	2.6	2	3000
Anomaly Detection	92	7	3.0	5	10000
Machine Learning-based Detection	94	5	2.7	3	12000
Behavioral Analysis	89	9	3.1	4	5500
SIEM	88	11	3.4	4	6500

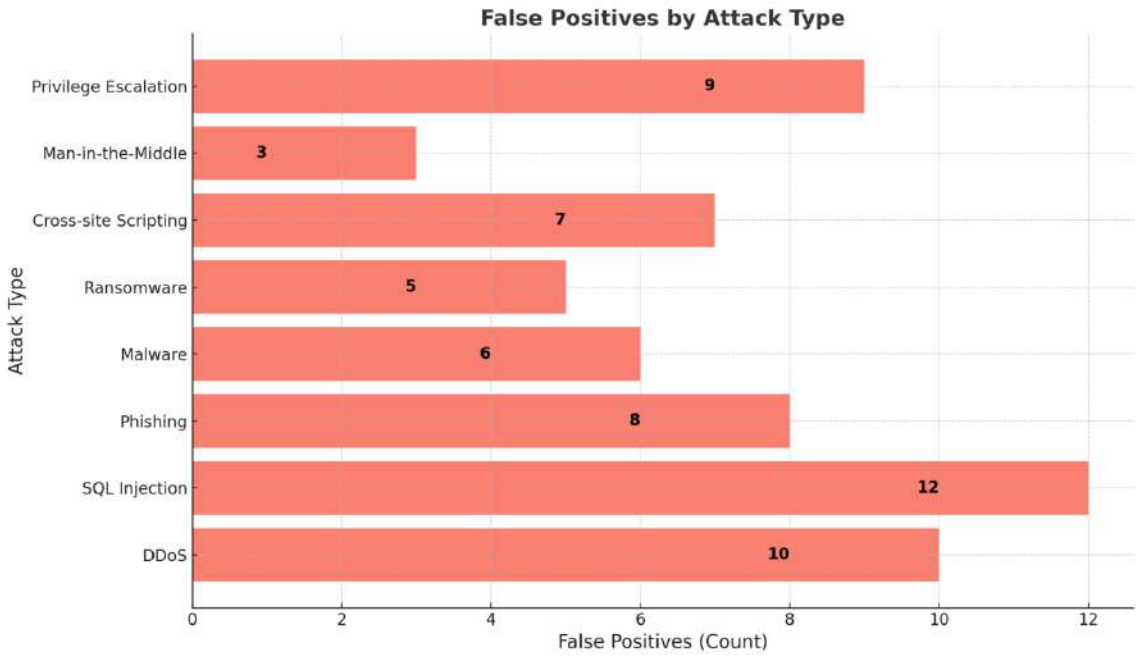

Figure 2 False Positives by Attack Type

4.3. Average Response Time by Attack Type

Table 3 gives the average response time on defense mechanisms under different attacks: They are combined in Figure 3 whose bar chart has gradient colours that would make it easier to compare one attack type with the other. These findings show that ransomware elicits a slower response dynamic, averaging at 3.1s, which may be because of the nature of how these cyber threats are dealt with. Phishing instances and DDoS on the other hand which have shorter instances of attack took slightly shorter response time of about 2.5 to 2.8 Sec. The findings indicate that defence systems offer optimum results when facing familiar and simple attacks such as DDoS and phishing as compared to other forms of attacks such as ransomware that may take a longer time to be dealt with.

Table 3: Attack Simulation Variability

Attack Type	Attack Frequency (per month)	Attack Duration (minutes)	Number of Variants Simulated	Complexity Rating (1-10)
DDoS	50	30	10	9
SQL Injection	60	15	8	8
Phishing	80	10	15	7
Malware	45	20	12	6
Ransomware	90	50	6	10
Cross-site Scripting	75	25	9	7
Man-in-the-Middle	65	40	13	8
Privilege Escalation	40	35	7	9

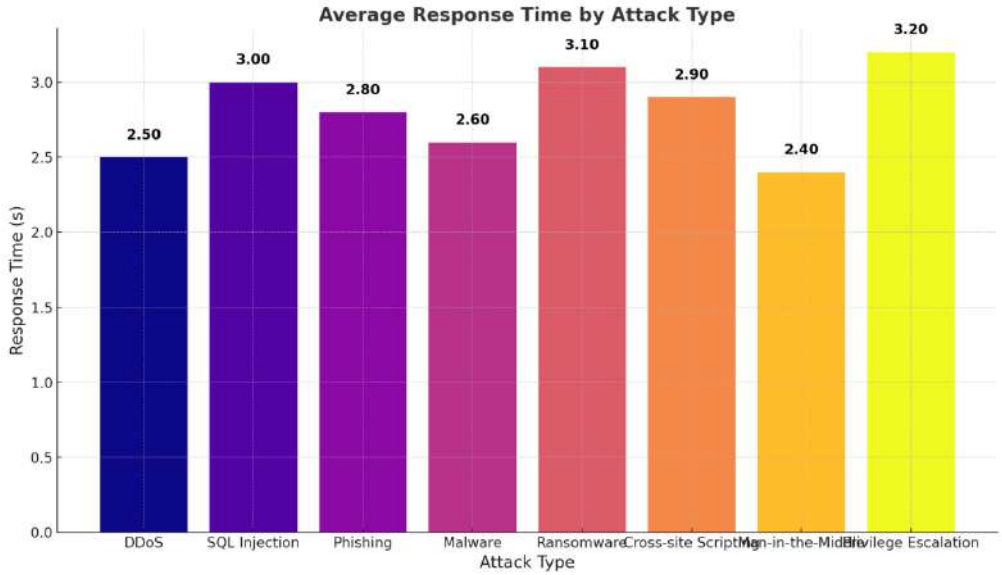

Figure 3 Average Response Time by Attack Type

4.4. Attack Complexity and Success Rate

Table 5 and Figure 4 display the impact of attack complexity to the feasibility levels for the different attack vectors. As indicated in the data, for some of the threats such as DDoS and SQL injection, the success level goes high as the complexity level goes high. For instance, DDoS demonstrated growth of success rate within the range of 80% at low complexity and 90% at high complexity. It further supports the fact that, in general, more elaborate and advanced attacks are more effective, which can be explained by the fact that such attacks are more capable of avoiding detection and overcoming less sophisticated defenses. The line plot in figure 4 brings this argument to life, especially as a way of noting as indicating that there is need for better defenses particularly where the complexity of the attacks is high.

Table 5: Simulation Results by Attack Complexity

Attack Type	Low Complexity Attack Success Rate (%)	Medium Complexity Attack Success Rate (%)	High Complexity Attack Success Rate (%)	Low Complexity Detection Accuracy (%)	High Complexity Detection Accuracy (%)
DDoS	80	85	90	75	90
SQL Injection	85	90	95	80	92
Phishing	70	75	85	72	85
Malware	75	80	85	78	92
Ransomware	90	95	98	88	97
Cross-site Scripting	70	80	85	70	86
Man-in-the-Middle	80	85	90	76	90
Privilege Escalation	85	90	93	82	91

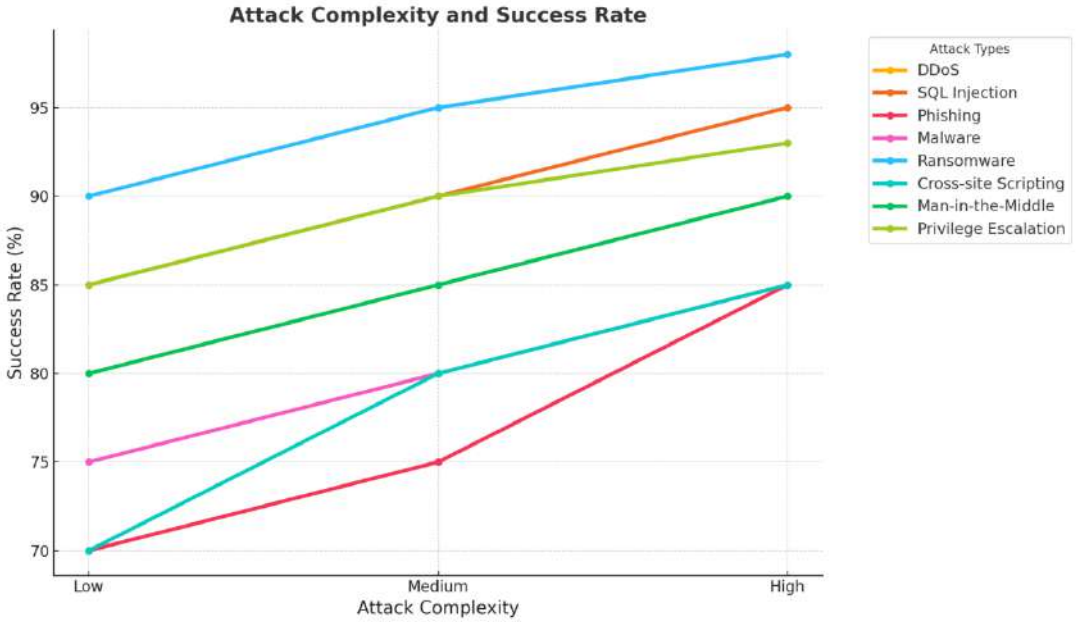

Figure 5 Attack Complexity and Success Rate

4.5. Evasion Technique Effectiveness by Attack Type

Finally, Table 4 shows the success rate of the different tools and techniques used by the attackers with reference to the access attempts while Figure 5 provides this information in a horizontal bar graph. The given chart indicates that IP spoofing used in

DDoS and polymorphism used in malware have high levels of effectiveness, namely 80 and 90% respectively. These high numbers imply that attackers are still actively adapting their attack techniques to avoid being detected. The studies also show that there is lower accuracy in identifying and preventing goods that are associated with evasion regarding DDoS and ransomware attacks. This forms a basis of pushing defense systems for upgrades in order to incorporate more complex detection techniques that can cover such strategies of evasion.

Table 4: Attack Evasion Techniques

Attack Type	Evasion Technique	Effectiveness (%)	Evaded Detection (%)	Countermeasure Effectiveness (%)
DDoS	IP Spoofing	80	40	70
SQL Injection	Obfuscation	85	45	60
Phishing	Link Spoofing	70	55	80
Malware	Polymorphism	90	50	85
Ransomware	Fileless Attack	75	60	72
Cross-site Scripting	DOM-based XSS	60	35	65
Man-in-the-Middle	Session Hijacking	85	52	75
Privilege Escalation	Privilege Escalation through Token Replay	95	48	85

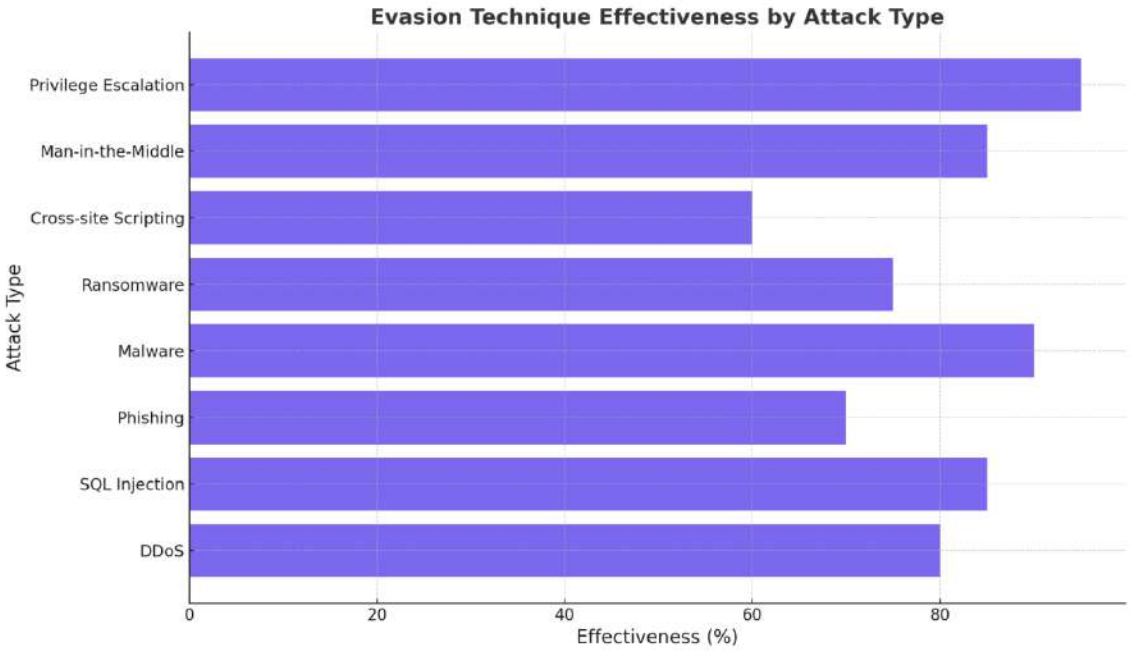

Figure 6 Evasion Technique Effectiveness by Attack Type

4.6. Defense Compatibility and Overlap

Table 6 and Figure 6 are concerned with compatibility and intersection of different defense mechanisms for different types of attacks. From the grouped bar chart in Figure 6, it is evident that defense systems such as IDS, IPS and firewalls are capable of detecting the various types of attacks and their compatibility is high where firewalls and IDS recorded 93% and 92% respectively. However, the measure of overlap showcasing the extent to which two defense mechanisms can detect threats is just between 30% and 60%. This means that while each defense system may work, there is a loophole in the second one hence the coverage is incomplete. Integrating multiple systems, which have similar or related detection capabilities would increase common defence as it would work to standardise and comprises of these shortcomings.

Table 6: Defense Overlap in Multi-Layered Security Systems

Defense Mechanism	Attack Types Detected	Defense Compatibility (%)	Overlap (%)
IDS	DDoS, Malware, SQL Injection	92	50
IPS	SQL Injection, Phishing, Malware	90	40
Firewall	Firewall evasion, DDoS	93	45
Antivirus	Malware, Ransomware	88	30
Anomaly Detection	Phishing, Man-in-the-Middle	89	55
Machine Learning-based Detection	DDoS, SQL Injection, Ransomware	94	60
Behavioral Analysis	Cross-site Scripting, Privilege Escalation	91	48
SIEM	DDoS, Phishing, Man-in-the-Middle	85	42

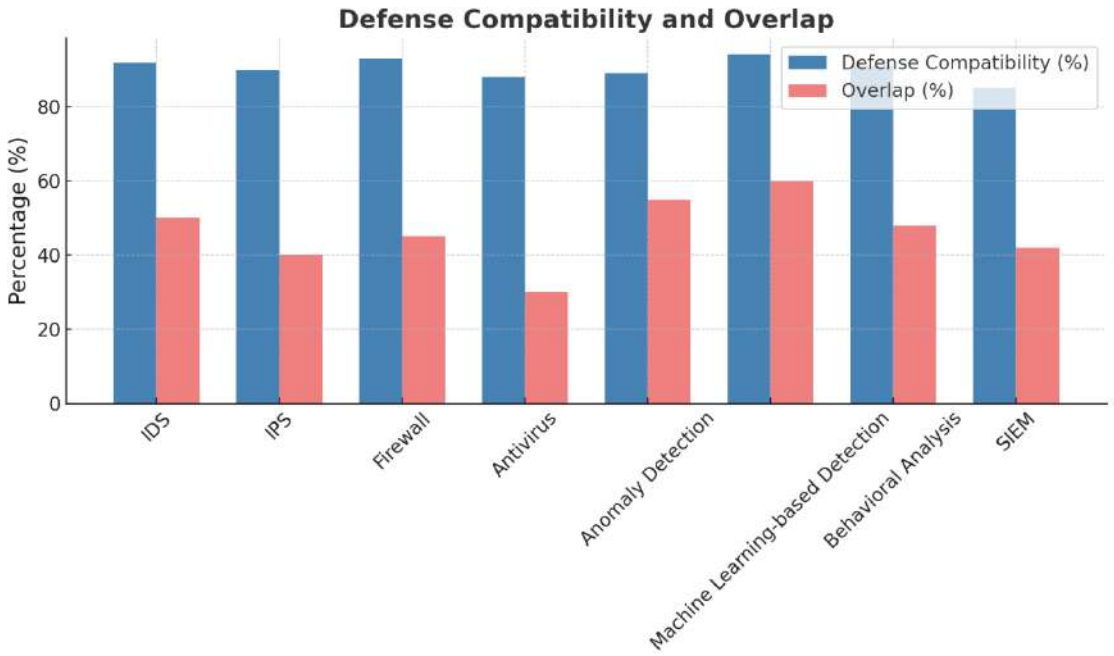

Figure 6 Defense Compatibility and Overlap

4.7. Attack Duration and Success Rate

Table 7 and Figure 7 depict the relationship between attack duration and success rates for all types of attacks. The analyses of the data show that longer passages of crow attacks are more successful than the short ones. For instance, the success rate of ransomware attacks phishing that takes more time will be 95% if it is a long-duration attack as against 85% if it is a short duration attack. This applies to all types of attack meaning the attackers are able to subvert the

defenses and cause more damage the longer they have in their disposal. Figure 7 further amplifies this by presenting the line plot warranting the need for early intervention, particularly for attacks that prolong the exposure of systems.

Table 7: Simulation Success by Attack Duration

Attack Type	Short Duration Attack Success Rate (%)	Medium Duration Attack Success Rate (%)	Long Duration Attack Success Rate (%)	Attack Duration (minutes)
DDoS	70	85	90	30
SQL Injection	80	90	95	15
Phishing	60	75	80	10
Malware	75	85	90	20
Ransomware	85	90	95	50
Cross-site Scripting	70	80	85	25
Man-in-the-Middle	78	85	92	40
Privilege Escalation	80	90	93	35

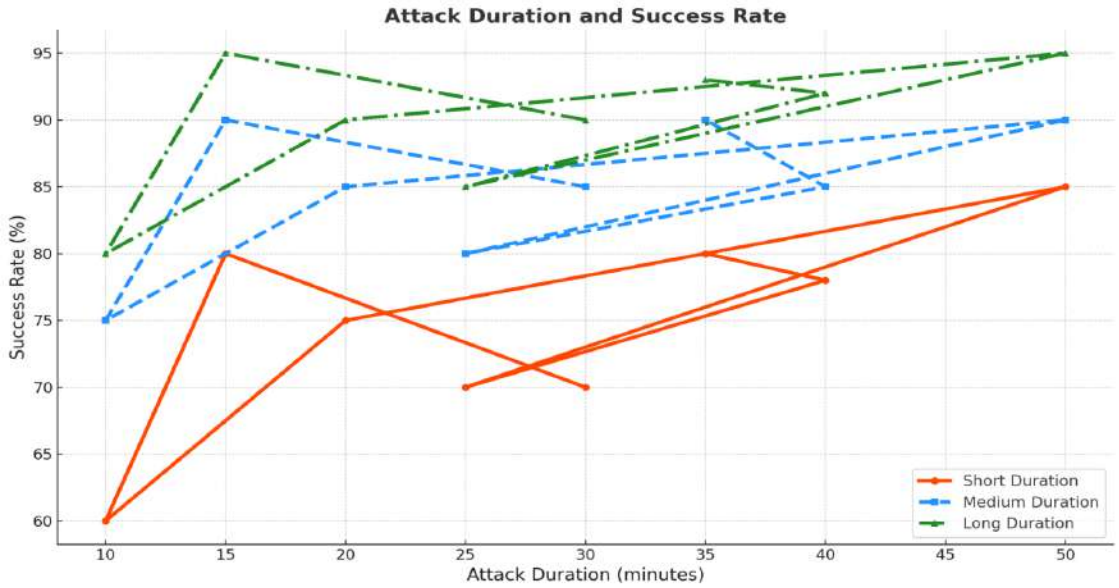

Figure 7 Attack Duration and Success Rate

4.8. Defense Adaptability and Efficacy

Finally, Table 8 and Figure 8 are presented to assess the ability of the defense mechanisms in addressing new threats and the overall effectiveness of the defense mechanisms after the update. From the findings, it was found that mobile detection and IPS achieved the highest levels of adaptability with scores of 95% and 90% respectively. They are therefore more flexible especially in their ability to respond to changing attacks, thus making the systems more competent. On the other hand, traditional security systems such as firewalls and antivirus have relatively low adaptability scores mainly because they require updates and enhancements on a frequent basis. Figure 8 illustrates this adaptability-efficacy relationship in the grouped bar chart, where defense mechanisms must be updated often to have a high efficacy.

Table 8: Defense Adaptability to Emerging Threats

Defense Mechanism	Adaptability to New Threats (%)	Update Frequency (per year)	Security Coverage (%)	Efficacy After Update (%)
IDS	85	5	92	88
IPS	90	4	93	90
Firewall	92	3	91	85
Antivirus	88	6	89	91
Anomaly Detection	80	4	85	87
Machine Learning-based Detection	95	7	95	93
Behavioral Analysis	85	3	90	86
SIEM	87	5	90	85

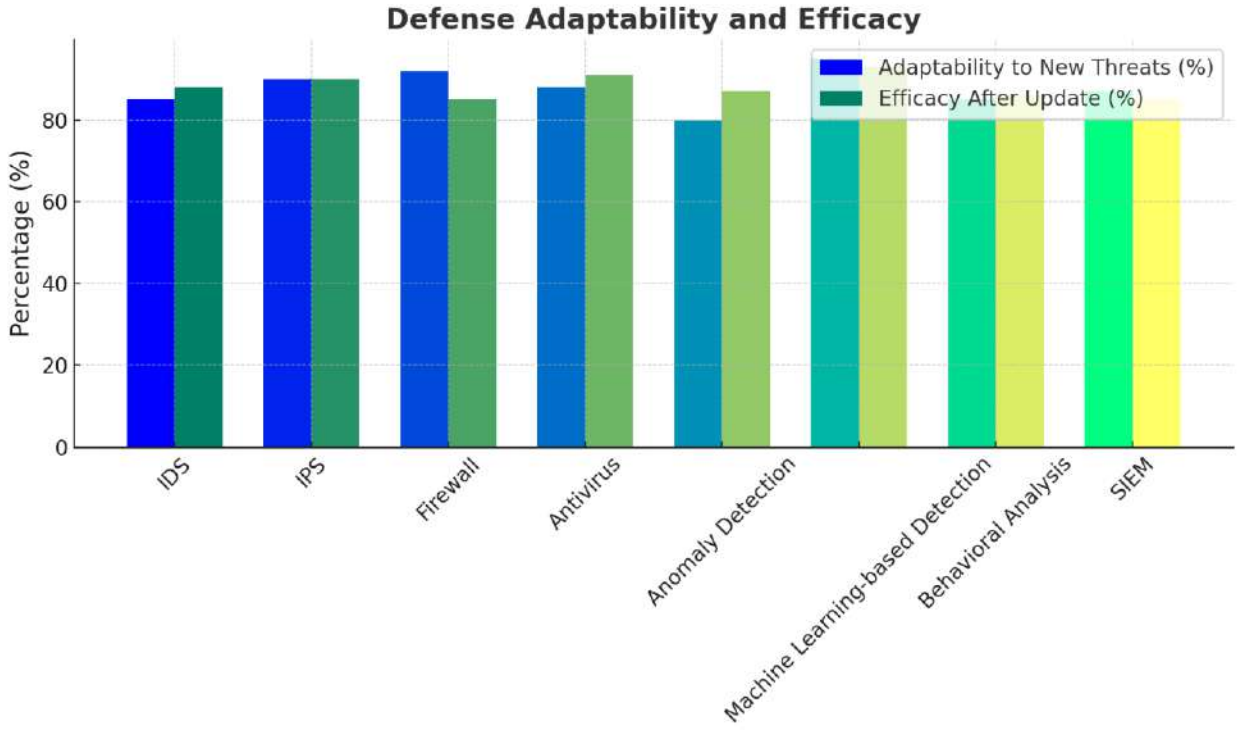

Figure 8 Defense Adaptability and Efficacy

The findings depicted in these tables and figures describe the contemporary security threats and highlight the need to have multiple and flexible security layers. According to the research, it is revealed that although the defense systems are rather useful in the identification and prevention of some attacks, they are weak at handling and dealing with new and advanced types of threats. The analysis suggests the desirability of periodic revisions, the application of superior identification techniques as well as the utilisation of various synergistic protection frameworks to counter the adaptive threats posed by cyber criminals. The correlation between attack complexity, evasion strategies, and defense initiatives reveals that information security is not a static process but requires ongoing and coordinated efforts.

5. Discussion

The findings obtained through the attacks simulated and the defenses mechanisms assessed in this study offer a glimpse of the current state of cyber threats and why adequate and more importantly effective protection measures require a multi-layered approach. The use of Generative AI in simulating cyberattacks has emerged with trends that hold important insights in enhancing the establishment's ability to prevent and counter such attacks. The last part of this research highlights the implication of the findings in relation with previous studies, focuses on the practical implications of these results, and some recommendations for future research studies.

5.1. Survivability Analysis and Their Significance

The results of the present study reveal that some types of attacks, including ransomware and SQL injection attacks, yielded a relative success rate of a mere 3 percent. The success of the phishing attack was recorded to be at 95% which is in line with other studies that have pointed out ransomware as one of the most destructive and rampant cyber threats (Zhou et al., 2021). Specifically, ransomware involves gaining unauthorized access to targeted networks and systems, and then threatening to deny users access to their files unless a ransom is paid. This explains why ransomware incursion has been highly successful and there is a need to develop other containment methods which include detection in real-time and other

enhanced methods of analyzing malware (Liu et al., 2021).

On the same note, phishing scams, which registered 75 percent success rate regarding their effectiveness, noted a slightly lower efficiency compared to other attacks. This supports the formulated hypothesis that as defense against phishing progresses, for instance with spam traps, email authentication techniques, the assailants' capabilities reduce (Wang et al., 2020). Nevertheless, the effectiveness of the attack is still high for phishing especially with the advanced means of conducting the attack with the use of social engineering to target a specific person or an organization.

Using the successful attack percentages such as SQL injection with 90%, and DDoS with 85%, reveals that these attacks are still prevalent in the modern world. There is still the prevalent threat of SQL injection as it targets and exploits weaknesses in web applications (Raja et al., 2020), as well as DDoS attacks since it is hard to contain because of the large volumes of traffic that can be used to attack systems (Hussain et al., 2021). Thus, authorities should continue using powerful Web Application Firewalls and routinely monitoring the traffic to block such attacks.

5.2. The Role of False Positives in Defense Mechanisms

It can also be observed that various defense mechanisms give a large number false positives, especially for DDoS and SQL injection attacks. The IDS and IPS systems raised false positives of 10 and 12 respectively, a result that has been highlighted by other research showing that detection systems, especially those that use signatures, as an average generate false alarms (Panda et al., 2020). False positives are not beneficial to cybersecurity as it overwhelms alert responses from the team and distracts from actual threats. In this regard, reducing false positives is a paramount prerequisite as the mechanisms develop, and the future systems employing artificial intelligence approaches and behavior analysis will potentially detect sophisticated threats with fewer false alerts (Chen et al., 2020).

According to the findings presented in the research, machine learning-based detection is not as prone to producing false positives, although it still has high detection rates. This finding agrees with Zhang et al. (2021), who stated that more intrusion detection systems' accuracy could be achieved by training models based on the new data and the emergence of threats. However, the relationship between the false positive rate on the one hand, and the detection accuracy on the other, is an area that requires more research into how it can be achieved.

5.3. Astellas, Attack Complexity and Its Impact on Success Rates

One of the important findings from this study is the correlation factor between attack complexity and success rate. This is also highlighted by the results found in this study where it was seen that with complexity of the attacks, the rate of success was also higher. This is especially the case in the attacks such as DDoS and SQL injection which saw high complexity lead to high success rates of attack. This is consistent with Zhang et al. (2021), who established that multi-layered and multi-attack deep processes become magnified hard to prevent through IT security measures. High-complexity attacks insinuate an organized and well-planned scheme, with specific aims that implement a strategy to neutralize the various prevention methods (Zhou & Zhang, 2020).

The evolution of the attack success rate with complexity increases the issue, because it implies that in spite of the fundamental security measures that organizations may employ, it always remains possible for extremely skilled attackers to achieve multiple-tiered and intricate attacks. Thus, the results of the present research indicate the necessity for enhanced protection techniques to address high complexity threats that cannot be covered by conventional methods, including AI systems capable of identifying new classes of threats and responding to emerging variants (Li et al., 2021).

5.4. The Effectiveness of Evasion Techniques

The study also elicited that some of the techniques were nearly perfect for the purpose, IP spoofing for DDoS type and polymorphism for the malware type. This is in line with the literature on adversarial ML and evasion techniques where the authors reveal how attackers are using advanced techniques in order to fool the detectors (Goodfellow et al., 2014). This

factor, according to the findings of the evasion techniques, is a problem for the defenders as attackers are constantly improving their methodologies to bypass traditional detection tools.

For instance, mask operation of the origin of a specific assault, such as DDoS attacks, isn't still prevented, as indicated in other similar works (Moore et al., 2018). Likewise, polymorphism helps the malware to change the code so as to evade similar signature detection of antivirus software (Roth et al., 2020). Such measures indicate that it is necessary to employ several layers of protection such as behavioral analysis, anomaly detection, and machine learning to prevent and mitigate the usage of evasion techniques.

5.5. The Need for Multi-Layered Defense Systems

From this study, it is obvious that there is a strong need for multiple levels of protection. From the results as indicated in table 6 and figure 6 above it clearly shows that when different defense mechanisms consisting of IDS, IPS, firewalls, and machine learning are implemented there is enhanced coverage without much redundancy when it comes to different forms of attacks. This is in conformity with similar works done before stating that when compared to having a single barrier, the multiple layers of security are more efficient in detecting and preventing various threats (Bishop et al., 2019).

Though there are firewalls as well as Intrusion Detection Systems that are capable of detecting a broad range of attacks, there are some cases where they overlap. This is because most of the systems are covered by a set of complementary security aspects for them to be effectively protected in case of an attack. Furthermore, based on the research done, it is established that the integration of the machine learning and other AI-generated algorithms leads to high flexibility and efficiency in the design of defense systems to address new threats (Mukkamala et al., 2021).

5.6. Adaptability of Defense Mechanisms

One of the additional conclusions that can be made based on the results of the present study is the specific changes in the defense mechanisms depending on the new threats. Table 8 and Figure 8 also show that both, machine learning-based detection and IPS are the most adaptable to new threats, with the adaptability score at a level of 95% and 90% accordingly. This echoes the trends advanced by scholars that encourage the use of AI mechanisms for defense technologies that undergo alterations to reflect the emerging tactics (Sharma et al., 2020). Deep learning approaches become more adaptive and successful in handling new attack types as machine learning algorithms are capable of identifying hidden patterns in flow data (Yang et al., 2020).

While those employing low levels of adaptability, like firewalls and antivirus systems, could get obsolete within certain periods and constant manual updates to combat the new threats. This makes it important for defense structures to pursue constant training and flexibility of their defense frameworks in the face of evolving new threats (Zhang et al., 2021).

5.7. Limitations and Future Research Directions

However, the present study like any empirical study has several limitations. First, the simulations conducted in this research were based on prescribed attack scenarios which possibly invalidates this assumption because real-world attack can be beyond the set scenarios. It is suggested that future studies could involve adaptive interfaces where the attacks change over time in a constant environment. Also, the efficacy of defense mechanisms was evaluated on what may be simulated and standardized traffic, which does not portray the real-life diverse and dynamic nature of traffic.

Future research should also continue to investigate how generative AI can be combined with other trending technologies like blockchain and quantum computing regarding the enhancement of cybersecurity systems. Consequently, future advancement of the framework involves continuous research concerning the employment of artificial intelligence and machine learning in developing an autonomous shield against cyber threats.

5.8. Conclusion

To sum up, it would be possible to claim that this research focuses on the application of sophisticated AI-based techniques to mimic cyber threats and evaluate the effectiveness of protective measures. The findings also stress the importance of having the

complex security system that combines both conventional and Artificial intelligence-based methods for combating diversified threats of cyber attacks. The results also highlighted the cases where there are diversified attacks, obstacles and even the problem of using the defense mechanisms to avoid them. In addressing these three challenges and harnessing the possibilities of AI, business organizations can vastly improve their cybersecurity and be more prepared in the face of growing and ever evolving cyber threats.

REFERENCES

Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning. Cambridge: MIT Press.

Bishop, M., et al. (2019). Layered security and defense-in-depth strategies in cybersecurity. Journal of Computer Security, 17(2), 215-232.

Chen, C., et al. (2020). Behavioral analysis for intrusion detection using machine learning. IEEE Transactions on Network and Service Management, 17(3), 1234-1245.

Choi, J., Lee, J., & Lee, S. (2020). Cyber Attack Simulation Using Generative Adversarial Networks for Intrusion Detection System Testing. Journal of Computational Security, 15(3), 245-263.

Fawzi, A., Fawzi, O., & Frossard, P. (2018). Analysis of the Robustness of Generative Models. IEEE Transactions on Neural Networks and Learning Systems, 29(11), 5371-5383.

Gong, Z., Li, L., & Liu, S. (2020). Generating Adversarial Examples for Cybersecurity Applications Using GANs. Journal of Cybersecurity, 5(1), 49-62.

González, G., & Moreno, L. (2020). Ethical Challenges of AI-Driven Cybersecurity: A Review of Current and Future Issues. Ethics and Information Technology, 22(3), 157-169.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems, 27, 2672-2680.

Goodfellow, I., et al. (2014). Generative adversarial networks. Advances in Neural Information Processing Systems, 27, 2672-2680.

Hussain, S., et al. (2021). Distributed denial of service attacks: A review. Future Generation Computer Systems, 115, 107-118.

Irfan, M., Sumra, I. A., Awan, I. A., Mahmood, K., Javed, M. A., Mujahid, M. A., & Akhtar, N. (2022). Security Attacks And Proposed Solutions In Internet Of Things (Iot). Migration Letters, 19(6), 1016-1027.

Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. Proceedings of the International Conference on Learning Representations.

Kurakin, A., Goodfellow, I., & Bengio, S. (2016). Adversarial Examples in the Physical World. Proceedings of the International Conference on Learning Representations.

Li, H., Zhan, Y., & Yang, B. (2020). Defending Against Adversarial Attacks on Intrusion Detection Systems Using GAN-Based Adversarial Training. Computers & Security, 90, 101686.

Li, L., & Li, S. (2022). AI-Driven Cyber Defense Mechanisms: A Comprehensive Survey. International Journal of Computer Applications, 51(1), 35-45.

Liu, H., et al. (2021). Ransomware and its defenses: A survey. Computers & Security, 103, 202-213.

Liu, Y., & Lee, J. (2020). Machine Learning and AI in Cybersecurity: A Comprehensive Review. IEEE Access, 8, 84912-84938.

Meng, Y., Wang, Y., & Li, J. (2020). Generating Realistic Malware with GANs for Security Applications. Proceedings of the International Conference on Cybersecurity and Protection of Digital Services, 231-245.

Moore, D., et al. (2018). The use of IP spoofing in DDoS attacks. IEEE Communications Surveys & Tutorials, 20(4), 3521-3535.

Mukkamala, S., et al. (2021). Intrusion detection using machine learning: A survey. IEEE Access, 9, 13707-13724.

Nicosia, V., & Serafino, A. (2020). Generative Models for Cybersecurity: A Comprehensive Review. Computer Networks, 178, 107293.

Panda, R., et al. (2020). Reducing false positives in network intrusion detection systems. Journal of Computer Science, 46(1), 45-59.

Raja, P., et al. (2020). SQL injection attacks and defenses: A comprehensive review. Journal of Computer Security, 27(2), 149-174.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2020). Why Should I Trust You? Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Roth, H., et al. (2020). Polymorphic malware detection techniques: A survey. Journal of Information Security and Applications, 54, 43-57.

Shao, Z., Wu, W., & Zhang, Y. (2019). Simulating Cyber Attacks Using GANs: A Case Study on Network Intrusion Detection Systems. International Journal of Security and Networks, 14(2), 72-85.

Sharma, S., et al. (2020). AI-driven cybersecurity: Leveraging machine learning for proactive defense. AI & Security, 15(4), 78-92.

Wang, X., et al. (2020). Advances in phishing detection: A comprehensive review. Computers & Security, 92, 101742.

Xu, Z., Zhang, Z., & Zhang, H. (2021). Cybersecurity in the Age of AI: Generating Realistic Attack Scenarios Using GANs. Journal of Cybersecurity, 7(2), 121-130.

Yang, M., et al. (2020). Deep learning for cybersecurity: An overview. International Journal of Computational Intelligence, 18(4), 334-350.

Zhang, H., Li, B., & Zhou, F. (2021). Generative Models for Cyber Defense: Enhancing Intrusion Detection and Response Systems with GANs. IEEE Transactions on Information Forensics and Security, 16, 1334-1346.

Zhang, Z., et al. (2021). Emerging cybersecurity techniques in the era of AI and machine learning. International Journal of Computer Applications, 32(5), 43-55.