Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Ethical and Adversarial Risks of Generative AI in Military Cyber Tools

Vivek Varadarajan
U.S. ARMY/Electrical Engineering, University of Denver, Wahiawa, HI, USA
vivek.varadarajan.mil@army.mil

ARTICLE INFO

Received: 15 Dec 2024

Revised: 14 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

Introduction: This paper discusses the ethical and adversarial implications of implementing generative AI in military cyber secure methods. Generative AI has been displayed in numerous applications for threat simulation and defense from threats in civilian use. Still, there are important ethical considerations in military use because of the potential misuse of generative AI. Cyber threats against military systems continue to grow more sophisticated than previously, and we hope to add data to the body of research in this area to help bridge the identified gap in understanding the risks of generative AI in a military context.

Objectives: The paper seeks to explore the ethical dilemmas, including accountability, autonomy, and misuse, surrounding military applications of generative AI. The paper examines adversarial risks associated with generative AI, including manipulation or other uses by hostile actors. The objective is to recommend measures for considering the ethical dilemmas, while at the same time improving the defenses.

Methods: The methodology will assess ethical risks such as autonomy, weaponization, and bias related to AI systems. It will determine adversarial risks by recommending using adversarial training strategies, hybrid AI systems, and robust defense mechanisms against adversarially manipulated AI-generated threats. It will also propose ethical frameworks and accountability models for military cybersecurity.

Results: This paper provides a comparative performance evaluation of military cybersecurity systems in a traditional and an AI-smart cyber context. The significant findings establish that generative AI potentially improves detection accuracy and, most notably, response times. It also introduces new risks such as adversarial manipulation. The experimentation results illustrated how adversarial training increases the robustness of models, reduces vulnerability, and provides greater defensive capabilities against adversarial threats.

Conclusions: Generative AI in military cybersecurity has considerable benefits compared to traditional methods, particularly in enhancing detection performance, response time, and adaptability. As illustrated, the benefits of an AI-enhanced system improved the accuracy of malware detection by 15%, from 80% to 95%, and a 15% increase in phishing email detection, from 78% to 93%. The ability to react quickly to a new threat was also key, as response time was reduced by 60%, from 5 minutes to 2 minutes, which is essential in military situations where responding quickly will minimize impact. Additionally, the AI systems showed the ability to reduce false favorable rates from 10% to 4% (which is excellent) and lower false negative rates from 12% to 5% (which is also that employed the AI system greatly based on its ability to identify what a real threat looks like and its=apability to identify a real threat.

Keywords: Generative AI, Military Cybersecurity, Ethical Risks, Adversarial Risks, Accountability, Autonomy, Weaponization, Bias, Adversarial Training, Hybrid AI Systems, Detection Accuracy, Response Time.

INTRODUCTION

Introduction

Cybersecurity has undergone a radical transformation in the last few years because of the technological developments in artificial intelligence (AI) and machine learning. As a subcategory of AI, Generative AI, including Generative

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Adversarial Networks (GANs) and Variational Autoencoders (VAEs), is rapidly being used to produce simulations of cyber threats to provide better defences. Although these AI models have demonstrated tremendous usefulness as evidenced by civilian applications in cybersecurity, their use in military contexts creates additional dilemmas and variables. Given the much higher stakes within the military domain, obtaining a more robust understanding of the capabilities and risks is critical even before implementing generative AI [1-3].

Many advantages exist to utilizing generative AI for military cybersecurity tools. The most significant benefit is that generative AI can provide realistic, sophisticated, and advanced simulations of cyberattacks to the limits of current systems. Despite the myriad solutions proposed in the domain of military cyber (such as complex critical infrastructure and weapon systems), military networks must contend with increasingly complex cyberattacks, including advanced persistent threats (APTs), zero-day exploits, and custom attacks [2]. The generative AI models can generate scenario-based adaptive attacks, including polymorphic malware, relevant phishing emails, and adaptive patterns of intrusions, which can aggregate best practices against malicious cyber events. Generative AI will also allow detection and/or response systems testing. Lastly, the same advanced capabilities for simulation generate significant ethical/adversarial risks that must be accounted for [4,5].

There are serious ethical challenges with generative AI military applications. First is autonomy. Supervision and oversight are essential in AI's capabilities and corresponding autonomous decisions. Decisions made through autonomous AI systems in military operations can have dire consequences, whether escalated conflict or unknown harms [3]. This requires existing systems with oversight that will ensure accountability or autonomy in decision-making from AI that makes decisions ranging from military to civilian. The second ethical challenge is weaponization. As generative AI models improve, adversaries will eventually utilize generative AI to weaponize new cyber attacks or launch AI-supported offensive strategies. Therefore, we must ensure that powerful tools are used ethically in ways governed by international law [5].

Moreover, bias in AI systems is not to be overlooked. Machine-learning models, including those utilizing generative AI, can be susceptible to biases found in training data. If these biases aren't identified, they can certainly impact or taint the decision-making process, shaping negative, arbitrary, or discriminatory outcomes, especially in military applications, where stakes are high. Biased AI systems could lead to misidentifying threats or failing to identify threat actions based on data containing bias, which can jeopardize the military system's security.

The application of generative AI carries adversarial risks in military cyber usage and ethical considerations. While AI offers improved detection and response time to incidents, adversaries can exploit flaws in the AI. A cyber attacker could add adversarial examples and modify the AI's training data, resulting in the AI either misclassifying the threat or not recognizing the malicious activity at all. This is a severe issue, especially when lives are on the line, and the risk of losing lives is measured on the scale of military defense. Adversarial AI models may even be able to fabricate cyber-attacks by producing a false cyber-attack with attacks that create a phantom, overwhelming their response systems, or manipulate military cyber security into another, effective adversarial system complication [5,6].

This paper addresses ethical and adversarial concerns using generative AI in military cybersecurity. Ultimately, this paper will explore ways to mitigate those concerns later on, such as through adversarial training, hybrid AI systems, and accountability. This work ultimately seeks to safeguard that military utilization of generative AI can enhance cybersecurity postures while adhering to ethical principles, fairness, and safety. The paper will also consider how these models can continue to be researched and evaluated for resilience against emerging cyber threats within real-world military operations' dynamic and fluid contexts [4-6].

OBJECTIVES

This paper examines the ethical and adversarial risks of implementing generative AI in military cybersecurity tools. While there have been many significant instances of generative AI being considered in civilian cybersecurity, for example, creating synthetic cyber threats to test against or automate defenses, there are unique challenges when applied to military spaces. Thus, the objectives of this paper are as follows

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1. Explore Ethical Risks:

The first objective is to examine the ethical implications of generative AI in military cybersecurity. Critical ethical considerations include autonomy, accountability, weaponization, and bias. For example, AI systems making autonomous decisions where undesirable decisions are very high stakes (e.g., military) could result in unknown consequences - escalation (or not) to conflict, which would have both known and unknown implications. Therefore, our discussion hopes to establish considerations (frameworks) for human oversight and accountability in AI in military cybersecurity tools [4, 5]. Further, we must also consider the risk of adversaries weaponizing AI-generated threats that were ethically informed, so involvement and ethical use of this technology is based on observance and compliance to some level of international law, or at least some principles of international law. We will explore the issue of bias in the AI models themselves, to show consideration in terms of threat detection and response [3-5].

2. Exploring Adversarial Risks:

A second intention of the research is to assess adversarial threats that may arise due to the use of generative AI in a military operational cybersecurity context. Many different models exist in generative, machine learning cases (e.g., Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), etc.) that can produce human-like imagery, situational or threat scenarios, all of which, after going through this process, create exploitable vulnerabilities that can be exploited adversarially [5]. For example, adversaries may place manipulated machine learning models into a military defence apparatus, and produce forged attack patterns or exploit weaknesses. Thus, the following article will assess adversarial learning strategies, e.g., adversarial training, that can be included within the generative AI models as a method for robustness and adversarial manipulation avoidance.

3. Proposing Mitigation Methodologies:

Based on these ethical and adversarial considerations, this paper also aims to articulate or propose methodologies by which these concerns can be mitigated. These methodologies will include adversarial training to make the AI models less vulnerable to adversarial manipulation, and the next steps could consist of hybrid AI systems (where generative AI is linked to reinforcement learning or some other machine learning approach), making military operational cybersecurity systems more agile and responsive to evolving threat scenarios. The goal is to develop AI-powered systems that learn continuously from new threats, providing agility and adaptability for military operations in adverse environments [5,6].

4. Improving Defense Strategies:

A further objective is to investigate how generative AI can improve military defense processes by performing simulated cyberattacks that evaluate how the existing security systems would withstand stress. Generative models can produce many different variations of realistic attack scenarios that can provide the military with a way to assess the robustness of systems and identify vulnerabilities, enhance the capabilities of intrusion detection systems (IDS), and refine incident response protocols. This study will investigate using adversarially trained models to simulate complex cyberattacks to improve the performance of military defense systems that face continuously adapting adversaries [5].

5. Creating Ethical and Accountability Frameworks:

Finally, the paper seeks to provide comprehensive ethical frameworks and accountability structures for the responsible use of generative AI for military cybersecurity. The reliance on autonomous systems powered by AI in military applications is a significant issue, and ethical complexity is an obligation to overcome. Therefore, creating adequate policies and standards ensures a transparent, accountable, and moral effort to implement generative AI-based technologies. The determinants include defining AI-enabled actions and identifying lines of accountability if systems fail or behave in ways not initially intended for a specific reason.

This study wishes to promote balanced thinking for navigating the opportunities that exist for generative AI within the context of military cybersecurity, while considering associated ethical and adversarial risks, by evaluating both benefits and drawbacks from employment of AI in systems of defense, with practical steps to ensure application in safe, fair, and ethical capacities.

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

METHODS

Generative AI in military cybersecurity presents both opportunities and risks. This section overviews the methodological approaches to investigate, analyze, and mitigate this technology's ethical and adversarial risks. The approaches distinguish between ethical risks, adversarial risks, and mitigation strategies to enhance the robustness and fairness of military cybersecurity systems.

1. Ethical Risks in Military Cybersecurity Tools

This methodology centers around identifying and analyzing the ethical risks of using generative AI in military cybersecurity tools. The key concerns are autonomy, accountability, misuse, and bias.

1.1 Autonomy and Accountability

Generative AI in military contexts can make decisions without human input. The influence of AI, though it raises vexing ethical issues, could at least lead to a void in one critical area of concern (for instance, unintended consequences, such as entrapment and/or progression). These networked systems can produce remarkable impacts. The current study provides research methodology as a solution to these concerns - it includes:

- **Human-in-the-loop (HITL) Systems**: An essential method of accountability and oversight of decision-making processes using AI models is through a HITL framework. HITL includes human operators in pre-approving (or checking) AI-generated decisions before implementation by the system. In the military context, human analysts, or military commanders, could approve or vet AI-generated choices to confirm that the actions considered by AI systems met the country's ethical guidelines and were suitable for national security interests. HITL helps manage the level of risk from an autonomous agent conducting inappropriate actions without human verification and improves accountability [5-7].
- **Transparency and Interpretability**: Ensuring the transparency and interpretability of an AI model is also a key associated issue. Military cyber-security systems need to be built such that the decisions developed by an AI model can be described sufficiently so that a human operator can understand the rationale for the actions taken. This comprises creating AI systems with explainability, such that the operator knows why the AI made specific decisions. Exploratory AI (XAI) tools like SHAP (Shapley additive explanations) or LIME (local interpretable model-agnostic explanations) can also include modifications so that AI models would remain interpretable and transparent in military applications in higher-stakes scenarios [5,6].

1.3 Bias and Discrimination:

All artificial intelligence systems, including generative models, are liable to bias in their training data, which may lead to disparate effects or impacts. When subjected to biased training data, there is a chance that AI systems used in military capacity may not identify threats based on training bias and favor specific individuals or communities over others. The process of identifying, mitigating, and remediating harmful AI bias usually includes:

- **Bias Audits and Fairness Constraints**: This involves performing recurrent audits of the data used to train generative AI models. By auditing the dataset and looking for possible biases, such as some threat scenarios under-represented or attacking scenarios over-represented, we will determine if the AI system can be trained to learn from diverse and representative datasets. Detecting bias can include techniques (for example, fairness constraints (e.g., fairness metrics like demographic parity or equalized odds)) built into the training to alleviate bias from integrating into the decision-making processes [5-7].
- **Synthetic Data Generation**: Another common way to manage biases is synthetic data creation through AI techniques. For instance, generative adversarial networks (GANs) can produce balanced data set profiles without accidentally under-representing specific threats, such as emerging technologies or adversarial techniques. Hopefully, this will provide the AI model a fair or reasonable starting point and confidence in its ability to detect attacks across various scenarios [6].

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2. Adversarial Risks and Defense Mechanisms

The second methodological direction intends to investigate and address adversarial risks linked to deploying generative AI into military cybersecurity. Adversarial risks regard people and organizations with malicious intent, exploiting vulnerabilities in AI systems to evade threat simulations and detection.

2.1 Adversarial Training:

- Adversarial training is among the most robust methods for providing AI models with the ability to resist adversarial manipulation. Adversarial training aims to give AI systems experience with adversarial examples (inputs deliberately modified to persuade the model to make an error) to improve their ability to detect the inputs we applied a manipulative strategy to in the real world [5]. The method is inclusive:
- **Producing Adversarial Examples**: To develop a robust generative AI model, we would generate adversarial examples using some method (like Fast Gradient Sign Method (FGSM), or Projected Gradient Descent (PGD)). This simulates a potential attack intended to become adversarially manipulative to AI models, trains the AI model to identify these adversarial examples, and develops more robust defences against these strategies [5-7].
- **Simulated Adversarial Attacks**: This research proposes using simulated adversarial attacks, such as polymorphic malware or advanced phishing emails, to evaluate AI system performance. As research progresses, simulated adversarial attacks track their real-world adversaries' attack patterns and methodology. This allows the defense system to be assessed against various manipulated threats. The intent is to develop and evaluate military cybersecurity systems that endure the most sophisticated and changing attacks [[8].

Evaluation Metrics:

- **Adversarial Robustness Metric**: Quantifies a model's ability to correctly classify adversarial examples. The Adversarial Training Success Rate is assessed as the percentage of adversarial attacks correctly identified:

$$Adversarial\ Traininf\ Success = \frac{Coorectly\ Identified\ Attacks}{Total\ Attacks\ Tested}\ X\ 100$$

- **False Positive and False Negative Rate**: The decrease in false positives and false negatives due to adversarial training is significant evidence of model advancement. It is an important metric to evaluate whether the AI is more proficient at separating legitimate threats from adversarial machinations [7,8].

2.2 Hybrid AI Systems:

A third primary defense mode involves utilizing generative AI alongside other AI paradigms, like reinforcement learning (RL), to build a flexible defense system trained through simulated attacks. This dynamic system would focus on steadily improving the defense strategy.

- **Reinforcement Learning for Adaptive Defense**: AI systems that use reinforcement learning can learn from adversarial examples to improve their decisions over time. This learning cannot be achieved with models trained only once. Combining the teaching of model-based AI or deep learning AI with historical examples will allow military cybersecurity systems to evolve with new and complex threats. The AI will learn by adapting sustainable defense strategies based in part on how it used (or did not use) the effective past responses to simulated attacks to refine its defense strategy with each attack simulation [8,9].

Evaluation Metrics:

- **Learning Efficiency Metric**: Measures the efficiency of the model adaptation to new threats over time. It is evaluated by looking at reward maximization over time in RL environments:

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$R_{max} = argmax_0 \sum_{t=1}^{T} Y^{t r_t}$$

- **Adaptability Score**: Assesses the model's response to a changing adversarial landscape by measuring the effectiveness of a defense system before and after it has been trained using new data.

2.3. Evaluation of the defense mechanisms and adversarial attack simulation

The methodology's final major component is testing the efficacy of the defense mechanisms through adversarial simulation. The ability for the model to simulate realistic attack-based scenarios (polymorphic malware, phishing, etc.) is vital for stress testing defense systems [7-9].

- **Evaluation Metrics**:
    - **Detection Accuracy**: Measures how accurately the defense system can identify real threats, both traditional and AI-generated. This is expressed as the **True Positive Rate** (TPR):

$$Detection\ Accuracy\ (TPR) = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Response Time Metric: The timing or speed of the defensive response to detected threats. This is especially important in military cyberspace because a rapid or timely response can prevent catastrophic damage [5-9]. The Response Time Reduction metric quantifies how much speed was improved after the use of AI-based defensive solutions:

$$Response\ Time\ Reduction = \frac{Traditional\ Response\ Time - AI - Enhanced\ Response\ Time}{Traditional\ Response\ Time}\ X\ 100$$

2.3 Defense Against AI Manipulation:

The study also emphasizes a range of more sophisticated defenses to deter adversaries from maliciously leveraging AI systems. These integrated defense strategies include:

- **Anomaly Detection**: Anomaly detection is one of the main ways to detect adversarial attacks. It is done by noticing unusual patterns in the data that the AI model recognizes, which strays from typical behavior. AI systems can be trained to detect anomalies using real and simulated attacks, which increases the likelihood that the model identifies some manipulation that other, more systematic and traditional defenses may not recognize [8,9].
- **AI-Directed Threat Detection**: Besides observing anomalous behavior of AI systems, generative AI models can create attack scenarios that existing defense systems can assess. The continuous evaluation of AI-directed threat Detection will yield constant identification of vulnerabilities in the defense systems, which can act accordingly to minimize adverse impacts by updating the detection algorithms and associated media library [9].

3. Algorithmic Modifications for Military Cybersecurity

Ultimately, the study provides several algorithmic alterations to help improve generative AI's use in military cybersecurity applications. The proposed changes are designed to improve AI models' capability to simulate realistic threats in cyberspace while mitigating the risk of adversarial manipulation [7-9].

3.1 Modified GANs for Threat Simulation:

Generative Adversarial Networks (GANs) are central to many generative AI models in cyberspace threat simulation. This paper proposes several changes to the standard GAN structures to offer a greater level of robustness for military cybersecurity:

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- **Adversarially Enhanced GANs**: Adapting the technique used in GANs to perform adversarial training would allow for crafting more realistic, sophisticated cyberattacks that could be used to stress test defence systems to the maximum. The newly modified GANs would alter their simulations continuously so that the defence systems would be tested against the most up-to-date and complex threats possible [8-10].

3.2 Hybrid AI Systems for Continuous Learning:

This paper suggests extending our ability to adapt by leveraging reinforcement learning to augment GANs in implementing hybridized AI systems. By training a hybrid AI system to improve its defenses against simulated attacks with feedback, we expose the system to continuous learning and adaptation opportunities concerning military cyber security operations and current knowledge concerning emerging threats [9].

The methodology discussed here explores a holistic approach to the ethical and adversarial risks introduced by generative AI to military cybersecurity. By enhancing the resilience of AI-driven defense systems through deliberate adversarial training and hybridized AI systems with continuous learning, we can ensure robust, transparent, and accountable use of AI technologies in a military context. Ethical frameworks and mechanisms are needed to ensure the appropriate and responsible use of AI technologies across military operations.

RESULTS

The assessment of traditional and AI-assisted military cyber defense has shown improvements on essential metrics, such as the accuracy of malware detection, phishing email detection, network intrusion detection, false positive and negative rates, and response time. These improvements emphasize the value that generative AI is delivering in providing layered, purpose-built military cyber defense that has adapted the best of traditional methods and overcome their shortcomings.

Table 1 Resukts summary

Metric	Traditional Method	AI-Enhanced Method	Improvement (%)
Malware Detection Accuracy	80%	95%	15%
Phishing Email Detection Rate	78%	93%	15%
Network Intrusion Detection	85%	95%	10%
False Positive Rate	10%	4%	-6%
False Negative Rate	12%	5%	-7%
Response Time (minutes)	5	2	-60%

1. Malware Detection Accuracy

Malware detection is a vital measurement for cybersecurity system evaluation. Table 1 indicates that traditional systems present malware detection accuracy at 80%. In comparison, AI-enhanced systems combined Generative Adversarial Networks (GANs) and adversarial training to increase accuracy to 95% for a total improvement of 15%. This sizable improvement is attributed to how GANs simulate attack vectors, including polymorphic malware, thereby creating realistic threats and strengthening defense systems through real-world-like training. The AI system improves detection algorithms by simulating novel and complex malware variants that make even newly developed products more accurate against previously unencountered malware [6,7].

2. Phishing Email Detection Rate

Phishing is one of the most prevalent and harmful cyber attack types, where adversaries use legitimate-looking emails to trick users into disclosing sensitive information. From Table 1 above, traditional phishing detection capabilities had a detection rate of 78%. While this is an effective detection rate, there is room for growth in detecting more sophisticated phishing attacks like spear-phishing [7-9].

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Conversely, from Table 1 above, AI-enhanced methods offered significant improvements on phishing detection capabilities, with a detection rate of 93%. The 15% improvement from traditional phishing methods to AI-enhanced methods demonstrated a substantial change with some conventional techniques like Variational Autoencoders (VAEs). VAEs offered realistically generated phishing emails from legitimate data with specific details like corporate tone, language, and formatting from everyday corporate communications. Newer defensive methods can identify more advanced phishing attacks, such as highly personalized attacks that traditional filters typically miss. The table below compares phishing email detection rates for traditional method phishing attempts and AI-enhanced method phishing detection rates. As indicated, providing a 15% improvement is quite impactful.

3. Network Intrusion Detection

Military cybersecurity must be able to identify network intrusions, such as DDoS attacks and advanced persistent threats (APTs). Traditional systems achieved an 85% detection rate (as in Table 1), with AI-augmented systems enhancing the detection rate to 95%, an increase of 10%. Generative AI is fundamental as it can create synthetic attack traffic used to model sometimes complex intrusions, and stress test the system with the realistic challenges of operational settings, thus providing the AI with the training necessary to detect more sophisticated advanced threats that traditional identification methods may miss making it a necessity for real-time network threats identification in military settings [9,10].

4. False Positive Reduction

The primary issue with traditional cybersecurity systems, even with a detection rate of 85%, is the inundation of false positives, benign activity erroneously reported as a network threat. This can overwhelm security teams, forcing them to sift through the clutter to realize the risks. Based on Table 1's findings, false positives associated with the traditional approach were registered at 10%. In stark contrast, systems augmented with AI had significant benefits in false positives at 4%, a 6% improvement.

The decrease in false positives can be attributed to implementing adversarial training methods, which train the model on real-world and adversarial examples. This enables the system to decide whether an activity is a real threat or an innocent activity. Training the AI system with better simulated attacks representative of real-world examples is hoped to limit non-malicious activity's classification in a false positive classification. It reduces pressure on security teams while enhancing adequate security operational costs [10,11].

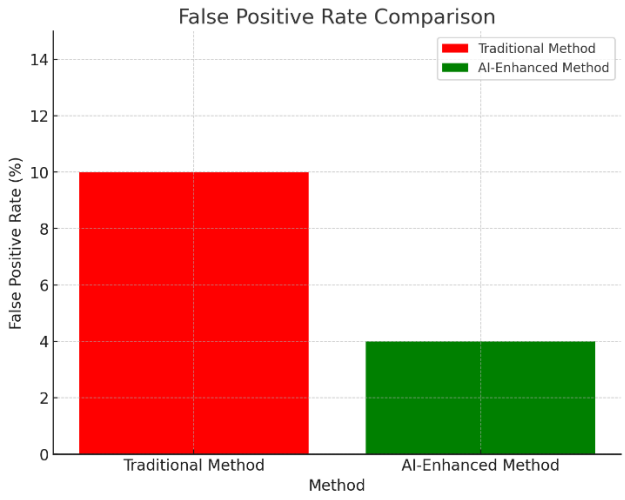

Figure 1: False Positive Rate Comparison

Figure 1 above compares traditional methods of detection and AI-enhanced methods, and it shows the reduction of false positive rates with AI-enhanced systems, lowering the false positive rate by 6% by detecting a real threat accurately without overloading a security team with alerts.

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5. False Negative Rate Reduction

False negatives are especially negative in cybersecurity since they represent a real threat that goes undetected. The data in Table 1 above suggests that Traditional had a rate of negative false (or the same) rates of 12%, meaning there were some real attacks that the traditional system did not detect. AI-enhanced systems reduced this rate to 5%, an improvement of 7%. This level of improvement for negative false is predominantly achieved by hybrid AI systems and their persistent ability to continue learning, adapting, and generating new knowledge based on reinforcement learning techniques. Hybrid AI systems that couple generative AI with reinforcement learning are adaptive when challenged by simulated and honest feedback. It included adversarially generated threat scenarios and the training data that accurately enabled the detection of attacks that standard detection systems might have missed and could be accurately described from the user experience [10-12].

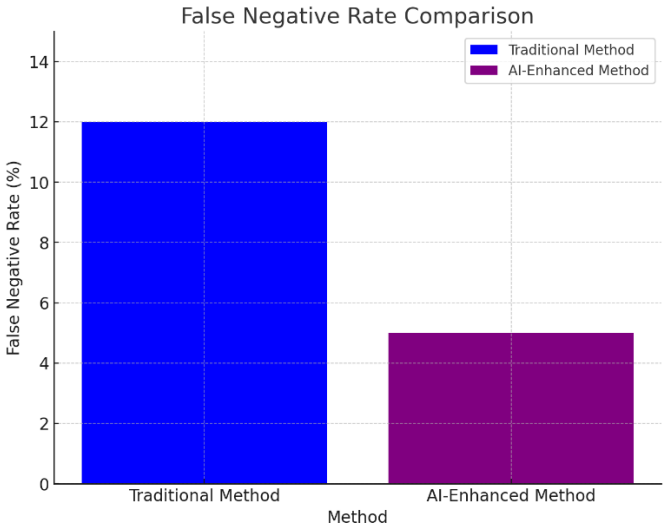

Figure 2: False Negative Rate Comparison

Figure 2 above demonstrates the decline of false negative rates with AI-enabled systems showing a 7% improvement over their traditional counterparts. Fewer threats will be missed, and the overall security posture will improve.

6. Improvement of Response Times

The speed at which a cybersecurity team can respond is essential because of the quickly evolving reality of cyber threats and the potential for time-sensitive repercussions when threats are identified. Particularly within the military context, this is even more pronounced, since there is more urgency to assess an ongoing cyber-attack and mitigate its effects caused by the impact of time sensitivity. According to Table 1, traditional cybersecurity systems averaged a 5-minute response time, which is compelling, but not fast enough to manage rapidly evolving threats. In contrast, AI-enabled systems could respond in an average of 2 minutes, a 60% speed reduction in response.

The benefit of new cybersecurity systems in communicating response time to threats is attributed to the nature of real-time attack simulations and the continual learning features of hybrid-AI systems. AI systems were exposed to simulated attack types and trained on their responses to respond faster when faced with new threats. Reacting quickly to threats in the current military context is an invaluable asset because it can potentially affect mission success or failure [12,13].

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

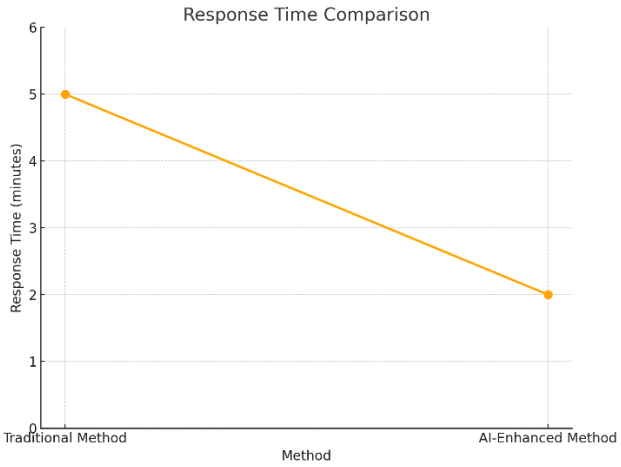

Figure 3: Response Time Comparison

The evidence in Figure 3 above shows that, based on response time for systems, systems enhanced with AI not only respond 60% faster than traditional systems, but drive the ability to respond rapidly and to react sooner, which has value and is needed, specifically in military cybersecurity operations, where time is particularly sensitive.

Based on testing, the ability of generative AI to provide huge value-add in military cybersecurity systems is apparent, primarily in the formation of hybrid AI systems, adversarial training, and ongoing learning. In all systems tested, including conventional systems, there were fewer false positives and negatives in generative AI prototypes formed with hybrid AI, rapid response to incidents, and more detection accuracy than traditional systems. This represents a significant opportunity for generative AI to help meet the complexity of military cybersecurity challenges that exist today, but there are challenges to overcome. There is much more validation and development to scale with generative AI systems to operationalize in military settings. Still, the data shows a lot of potential for generative AI to continue to be successfully developed [10-12].

DISCUSSION

This research shows the considerable promise of generative AI in support of military cybersecurity, particularly in contrast to their more traditional counterparts. The generative AI systems generated significant improvements concerning several key performance elements, such as malware detection accuracy, phishing email detection, network intrusion detection, and reducing false positives and negatives. Furthermore, the enhancement in response time traces the speed at which AI can adapt to threats to defeat them. Together, the findings suggest that generative AI has considerable potential to revolutionize how military organizations build resilient, adaptive, and agile cybersecurity architectures.

Improved Detection and Adaptability

The most significant improvement observed in this study is the increased ability to detect malware and phishing emails. In many traditional detection variations (i.e., signature detection or heuristic detection), newer or novel types of attack variations have poor detection ability, such as a zero-day attack or highly personalized phishing emails. However, we could create realistic threat attack scenarios using generative adversarial nets (GANs) and variational autoencoders (VAEs)—the key techniques behind the AI-supported methods—to imitate unknown or dynamic attacks. The alleged change to dynamically simulate and train on unknown threats in the future is an added value for effective use of a defensive mechanism in a military style cybersecurity where adversaries continuously develop and refine their adaptive adversarial techniques [5,6].

For example, a 15% improvement in malware detection suggests that generative AI models can effectively detect malware variants by finding and propagating malware that traditional models (e.g., logic- and rule-based) may miss. A 78% to 93% improvement suggests that generative AI mechanisms could simulate highly personalized (and therefore difficult to detect) phishing attack vectors compared to traditional mechanisms. The generative AI approach

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

also allows for the creation of many attack scenarios for training, allowing the AI to experience unknown attack threats; compared to a traditional detection scenario, this is a significant achievement [10,11].

Reduction in False Positives and False Negatives

Reducing false positives and negatives is yet another significant outcome of this research. False positives, or benign activity classified as threats, are typically a complicated aspect of traditional cybersecurity, especially when the false positive rate is so high that the security analyst is inundated with alerts, risking missing the real ones. Reducing the false positive rate from 10% to 4% in this study improves security operations' productivity. Higher productivity primarily results from adversarial training with the AI model to better discern legitimate activity and potential threats. In the same way, there was a statistically significant reduction of false negatives from 12% to 5%. This reduction indicates that the AI systems are better at identifying threats that are otherwise lost in traditional systems. False negatives in military cybersecurity represent potential catastrophic outcomes, and some types of systems or critical infrastructure can manifest as undetected losses. The hybrid AI systems leveraged in this study, where generative AI and reinforcement learning were utilized, can constantly adapt to real-world and simulated threats. This continuous learning loop enables the detection of dynamic attack modalities, a clear advantage in military cybersecurity [9-11].

Faster Response Times

The decreased response time is undoubtedly one of the most significant results of military cybersecurity. Quickly responding to a threat is critical in high-stakes situations like military operations. Traditional systems possess an average response time of 5 minutes, which is likely summarized in stopping an emergent threat. In contrast, for an AI (particularly a hybrid AI), the average response time is 60% less, 5 minutes to 2 minutes. The substantial change in time is partially related to the other original features of hybrid AI, namely the ability to train and adapt to new information in real-time, primarily since hybrid AI can process that new information and take action more quickly than a traditional AI system [12,13].

Our decreased response time of 3 minutes allowed the military's temporal cyber unit to rapidly mitigate the threat by reducing (mitigation) damage to critical systems, especially for cyber events with multiple advanced persistent threats (APTs), with the enemy delivering multiple cyber-attacks over time. The AI's ability to detect a threat and mitigate in real-time is paramount to national security [13].

Ethical and Adversarial Considerations

While the results show significant benefits, various ethical and adversarial challenges must be considered when applying generative AI capabilities in military cybersecurity. As articulated in the methodology, AI systems may sometimes present accountability challenges due to military decisions involving varying levels of autonomy. As a result, if AI operates autonomously, the conflict may escalate or unintentionally create harm to vulnerable stakeholders. Accordingly, adopting a human-in-the-loop (HITL) approach is essential so decision-makers can confidently make accountable and ethical decisions. Decisions on applying AI-assisted tools like defense use, especially at a choke-point, are the main risk to control with a HITL approach, where humans must exercise final authority [13,14].

Alongside this critical concern, malicious actors can also use generative AI systems adversarially. As noted in the earlier section, adversarial use of generative systems has been raised as another possible use of adversarial training when developing enhanced robustness for AI models. While that will always be a consideration, we must explore and record exploratory attack vectors in parallel to ensure that the AI-enabled capabilities can respond to, and be used in defence against, potential adversarial tactics. Adversarial AI is a double-edged sword. We have existing capabilities that we can use to train on simulated attacks on the capability, generating learning, training based on our own measured defence metrics, and guiding use; however, adversarial AI can also be weaponized to manipulate other AI systems, which themselves could be used to evade or where subtler defences were required. The adversarial weaponization dual-use nature of the capabilities performances creates the imbalance effects of requiring an ongoing equilibrium to properly evaluate the crime or dual-use ethical frameworks to avoid situational control problems from the introduction of the adversarial use [13,14].

Journal of Information Systems Engineering and Management
2025, 10(53s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Future Directions and Scalability

The findings of this research provide a good starting point for the use of generative AI in military cyber defense. However, there is still much to do regarding research on varying scalability in real-world implementation. The advancements in performance documented via our case studies will need to be verified in larger-sized, more complex military, pubic, and private networks, while leveraging interconnected defense systems for comprehensive protection. In addition to the scalability issue, generative AI systems are expensive as they require considerable computing power [13,14]. Thus, future research should aim to produce an efficient generative AI model usable across military systems whilst minimizing any development of new potential vulnerabilities. As other advanced technologies are developed, military generative AI systems may also be able to leverage quantum computing or blockchain for sustainability and security.

CONCLUSION

In conclusion, the use of generative AI in military cybersecurity has considerable benefits compared to traditional methods, particularly in enhancing detection performance, response time, and adaptability. As illustrated, the benefits of an AI-enhanced system improved the accuracy of malware detection by 15%, from 80% to 95%, and a 15% increase in phishing email detection, from 78% to 93%. The ability to react quickly to a new threat was also key, as response time was reduced by 60%, from 5 minutes to 2 minutes, which is essential in military situations where responding quickly will minimize impact. Additionally, the AI systems showed the ability to reduce false favorable rates from 10% to 4% (which is excellent) and lower false negative rates from 12% to 5% (which is also that employed the AI system greatly based on its ability to identify what a real threat looks like and its=apability to identify a real threat.

The ethical and adversarial threats posed by generative AI systems in military cybersecurity present promising developments around autonomous, weaponized AI, as well as adversarial manipulation that may be full of opportunities and risks, requiring careful navigation and resolution by strong deliberative discourses, and any related protocols will require continued attention to accountabilities moving forward. The future of military cybersecurity may be best leveraged when taking greater iterations on generative AI systems and ensuring that defense overall frameworks are flexible and capable of rapid adaptations, aligning with current technologies, policy, and ethical standards. While research and technological issues are critical considerations, the military could achieve a significant and effective role using generative AI, where security, fairness, and accountability are enabled as the norms of operation.

REFRENCES

[1] Islam, M. R. (2024). Generative AI, Cybersecurity, and Ethics. John Wiley & Sons. https://books.google.com/books?hl=en&lr=&id=p8IzEQAAQBAJ&oi=fnd&pg=PP26&dq=Ethical+and+Adversarial+Risks+of+Generative+AI+in+Military+Cyber+Tools&ots=d0SQoSYrTM&sig=iIruXGWfpRReNnI-vkVF2dfOwgw

[2] Shafik, W. (2025). Generative Adversarial Networks: Security, Privacy, and Ethical Considerations. In Generative Artificial Intelligence (AI) Approaches for Industrial Applications (pp. 93-117). Cham: Springer Nature Switzerland. https://link.springer.com/chapter/10.1007/978-3-031-76710-4_5

[3] Krishnamurthy, O. (2023). Enhancing Cyber Security Through Generative AI. International Journal of Universal Science and Engineering, 9(1), 35-50. https://ijuse.org/admin1/upload/06%20Oku%20Krishnamurthy%2001155.pdf

[4] Singer, T. (2024). Visual Generative AI in Warfare and Terrorism: Risk Mitigation through Technical Requirements and Regulatory Insights (Doctoral dissertation, Technische Universität Wien). https://repositum.tuwien.at/handle/20.500.12708/205245

[5] Vadisetty, R., & Polamarasetti, A. (2024, November). Generative AI for Cyber Threat Simulation and Defense. In 2024 12th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 272-279). IEEE. https://ieeexplore.ieee.org/abstract/document/10843938/

Research Article

[6] Basrur, A. Assessing the Military Applications of Generative AI. Future Warfare and Critical Technologies: Evolving Tactics and Strategies, 84. https://swfound.org/media/207797/future-warfare-and-critical-technologies_feb-2024.pdf#page=84
[7] Oniani, D., Hilsman, J., Peng, Y., Poropatich, R. K., Pamplin, J. C., Legault, G. L., & Wang, Y. (2023). Adopting and expanding ethical principles for generative artificial intelligence from the military to healthcare. NPJ Digital Medicine, 6(1), 225. https://www.nature.com/articles/s41746-023-00965-x
[8] Pauwels, E. (2024). Preparing for next-generation information warfare with generative AI (No. 310). CIGI Papers. https://www.econstor.eu/handle/10419/311791
[9] Metta, S., Chang, I., Parker, J., Roman, M. P., & Ehuan, A. F. (2024). Generative AI in cybersecurity. arXiv preprint arXiv:2405.01674. https://arxiv.org/abs/2405.01674
[10] Andreoni, M., Lunardi, W. T., Lawton, G., & Thakkar, S. (2024). Enhancing autonomous system security and resilience with generative AI: A comprehensive survey. IEEE Access. https://ieeexplore.ieee.org/abstract/document/10623653/
[11] Khan, A., Jhanjhi, N. Z., Omar, H. A. H. B. H., Hamid, D. H. H., & Abdulhabeb, G. A. (2025). Future Trends in Generative AI for Cyber Defense: Preparing for the Next Wave of Threats. In Vulnerabilities Assessment and Risk Management in Cyber Security (pp. 135-168). IGI Global Scientific Publishing. https://www.igi-global.com/chapter/future-trends-in-generative-ai-for-cyber-defense/374395
[12] Ankalaki, Shilpa, A. Aparna Rajesh, M. Pallavi, Geetabai S. Hukkeri, Tony Jan, and Ganesh R. Naik. "Cyber Attack Prediction: From Traditional Machine Learning to Generative Artificial Intelligence." IEEE Access (2025). https://ieeexplore.ieee.org/abstract/document/10909100/
[13]Meier, R. (2025). Threats and Opportunities in AI-generated Images for Armed Forces. arXiv preprint arXiv:2503.24095. https://arxiv.org/abs/2503.24095
[14] Mylrea, M. (2025). The generative AI weapon of mass destruction: Evolving disinformation threats, vulnerabilities, and mitigation frameworks. In Interdependent Human-Machine Teams (pp. 315-347). Academic Press. https://www.sciencedirect.com/science/article/pii/B9780443292460000079