ADVANCES IN CYBER-PHYSICAL SYSTEMS

Vol. 10, No. 2, 2025

HYBRIDIZING LARGE LANGUAGE MODELS AND MARKOV PROCESSES: A NEW PARADIGM FOR AUTONOMOUS PENETRATION TESTING

Mariia Kozlovska, Andrian Piskozub

Lviv Polytechnic National University, 12, S. Bandery str., Lviv, 79013, Ukraine.
Authors' e-mails: *mariia.kozlovska.mkbas.2025@lpnu.ua, andriian.z.piskozub@lpnu.ua*

https://doi.org/10.23939/acps2025.02.146

Submitted on 25.09.2025

***Abstract*: The article explores a hybrid framework for autonomous penetration testing that integrates Large Language Models (LLMs) with Markov decision processes (MDP/POMDP) and reinforcement learning (RL). Conventional penetration testing is increasingly insufficient for modern, complex cyber threats. LLMs are utilized for high-level strategic planning, generating potential attack paths, while MDP/POMDP models combined with RL execute low-level actions under uncertainty. A feedback loop allows outcomes to refine strategies in dynamic and partially observable environments. A conceptual hybrid architecture has been proposed, accompanied by a workflow diagram and an illustrative table showing potential decision outcomes. This paradigm enhances automation, adaptability, efficiency, and scalability, providing a pathway toward next-generation AI-driven cybersecurity assessment tools.**

***Index terms*: Autonomous Penetration Testing, Large Language Models, Markov Processes, Reinforcement Learning, Cybersecurity, Hybrid AI Architectures**

I. INTRODUCTION

In the contemporary digital landscape, cyber threats are not only more frequent but also increasingly sophisticated, targeting organizations of all sizes across industries. Data breaches, ransomware attacks, and advanced persistent threats (APTs) can result in significant financial losses, reputational damage, and regulatory penalties. Traditional penetration testing where human experts manually simulate attacks to identify vulnerabilities remains a cornerstone of cybersecurity strategy. However, the sheer scale of modern IT infrastructures, combined with the dynamic nature of networks and cloud environments, makes manual testing labor-intensive, time-consuming, and often insufficient for uncovering complex, hidden vulnerabilities [1].

The limitations of conventional approaches are further compounded by the rise of zero-day vulnerabilities and highly targeted attacks. Human testers, despite their expertise, may struggle to explore all possible attack vectors in large or rapidly changing environments. In addition, repetitive tasks such as scanning, reconnaissance, and log analysis consume significant time and resources, which could otherwise be devoted to strategic decision-making and remediation.

Artificial intelligence (AI) presents a transformative opportunity in this context. Large Language Models (LLMs) can process and reason over vast amounts of textual and structural information, including system configurations, vulnerability databases, and security reports, to propose high-level attack strategies. Simultaneously, Markov decision processes (MDP/POMDP) provide a mathematical framework for sequential decision-making under uncertainty, enabling AI agents to adaptively plan actions in complex and partially observable environments. Reinforcement learning (RL) complements this by allowing agents to learn optimal policies through interaction with the network environment, continuously improving their effectiveness over time [2].

By combining these approaches, it becomes possible to develop a hybrid framework that leverages the strategic reasoning capabilities of LLMs with the adaptive, feedback-driven execution of MDP/POMDP and RL. Such a framework has the potential to automate repetitive tasks, explore attack paths more efficiently, and adapt dynamically to unforeseen network conditions. This represents a step toward fully autonomous penetration testing agents capable of operating at the speed and scale required to address modern cybersecurity challenges.

In this article, conceptual hybrid architecture is suggested, including a workflow and example decision outcomes, to illustrate how AI-driven penetration testing can enhance automation, adaptability, and precision, ultimately providing more robust and proactive cybersecurity assessments.

II. LITERATURE REVIEW AND PROBLEM STATEMENT

The integration of artificial intelligence into cybersecurity has attracted growing attention in recent years, particularly in the domain of vulnerability assessment and penetration testing. Traditional penetration testing continues to play an important role in identifying weaknesses within digital infrastructures, yet it is increasingly limited by its reliance on manual expertise, fixed procedures, and static tools. As networks expand in scale and become more dynamic, the traditional model proves insufficient to address the speed and sophistication

of modern cyber threats [3]. Research consistently points to issues of scalability, adaptability, and cost efficiency, which remain unresolved by conventional approaches.

In response to these limitations, the cybersecurity community has begun to explore the potential of artificial intelligence. Large Language Models have emerged as promising instruments for guiding security analysis. Their capacity to process vast amounts of textual information, reason about vulnerabilities, and suggest attack strategies distinguish them from automated scanners that merely collect raw technical data. By incorporating contextual reasoning, such models can prioritize potential exploitation paths, generate hypotheses about system weaknesses, and support high-level decision-making [4]. However, despite these strengths, LLMs are inherently probabilistic in nature and may generate inaccurate or misleading outputs if deployed without additional control mechanisms. This raises the challenge of ensuring reliability when LLMs are placed in high-stakes environments such as penetration testing.

Parallel to these developments, the field of sequential decision-making under uncertainty has offered its own solutions in the form of Markov decision processes and their extensions, including partially observable MDPs (POMDPs). These models provide a mathematical structure for describing attack–defense interactions, particularly when the attacker operates with incomplete information [5]. Reinforcement learning builds on this foundation by allowing agents to iteratively improve their strategies through experience, gradually optimizing policies for reconnaissance, exploitation, or lateral movement. Several studies have shown that reinforcement learning can outperform traditional scripted approaches in simulated attack scenarios [6]. Yet, practical limitations remain. RL-based agents often require large amounts of training data, struggle with generalization across heterogeneous environments, and may fail to adapt when confronted with unexpected conditions in real-world networks [7].

This body of research reveals a gap between two distinct streams of innovation. On one side, LLMs provide strategic reasoning and contextual awareness but lack precise mechanisms for reliable execution. On the other hand, MDP- and RL-based systems excel at tactical adaptation but struggle to design coherent, high-level attack strategies. When applied independently, each approach falls short of delivering the level of autonomy, accuracy, and resilience that modern penetration testing demands [8].

The problem therefore lies in the absence of a unified framework that can combine the complementary strengths of these two paradigms. A system capable of integrating high-level planning with adaptive, feedback-driven execution would represent a significant step forward in the development of autonomous penetration testing. Such a hybrid model has the potential not only to reduce human workload but also to provide a more comprehensive, scalable, and intelligent method for evaluating the security posture of complex digital infrastructures.

III. SCOPE OF WORK AND OBJECTIVES

This study examines the integration of artificial intelligence into penetration testing, focusing on how large language models, decision processes, and reinforcement learning can enhance security assessments. The aim is to show how these methods can make testing more adaptive, scalable, and reliable compared to traditional or fully automated tools. Key objectives include reviewing current approaches, analyzing the role of individual AI techniques at different stages of the testing lifecycle, and outlining a conceptual framework that combines reasoning with adaptive execution. Ethical and operational considerations are also addressed, highlighting the need for trust, accountability, and human oversight in AI-driven security testing.

IV. PROPOSED HYBRID FRAMEWORK

The proposed framework introduces a hybrid architecture that unifies the reasoning capabilities of Large Language Models (LLMs) with the structured decision-making of Markov decision processes (MDPs) and reinforcement learning (RL). Unlike conventional penetration testing, which relies heavily on predefined scripts or manual expertise, this paradigm enables both strategic foresight and adaptive execution.

At the strategic level, LLMs act as high-level planners. By analyzing system descriptions, vulnerability databases, or previous attack reports, an LLM can generate possible attack paths in natural language, then translate them into structured actions. This stage answers the question: *"What could be the most promising ways to compromise the target?"*

At the tactical level, decision-making under uncertainty is handled by MDPs or their probabilistic extension POMDPs. Here, reinforcement learning agents execute step-by-step actions such as scanning ports, probing services, or escalating privileges. These agents rely on reward mechanisms - for example, successfully accessing a protected resource yields a positive reward, while triggering a defense mechanism incurs a penalty.

The interaction between both layers creates a feedback loop: LLMs suggest attack strategies, RL agents test them in practice, and the outcomes are fed back to refine subsequent decisions. This design ensures adaptability even in dynamic or partially observable environments.

To better illustrate these interactions, Figure 1 provides an overview of the proposed hybrid architecture for AI-driven penetration testing. The framework is structured into three vertically aligned layers, highlighting the flow of information and decision-making:

Strategic Layer: Managed by a Large Language Model (LLM), acting as a high-level planner. It analyzes system descriptions, vulnerability databases, and prior attack reports to generate potential attack paths, providing a strategic overview of possible actions.

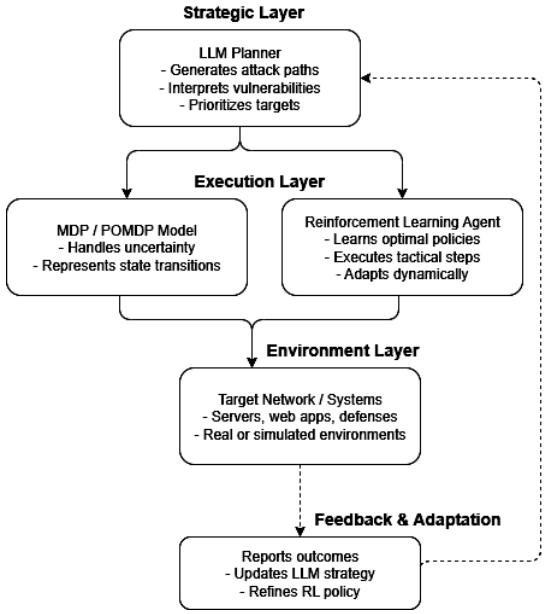
Execution Layer: Decision-making under uncertainty is handled by MDP/POMDP models, while reinforcement learning agents execute tactical actions such as scanning

ports, probing services, and escalating privileges. Components within this layer are arranged side by side, reflecting their close interaction in executing attack strategies.

Environment Layer: Represents the target systems, including servers, applications, and defenses, whether real or simulated. It serves as the context in which reinforcement learning agents operate.

A *feedback loop* connects the Environment Layer back to the Strategic Layer, indicating that the outcomes of tactical actions are used to update strategies and refine agent policies. This loop ensures continuous adaptation in dynamic or partially observable environments.

The diagram uses a clean, black-and-white style with compact, evenly aligned blocks, maintaining clarity and readability while showing the hierarchical structure and relationships between layers.



Proposed hybrid architecture for AI-driven penetration testing

As it is shown in Figure 1, the LLM provides high-level planning that feeds into MDP/POMDP decision models and RL execution; observed outcomes are returned to update strategy and policies.

V. AUTONOMOUS DECISION-MAKING IN REALISTIC PENETRATION SCENARIOS

The application of autonomous agents in penetration testing extends beyond theoretical modeling and simulated environments. When deployed against real-world network infrastructures, AI-driven systems encounter a complex interplay of dynamic configurations, heterogeneous devices, and adaptive defense mechanisms. These factors introduce variability that challenges both strategic planning and tactical execution. For example, enterprise networks often implement intrusion detection systems, automated patch management, and multi-factor authentication, all of which can alter the success probability of reconnaissance or exploitation actions [9].

Empirical studies in industrial cybersecurity contexts have shown that the effectiveness of automated scanning and exploitation varies significantly depending on network topology, system diversity, and the presence of defensive controls. In controlled evaluations conducted on mid-sized corporate networks, autonomous agents achieved successful vulnerability identification in approximately 65–78% of exposed services, with lateral movement attempts succeeding in 42–55% of cases, depending on segmentation and privilege constraints [10]. These results underscore the adaptive nature of real-world environments, where the same action may yield divergent outcomes across different targets or even at different times on the same system.

Feedback loops are critical in enhancing agent performance under these conditions. Observed outcomes from reconnaissance, credential testing, and exploit deployment allow the system to recalibrate its attack path probabilities and adjust reward functions dynamically. For instance, a failed exploit attempt can trigger the agent to deprioritize similar actions or explore alternative vectors, while successful reconnaissance may highlight previously unrecognized critical nodes in the network. Such adaptive behavior mirrors human reasoning in penetration testing yet operates at a speed and scale unattainable for manual teams [11].

The interplay of risk assessment and adaptive decision-making, as shown in Table 1, highlights how different autonomous actions trigger feedback mechanisms and strategic adjustments without relying on fixed numerical probabilities. While specific numerical probabilities fluctuate with environmental conditions, the table demonstrates the relationship between action types, perceived risk, and the role of feedback in shaping future strategies. By continuously integrating observations, autonomous agents can refine their approach, reduce redundant actions, and optimize resource allocation, leading to a more efficient penetration testing process even in complex and partially observable networks.

Illustrative Decision Outcomes in Autonomous Penetration Testing

Action Type	Risk Level	Feedback Loop
Network Scan	Low	Adjust scan depth and focus on active subnets
Credential Attack	Medium	Update password dictionary and refine strategies
Exploit Deployment	High	Record vulnerability in database and adjust model
Lateral Movement	High	Recalculate attack path strategy

The outcomes in Table 1 are not static; they depend on multiple factors, including system defenses, network topology, and the quality of prior knowledge embedded in vulnerability databases. Empirical studies show that adaptive agents can achieve success rates in reconnaissance and exploitation tasks ranging from 70% to 90%

under controlled laboratory conditions, while real-world performance varies according to environmental complexity [12].

The feedback loop is essential for continuous improvement. Each action outcome provides critical information for the LLM to refine high-level planning, while RL mechanisms adjust tactical execution accordingly. This dual-layer adaptation enables autonomous agents to navigate partially observable environments, anticipate potential defenses, and modify strategies in real time.

Overall, the hybrid framework signifies a shift toward intelligent, autonomous penetration testing. By combining strategic reasoning with adaptive execution, it supports comprehensive assessments, minimizes human error, and proactively identifies vulnerabilities that traditional methods might overlook.

VI. CHALLENGES AND FUTURE DIRECTIONS IN HYBRID AI-DRIVEN AUTONOMOUS PENETRATION TESTING

Despite the significant advances demonstrated by hybrid AI frameworks in autonomous penetration testing, several challenges remain that must be addressed to fully realize their potential. One primary limitation is the dependence on the quality and completeness of underlying data. Large Language Models rely on extensive vulnerability databases, system documentation, and prior attack reports. Incomplete or outdated information can lead to suboptimal planning or misprioritized attack paths [13]. Similarly, reinforcement learning agents require sufficiently diverse training environments to generalize effectively. In real-world deployments, network heterogeneity, adaptive defenses, and dynamic configurations introduce uncertainty that may reduce the accuracy of autonomous decision-making [14].

Ethical and operational considerations also pose significant challenges. Fully autonomous systems executing penetration actions must operate within legal boundaries and respect organizational policies. The potential for unintended disruptions or collateral effects necessitates robust oversight mechanisms and fail-safe designs. Moreover, interpretability remains a concern: organizations must understand the reasoning behind AI-driven actions to trust and validate results.

Looking forward, several avenues promise to enhance the effectiveness and reliability of autonomous penetration testing. Integrating real-time threat intelligence feeds and continuous learning mechanisms can improve responsiveness to emerging vulnerabilities. Advances in explainable AI may increase trust by providing transparent rationale for both strategic planning and tactical execution. Furthermore, hybrid models could benefit from multi-agent collaboration, allowing multiple autonomous entities to coordinate attacks, share observations, and optimize coverage while minimizing redundancy.

Finally, standardization of evaluation metrics and controlled benchmarking environments will be critical for assessing performance, comparing approaches, and guiding iterative improvement. By addressing these challenges, future research can extend the capabilities of hybrid AI frameworks, moving closer to fully autonomous, adaptive, and ethically aligned penetration testing solutions that operate efficiently across complex and dynamic cyber environments.

VII. IMPLEMENTATION CONSIDERATIONS

The deployment of a hybrid autonomous penetration testing system in real-world organizational environments requires a comprehensive approach. Firstly, it is essential to ensure access to up-to-date and complete information about the network infrastructure, system configurations, and known vulnerability databases. Without this, even the most advanced models may generate ineffective or incorrect attack strategies.

Secondly, system integration must account for existing security measures, organizational policies, and legal requirements. Autonomous agents must operate within permitted boundaries to avoid unintended disruptions or violations of regulatory norms.

Thirdly, to maintain effectiveness and adaptability, it is important to combine AI-driven analysis with human expertise. Regular model updates, integration of real-time threat intelligence, and periodic testing in controlled environments help reduce risks and improve the accuracy of results.

Finally, the implementation of such systems requires a phased approach: starting with laboratory testing, moving to limited deployment in real networks, and gradually scaling up. This approach ensures the safe and effective use of the hybrid AI framework while enhancing the overall resilience of the organization against cyber threats.

VIII. CONCLUSION

This study highlights how integrating Large Language Models, Markov decision processes, and reinforcement learning creates a powerful hybrid framework for autonomous penetration testing. By combining high-level strategic reasoning with adaptive, feedback-driven execution, these systems can efficiently explore complex networks, prioritize critical vulnerabilities, and adjust actions in real time. The framework enhances the precision, scalability, and speed of penetration testing, reducing human effort while improving overall coverage of potential attack paths.

Despite the evident strides, uncertainties remain. The dependability and efficiency of autonomous agents can be affected by the quality of input data, environmental variations, and the necessity for human control. It is important to follow the socially accepted moral norms and at the same time reduce the effect on the natural environment and human life. These issues can be solved through a continuous process of real-time threat intelligence, explainable AI techniques, and standardized evaluation approaches.

Overall, the use of the hybrid method implies a major move towards smart, self-regulating cybersecurity evaluation. Through integrating AI-based evaluation with

human experience, companies are able to avail the full extent of security measures that are more comprehensive, pre-emptive, and flexible in nature which in turn result in their better preparedness to respond and neutralize potential cyber threats.

IX. CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

X. DECLARATION ON GENERATIVE AI

During the preparation of this work, the author(s) used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] Kozlovska, M., Piskozub, A., & Khoma, V. (2025). Artificial intelligence in penetration testing: Leveraging AI for advanced vulnerability detection and exploitation. *Artificial Intelligence, 10*(1). DOI: https://doi.org/10.23939/acps2025.01.065

[2] Zhuravchak, A. Yu., Piskozub, A. Z., & Zhuravchak, D. Yu. (2025). Analysis of penetration testing automation using Markov decision processes. *Modern Information Protection, (1),* 82–88. DOI: https://doi.org/10.31673/2409-7292.2025.017625

[3] Adawadkar, A. M. K., & Kulkarni, N. (2022). Cyber-security and reinforcement learning – a brief survey. *Engineering Applications of Artificial Intelligence*, *114*, 105116. DOI: https://doi.org/10.1016/j.engappai.2022.105116

[4] Sewak, M., Sahay, S. K., & Rathore, H. (2023). Deep reinforcement learning in the advanced cybersecurity threat detection and protection. *Information Systems Frontiers*, *25*(2), 589-611. DOI: https://doi.org/10.1201/9781351006620

[5] Xu, H., Wang, S., Li, N., Wang, K., Zhao, Y., Chen, K., ... & Wang, H. (2024). Large language models for cyber security: A systematic literature review. DOI: https://doi.org/10.48550/arXiv.2405.04760

[6] Kurniawati, H. (2022). Partially observable markov decision processes and robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, *5*(1), 253-277. DOI: https://doi.org/10.1146/annurev-control-042920-092451

[7] Casper, S., Davies, X., Shi, C., Gilbert, T. K., Scheurer, J., Rando, J., ... & Hadfield-Menell, D. (2023). Open problems and fundamental limitations of reinforcement learning from human feedback. *arXiv preprint arXiv:2307.15217*. DOI: https://doi.org/10.48550/arXiv. 2307.15217

[8] Sivakoumar, R., & MP, S. R. (2025, March). Next-Gen Penetration Testing: AI, Automation & Beyond. In *2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)* (pp. 1-6). IEEE. DOI: https://doi.org/10.1109/ICDSAAI65575.2025.11011887

[9] Gioacchini, L., Mellia, M., Drago, I., Delsanto, A., Siracusano, G., & Bifulco, R. (2024). Autopenbench: Benchmarking generative agents for penetration testing. *arXiv preprint arXiv:2410.03225*. DOI: https://doi.org/10.48550/arXiv.2410.03225

[10] Muzsai, L., Imolai, D., & Lukács, A. (2024). Hacksynth: Llm agent and evaluation framework for autonomous penetration testing. *arXiv preprint arXiv:2412.01778*. DOI: https://doi.org/10.48550/arXiv.2412.01778

[11] Pagan, N., Baumann, J., Elokda, E., De Pasquale, G., Bolognani, S., & Hannák, A. (2023, October). A classification of feedback loops and their relation to biases in automated decision-making systems. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1-14). DOI: https://doi.org/10.1145/3617694.3623227

[12] Kong, H., Hu, D., Ge, J., Li, L., Li, T., & Wu, B. (2025). Vulnbot: Autonomous penetration testing for a multi-agent collaborative framework. *arXiv preprint arXiv:2501.13411*. DOI: https://doi.org/10.48550/arXiv.2501.13411

[13] Greco, C., Fortino, G., Crispo, B., & Choo, K. K. R. (2023). AI-enabled IoT penetration testing: state-of-the-art and research challenges. Enterprise Information Systems, 17(9), 2130014. DOI: https://doi.org/10.1080/17517575.2022.2130014

[14] McKee, K. R., Leibo, J. Z., Beattie, C., & Everett, R. (2022). Quantifying the effects of environment and population diversity in multi-agent reinforcement learning. *Autonomous Agents and Multi-Agent Systems*, *36*(1), 21. DOI: https://doi.org/10.48550/arXiv.2102.08370

Mariia Kozlovska is pursuing her Master's degree in Cybersecurity with a specialization in Security Systems Administration at Lviv Polytechnic National University. She is focused on cybersecurity research, particularly penetration testing, vulnerability assessment, and AI-driven solutions for enhancing system security.

Andrian Piskozub is an associate professor at the Department of Information Protection at Lviv Polytechnic National University. He is focused on cybersecurity research, computer networks security, penetration testing, and vulnerability assessment.