International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol.7 No.22 March –April- 2026

Protecting Computer Systems and Data Using Artificial Intelligence

Amit Venunath Dange
Dr. D. Y. Patil Arts, Commerce and Science College, Akurdi, Pune – 44
Corresponding Author – Amit Venunath Dange
DOI - 10.5281/zenodo.19345262

Abstract:

The swift growth of digital systems and the increasing amount of data have heightened cybersecurity challenges, making traditional security measures insufficient. Artificial Intelligence (AI) offers a promising solution by improving threat detection, safeguarding data, and automating defensive actions. This paper investigates the role of AI in protecting computer systems and sensitive information, examining existing techniques and potential future developments. We propose a multi-layered AI-driven security framework, explore machine learning (ML) and deep learning (DL) methodologies, and assess privacy-preserving strategies such as federated learning and encryption. Furthermore, we tackle challenges such as adversarial attacks, data governance, and model interpretability. Our conceptual framework illustrates that AI-enhanced security can significantly bolster system resilience, mitigate risks, and facilitate proactive, adaptive defense strategies.

Introduction:

The occurrence and complexity of cyber-attacks and data breaches are escalating at an extraordinary pace, rendering the protection of computer systems and sensitive information critically essential. Traditional security measures, including firewalls, antivirus programs, and signature-based intrusion detection systems, frequently fall short in addressing intricate threats like zero-day vulnerabilities, polymorphic malware, and insider threats.

AI, especially ML and DL, provides the capability to analyze extensive datasets, identify anomalies, and react to threats instantaneously. AI-powered security solutions can predict attacks, adjust to changing patterns, and alleviate the operational load on human security personnel.

This research examines the integration of AI into cybersecurity solutions. We introduce a multi-layered architectural framework, review contemporary AI techniques, outline assessment methods, and emphasize the key challenges and future research avenues in AI driven security.

Literature Review:

1. AI for Cyber Threat Detection:

Artificial Intelligence (AI) has been widely utilized in intrusion detection systems (IDS). Research shows that Machine Learning (ML) and Deep Learning (DL) models are more adept at identifying novel and polymorphic attacks compared to traditional methods. Reviews published in prominent journals, including the Journal of Big Data, outline over 60 AI-based detection strategies—encompassing ML, DL, and metaheuristic techniques—and emphasize their effectiveness in recognizing a variety of threats. Additional analyses in publications such as Measurement: Sensors examine the benefits of AI-driven IDS in contrast to standard signature-based systems.

2. AI for Data System Security:

In addition to threat detection, AI plays a crucial role in protecting data systems. Within the

realm of database security, ML algorithms are utilized for anomaly detection, predictive analytics, and automated threat response, thereby ensuring the integrity, confidentiality, and availability of data. Studies featured in the International Journal of Computational and Experimental Science & Engineering highlight AI's dual function in bolstering data security while simultaneously introducing new risk elements.

3. Privacy-Preserving AI Techniques:

Ensuring data privacy in AI applications is of utmost importance. Approaches such as federated learning, differential privacy, homomorphic encryption, and Secure Multi-Party Computation (SMPC) enable AI models to learn from data without revealing sensitive information. These techniques guarantee that user privacy is maintained throughout the processes of model training and inference.

4. Risks and Vulnerabilities of AI:

AI systems are not immune to threats, facing vulnerabilities such as poisoning attacks, inference attacks, and adversarial inputs. Furthermore, generative models like Generative Adversarial Networks (GANs) can be utilized both offensively (to create attacks) and defensively within cybersecurity contexts.

Proposed Methodology:

Multi-Layer AI Security Framework:

We present a multi-tiered AI security architecture:

Data Collection Layer:

- Collect logs from applications, network traffic, databases, and user interactions.
- Implement logging mechanisms that prioritize privacy.

Preprocessing Layer:

- Clean, normalize, encode, and transform the data.
- Utilize feature extraction and dimensionality reduction techniques (e.g., PCA).

Detection & Protection Layer:

- Anomaly Detection: Employ autoencoders or clustering methods to identify deviations from standard behavior.
- Classification Models: Use supervised machine learning models (Random Forest, SVM) to recognize known threat patterns.
- Sequential Models: Apply LSTM/RNN models for the temporal analysis of logs and traffic.
- Behavior Profiling: Model typical user behavior to identify insider threats. 4. Privacy-Preserving Layer
- Utilize federated learning to keep sensitive data local.
- Implement differential privacy for model updates.
- Employ homomorphic encryption or secure multi-party computation (SMPC) when handling encrypted data.

Response & Mitigation Layer:

- Automated responses to isolate compromised systems, block suspicious IP addresses, or revoke sessions.
- Decisions driven by policy based on the severity and context of threats.

Explainability & Monitoring Layer:

- Incorporate Explainable AI (XAI) techniques such as SHAP and LIME to enhance transparency.
- Establish continuous monitoring and feedback loops to improve models using insights from analysts.

Evaluation Strategy:

1. Datasets:

- Public IDS datasets (such as CICIDS2017 and NSL-KDD) utilized for evaluating threat detection.
- Synthetic or actual logs employed for

modeling user behavior.
- Simulations of federated learning to assess privacy-preserving strategies

2. Metrics:
- Detection Metrics: Accuracy, Precision, Recall, F1-Score, ROC-AUC
- Privacy Metrics: Privacy loss, communication overhead
- Performance Metrics: Detection latency, computational cost, model training time
- Explainability: Evaluation of model transparency by analysts

3. Experiments
- Contrast traditional security tools with the proposed AI framework.
- Examine anomaly detection models trained on normal data in relation to attack scenarios.
- Assess federated learning configurations for privacy-performance trade-offs.
- Simulate response mechanisms to evaluate speed and effectiveness.

Discussion:

1. Advantages:
- Adaptive Defense: AI models continuously learn and adjust to new threats.
- Proactive Protection: Identify anomalies and forecast attacks prior to escalation.
- Privacy Assurance: Data remains safeguarded through privacy-preserving methods.
- Automated Mitigation: Minimizes the time required to contain threats.
- Explainability: Analysts can comprehend and trust the decisions made by the model.

2. Challenges:
- Adversarial Threats: AI models are vulnerable to poisoning and adversarial attacks.
- Data Governance: The challenge of balancing data collection with privacy compliance is intricate.
- Resource Requirements: Federated learning and encryption demand significant computational resources.
- Interpretability: Despite XAI techniques, deep learning models can remain opaque.
- Scalability: Implementing AI solutions across extensive systems presents challenges.

Future Research Directions:

1. Develop resilient AI models that can withstand adversarial attacks.
2. Investigate hybrid systems that merge AI with blockchain technology for secure logging.
3. Design lightweight edge AI models tailored for IoT devices.
4. Improve explainability and accountability to ensure trustworthy AI.
5. Examine compliance and governance frameworks aimed at AI security.
6. Explore the safe integration of generative AI models within cybersecurity systems.

Conclusion:

AI possesses transformative capabilities in securing digital infrastructures and protecting sensitive data. By incorporating machine learning, deep learning, and privacy-preserving methodologies, AI-driven systems can identify, respond to, and mitigate threats in real time while ensuring privacy is maintained. Despite challenges such as adversarial threats, computational expenses, and governance issues, continuous research into robust, transparent, and privacy-conscious AI solutions will be essential for the future of cybersecurity.

References:

1. Mohamed, I. M., & Alosman, K. (2025). Artificial Intelligence in Security and Privacy: A Study on AI's Role in Cybersecurity and Data Protection.

International Journal of Education and Management Engineering. mecs-press.org

2. Sangwan, R. S., Badr, Y., & Srinivasan, S. M. (2023). Cybersecurity for AI Systems: A Survey. Journal of Cybersecurity & Privacy. MDPI
3. Yellepeddi, S. M., Ravi, C. S., Vangoor, V. K. R., & Chitta, S. (2024). AI-Powered Intrusion Detection Systems: Real-World Performance Analysis. Journal of AI-Assisted Scientific Discovery. scienceacadpress.com
4. Susarla, V. S. T., & Kumbham, S. C. R. (2025). A Survey Paper on AI Based Intrusion Detection System. ResearchGate
5. Karri, N., & Jangam, S. K. (2025). Role of AI in Database Security. International Journal of Artificial Intelligence, Data Science, and Machine Learning. IJADML
6. Hasan, S. N., Kaur, H., Mohonta, S. C., et al. (2025). The Influence of Artificial Intelligence on Data System Security. International Journal of Computational and Experimental Science. and Engineering. Ijcesen
7. Prajapati, S. B. (2025). Strategies for Protecting User Data While Maintaining Functionality: Data Privacy and Security in AI. World Journal of Advanced Research and Reviews. World Affairs Review
8. Kakarala, M. K., & Rongali, S. K. (2025). Data Privacy and Security in AI. World Journal of Advanced Research and Reviews. World Affairs Review
9. Arifin, M. M., Ahmed, M. S., Ghosh, T. K., et al. (2024). A Survey on the Use of Generative Adversarial Networks in Cybersecurity. arXiv
10. Journal of Big Data. (2024). Enhancing Cybersecurity: A Thorough Review of AI-Driven Detection Techniques. SpringerOpen