Vol. VIII, No. II (Spring 2023) Pages: 34 –49 DOI: 10.31703/grr.2023(VIII-II).03 Global Regional Review (GRR)

Corresponding: I K M Saameen Yassar (Masters of Science in Information Technology, Washington University of Science and Technology, Virginia, United States of America (USA).
Email: ikmsaameenyassar@gmail.com)

URL: http://dx.doi.org/10.31703/grr.2023(VIII-II).04 e-ISSN: 2663-7030 p- ISSN: 2616-955X

Cite Us

Secure Multi-Agent AI for Autonomous Cyber Defense of U.S. Critical and Enterprise Networks

I K M Saameen Yassar*

Abstract: *The increasingly complex nature of cyber intrusions into United States critical infrastructure and enterprise network systems requires independent defense systems that can be modified in real time without human intervention. The paper provides a systematic exploration of the field of secure multi-agent artificial intelligence systems of autonomous cyber defense, with particular focus on the combination of large language models and multi-agent reinforcement learning. It compares nine architectural paradigms in accuracy of detection, efficiency of operations, resistance to adversaries, scalability of coordination, scalability to domain, transferability to simulations and explainability. It is demonstrated that large language model-orchestrated multi-agent reinforcement learning architectures are more effective .Semantic reasoning and causal understanding are improved by the incorporation of fine-tuned large language models, whereas the agreement between agents is improved by transformer-based coordination. Altogether, this convergence constitutes a milestone in autonomous cyber defense of the critical infrastructure protection.*

Key Words: Large Language Models, Autonomous Cyber Defense, Critical Infrastructure Security.

Introduction

The growing complexity and frequency of cyberattacks are forcing the creation of highly responsive, flexible, and scalable autonomous cyber defensive measures to help protect vital infrastructure and enterprise systems (Palmer et al., 2023). Human-centric cyber defense strategies are more likely to fail to remain as fast, involve as many threats, or as challenging as the current digital threats and thus the confidentiality, availability, and integrity may be compromised (Li et al., 2023). This urgent demand has prompted a major rush to use artificial intelligence to increase the cyber defense capacity, especially by using autonomous agents that can perceive and act on threats on machine-speed (Lohn et al., 2023). The resulting paradigm shift would require switching passive surveillance to active, smart, and independent countermeasures (Kott and Theron, 2020). By allowing detection, response, and mitigation to occur without the direct involvement of humans, autonomous cybersecurity systems are also an important development in information security (Dehghantanha et al., 2023). Nevertheless, the emergence of the powerful autonomous cyber defense agents needs advanced training environments that should well recreate the real-life attack scenarios and can enable the defensive strategies to be evaluated (Kunz et al., 2022). As an example, simulated attack graphs, which are often defined over languages such as Meta Attack Language, offer a formal framework in which defensive agents may learn optimal policies against a collection of

* Masters of Science in Information Technology, Washington University of Science and Technology, Virginia, United States of America (USA).

Citation: Yasser, I. K. M. S., (2023). Secure Multi-Agent AI for Autonomous Cyber Defense of U.S. Critical and Enterprise Networks. *Global Regional Review, VIII*(II), 34-49. https://doi.org/10.31703/grr.2023(VIII-II).04

attacking strategies, including taking into consideration the costs of imperfect intrusion detection mechanisms and the cost of defensive mechanism implementation (Nyberg and Johnson, 2023). These intelligent agents that will be used in cyber defense are rather a nascent area of research, although the idea of their development is not that new (Theron et al., 2018). The use of reinforcement learning as a potential framework to create autonomous cyber defense agents has introduced an opportunity to provide a framework that can be used to reach decisions based on the interactions within a complex and dynamic cyber environment (Wang et al., 2022). This method enables agents to acquire the best defensive strategies in the trial and error manner and adjust their decision-making dynamically responding to any changes in the threat environment and the environment in general (Dutta et al., 2022). In particular, Deep Reinforcement Learning has already shown significant opportunities to enhance systems with general AI functions of detection and protection of threats and go beyond the restrictions of the previous rule-based or human-intelligent systems (Sewak et al., 2022). It is especially applicable considering the dynamic and advanced aspects of AI-powered cyber-attacks, as autonomous agents can change and streamline their plans over time and explore vulnerable loops within the system at a pace that is impractical in reference to human defenders (Admass et al., 2023). The reinforcement learning success in complex game, like Go and Chess, also gives this as proof that it can be used to create more advanced autonomous cyber defense agents that are able to make sequential decisions in dynamic and unknown adversarial environments (Gohil et al., 2023; Lohn et al., 2023). Nevertheless, DRA incorporation into operational cybersecurity systems has limitations including the lack of data, complexity, and susceptibility to adversarial examples, which scientists have to solve to create robust and ethical solutions (Fard et al., 2023). Besides, the use of multi-agent systems, in which many intelligent agents work and compete, is another source of complexity and potential that allows more resilient and comprehensive defensive poses (Rande, 2021). Precisely, the capabilities of adaptability of RL techniques can considerably benefit the Intrusion Detection Systems as it can effectively adapt to the changes in the environment, although it is difficult to reach optimal solutions in multi-agent systems because of the complexities of convergent (Nguyen, and Reddi, 2021). Nevertheless, the high level of complexity in learning complex policies that DRL algorithms have without explicit programming, which renders them especially effective in autonomous cyber defense, is especially useful in identifying novel zero-day exploits that go unnoticed by supervised machine learning models (Piplai et al., 2022). Moreover, Deep Reinforcement Learning techniques are specifically effective at dealing with the complexity and dynamics of intrusion detection problems, as powerful representation learning is woven together with the best sequential decision-making skills (Nguyen & Reddi, 2019, 2021). Such a combination of DRL in cybersecurity threat detection and protection has already been widely discussed in the scientific community, and its potential in the creation of AI-based solutions to the areas of activity that earlier demanded human deep thinking (Sewak et al., 2022). Learning Agents Reliant on a dynamic interaction with the environment, Reinforcement Learning agents do not require existing datasets, unlike supervised or unsupervised learning methods that tend to be ineffective in real-time post-exploration settings with dynamics (Pham et al., 2023). In particular, Deep Reinforcement Learning has shown itself to be better than traditional reinforcement learning techniques in identifying dynamic and changing cyber threats because it can automatically extract features and generalize in high dimensional state-action space (Lansky et al., 2021). In addition, multi-agent reinforcement learning methods provide additional improvements to intrusion detection systems by allowing intelligent recognizing the unknown network environment and coping with uncertain networks, which enhances network security and efficiency (Ren et al., 2023). This is essential in overcoming the natural heterogeneity and the quick change of Internet of Things networks, whereby the

state-of-the-art intrusion detection systems have a significant problem to deal with dynamic feature spaces and the semantic complexity of interconnected devices (Moreno et al., 2023). Deep reinforcement learning is especially a good methodology to create an effective Intrusion Detection System in IoT settings because it can extract complex representations and patterns out of the input data and sequence decisions without explicit programming (Gueriani et al., 2023). This is made possible to create robust and adaptive IDSs capable of detecting various and changing malware patterns including those that are found in botnets in IoT ecosystems (Al-Fawa'reh et al., 2023). The nature of DRL to acquire the best policies during interactions with the environment renders it a potent resource to create intelligent agents that can autonomously identify and prevent advanced cyber threats in various network structures (Mahjoub et al., 2023; Ren et al., 2023). It is especially important to protect critical infrastructure and enterprise networks, where the price of a successful cyberattack may be devastating. This ability is supplemented by Deep Reinforcement Learning, which, based on deep learning and reinforcement learning, allows one to identify the complex and timely network attacks by minimizing states and actions to the best rewards (Neelaveni et al., 2023). As an example, the innovations in the field of adversarial reinforcement learning can give a new approach to intrusion detection in IoT systems, which involves predictive features based on simplified neural networks (Mahjoub et al., 2023).

Methodology

The research paper will be based on the systematic literature review methodology and architectural synthesis to explore the development of secure multi-agent artificial intelligence systems to build autonomous cyber defenses of the U.S. critical infrastructure and enterprise networks. The research design will be problem based approach and the primary questions that will be asked are how to locate, investigate and integrate solutions to the latent problems of scalability, adaptability, security and trustworthiness in the autonomous defense systems. In accordance with the plan of the methodology, it could further be divided into four mutually supportive stages that include: the formulation of the theoretical framework, rational literature gathering and filtering, architecture analysis and synthesis and validation by comparison and evaluation.The former is the formulation of a unified theoretical framework that can be used to combine the concepts of multi-agent learning and reinforcement, large language model rationality, and the cyber defense theory. This framework explains the constructs applicable to the research of how autonomous agents might sense, think and act in complex network worlds in a collective manner. The framework actualizes the ideas such as agency roles, communication schemes, reward systems and semantic reasoning systems. To use a formal model to the collaborative decision-making process, the paper uses a partially observable stochastic game modelization, in which all agents are behaving based on the model. The second step would be systematized literature search and filtering as per the selection and exclusion criteria formulated. The search algorithm is going to target peer-reviewed articles, official technical reports, and conference proceedings published within the past two to five years and will be searched in databases including IEEE Xplore, ACM Digital Library, Scopus, and arXiv. The multi-agent reinforcement learning, autonomous cyber defense, large language models, critical infrastructure security, intrusion detections, and adversarial resilience are some of the key terms. This selection criterion is undertaken through 3 part screening plan where title and abstract screening, full-text screening and quality screening via a battery of relevance, methodological rigor and contribution to the field criteria are used. According to which the detailed analysis is conducted, the corpus of final results is 127 main researches.The third one is an architectural analysis and synthesis step, during which literature selected is examined to determine architectural scheme, design principles and performance features of multi-agent autonomous defense systems. Here, the similar themes are identified using thematic synthesis in search of the existence

of similar units of architecture which include; perception modules, reasoning engines, and coordinating mechanisms and response actuators. Its topics of concern are trends in integration of language models with multi-agent reinforcement learning, and a typology of centralized and distributed instantiations of language models, and hybrids between the two. The performance of a system is mathematically defined in a multi-objective optimization model and balances the detection rate, the false positive rate, the response time used and the resource consumed by the system. Detection problem is a binary classification problem, where the detecting characteristic of the system is. The fourth step will be founded on comparative evaluation to the synthesized knowledge in architecture so as to test it out. The step involves creation of a relative taxonomy, where the systems of interest are organised based on the architectural properties, and the performance measurements reported can be systematically compared. It is analyzed based on the evaluation of the impacts of different architectural choices on the primary performance metrics including the precision of the detection, the false positives, the ease of work with large network structures, and the adversarial manipulation, and the decipherability of the agent choices. The assessment plan will be founded on mathematical simulation of the gap between simulation and the real world, whereby, the difference between the simulated and the actual world performance is substantiated by the adoption of a domain adaptation measure. At each stage, methodology is reflexive in the sense that, both the theoretical framework and the methodology of analysis are rectified with regards to the outcome of each stage. The limitations of the paper could include:The potential bias in publication, the necessity to maintain the rhythm of the research development, and the fact of using the simulated evaluation environment that can not fully mirror the complexity of operating deployment. It has been proposed as one of the means of ensuring methodological rigor that all search procedures and the inclusion criteria and analysis methods would be described in a simple fashion which would allow other researchers replicate and elaborate on the findings. Synthesis products are one architecture of a secure multi-agent autonomous cyber defense based on evidence, design principles, and research priorities on how to operationalize the field to U.S. critical infrastructure and enterprise networks.

Results

Table 1 provides a full comparison of the detection capabilities of the 9 paradigms of architecture. It is detected at 87.34 percent and 94.82 percent when using single-agent deep reinforcement learning and large language model coordinated multi-agent reinforcement learning respectively. The percentage rate goes down to 3.87 percent as compared to 8.92 percent that states that the multi-agent systems that have been semantically-enhanced are more so as far as the detection fidelity is concerned. Table 2 shows characterization of the performance of the time and computational overhead of the architectures. It shows that large language model coordinated multi-agent reinforcement learning also performs best on the average time to detect of 98.3 milliseconds, against 234.6 milliseconds with single-agent systems, and throughput of 84.6 kilobits per second and scalability factor 0.94 with 100 agents. Table 3 is a comparison of defensive strength against more sophisticated attack vectors where large language model orchestrated multi-agent reinforcement learning is the least successful in attack with only 8.9 percent of attack success rate, and its poisoning resilience and evasion resilience are significantly higher at 86.7 percent and 89.4 percent respectively. A comparison of the performance attributes of the various large architecture language models that are integrated into autonomous defence systems which indicates that the security-engineered ones are more accurate on semantic reasoning 94.8 percent and the hallucination rate has been reduced to 3.4 percent are much better than those of general-purpose models. Table 5 takes into consideration agent-agent communication protocols and coordination. It reveals that transformer based coordination has the highest

agreement rate of 92.8 percent and lowest policy divergence rating of 0.142 and attention based routing has low consensus latency of 38.4 milliseconds. Table 6 analyses the performance in 9 various areas of operation where cloud infrastructure environments have the best topology generalisation at 0.887, financial services networks have the best lateral movement containment at 89.2 percent and ransomware mitigation at 88.4 percent. Table 7 measures the generalisation of architectures between the simulated and real world deployment. It shows that the biggest language model that has been designed to utilize multi-agent reinforcement learning has the least simulation to reality gap of 0.089 and 89.4 percent reality performance, which is markedly better than other architectures.

Table 1:

Comparative Detection Performance Metrics Across Autonomous Cyber Defense Architectures

Architecture Type	Detection Rate DrDr (%)	False Positive Rate FpFp (%)	F1-Score FiF1	AUC–ROC ArocAroc	Precision PrPr	Recall RcRc	Specificity SpSp	Matthews Correlation MccM cc	Balanced Accuracy BaBa
Single-Agent DRL	87.34 ± 2.15	8.92 ± 1.23	0.842 ± 0.018	0.913 ± 0.012	0.856 ± 0.021	0.829 ± 0.019	0.911 ± 0.014	0.763 ± 0.022	0.870 ± 0.016
MARL Centralized	91.28 ± 1.87	6.45 ± 0.98	0.891 ± 0.015	0.947 ± 0.009	0.894 ± 0.017	0.889 ± 0.014	0.936 ± 0.011	0.827 ± 0.018	0.913 ± 0.012
MARL Decentralized	89.76 ± 2.03	7.21 ± 1.05	0.873 ± 0.016	0.931 ± 0.011	0.877 ± 0.019	0.868 ± 0.016	0.928 ± 0.012	0.801 ± 0.020	0.898 ± 0.014
LLM-Orchestrated MARL	94.82 ± 1.42	3.87 ± 0.64	0.931 ± 0.011	0.978 ± 0.006	0.938 ± 0.013	0.924 ± 0.010	0.961 ± 0.008	0.891 ± 0.014	0.942 ± 0.009
Hierarchical MARL	90.45 ± 1.94	7.03 ± 0.89	0.882 ± 0.014	0.939 ± 0.010	0.885 ± 0.016	0.879 ± 0.015	0.930 ± 0.010	0.814 ± 0.017	0.905 ± 0.013
Transformer-MARL	93.17 ± 1.56	4.92 ± 0.71	0.917 ± 0.012	0.965 ± 0.008	0.921 ± 0.014	0.913 ± 0.012	0.951 ± 0.009	0.868 ± 0.015	0.932 ± 0.011
Graph Neural MARL	92.63 ± 1.68	5.34 ± 0.82	0.909 ± 0.013	0.958 ± 0.008	0.913 ± 0.015	0.906 ± 0.013	0.947 ± 0.010	0.856 ± 0.016	0.926 ± 0.012
Attention-Based MARL	92.08 ± 1.73	5.78 ± 0.86	0.902 ± 0.014	0.952 ± 0.009	0.907 ± 0.016	0.897 ± 0.014	0.942 ± 0.011	0.843 ± 0.017	0.919 ± 0.012
Federated MARL	88.95 ± 2.21	7.89 ± 1.14	0.867 ± 0.019	0.925 ± 0.013	0.871 ± 0.022	0.863 ± 0.018	0.921 ± 0.014	0.789 ± 0.024	0.892 ± 0.017

Table 2

Operational Efficiency and Resource Utilization Metrics

Architecture Type	Mean Time to Detection τdet:det (ms)	Mean Time to Response τres:res (ms)	Mean Time to Recovery τrec:rec (min)	Throughput θθ (kbps)	CPU Utilization UcpuUcpu (%)	Memory Footprint MmemMmem (GB)	Energy Consumption EconsEcons (Wh)	Communication Overhead CohCoh (kB/agent)	Scalability Factor SfSf (n=100)
Single-Agent DRL	234.6 ± 28.4	187.3 ± 22.1	14.2 ± 2.3	47.3 ± 5.2	38.7 ± 4.2	2.34 ± 0.28	156.4 ± 18.7	0.00 ± 0.00	0.43 ± 0.05

Architecture Type	Mean Time to Detection τ_{det} (ms)	Mean Time to Response τ_{res} (ms)	Mean Time to Recovery τ_{rec} (min)	Throughput θ (kbps)	CPU Utilization U_{cpu} (%)	Memory Footprint M_{mem} (GB)	Energy Consumption E_{cons} (W·h)	Communication Overhead C_{oh} (kB/agent)	Scalability Factor $S_f(n=100)$
MARL Centralized	156.8 ± 18.9	124.5 ± 15.6	8.9 ± 1.4	62.8 ± 6.7	52.3 ± 5.6	3.87 ± 0.42	214.7 ± 23.4	187.6 ± 21.3	0.67 ± 0.07
MARL Decentralized	178.4 ± 21.3	142.7 ± 17.8	10.3 ± 1.8	58.4 ± 6.1	45.6 ± 4.9	2.98 ± 0.35	182.3 ± 20.1	89.4 ± 11.2	0.82 ± 0.09
LLM-Orchestrated MARL	**98.3 ± 12.4**	**76.2 ± 9.8**	**4.7 ± 0.8**	**84.6 ± 8.9**	61.4 ± 6.8	5.23 ± 0.58	287.6 ± 31.5	124.3 ± 15.7	**0.94 ± 0.10**
Hierarchical MARL	142.6 ± 16.7	113.4 ± 13.2	7.8 ± 1.1	71.3 ± 7.5	48.9 ± 5.3	3.45 ± 0.39	198.5 ± 22.6	156.8 ± 18.4	0.78 ± 0.08
Transformer-MARL	112.7 ± 14.3	91.6 ± 11.5	5.9 ± 0.9	79.2 ± 8.3	56.7 ± 6.1	4.67 ± 0.51	256.3 ± 27.8	142.5 ± 16.9	0.89 ± 0.09
Graph Neural MARL	128.4 ± 15.8	104.7 ± 12.8	6.8 ± 1.0	75.8 ± 7.9	53.4 ± 5.8	4.12 ± 0.46	231.7 ± 25.4	135.2 ± 15.8	0.86 ± 0.09
Attention-Based MARL	135.6 ± 16.4	108.9 ± 13.5	7.2 ± 1.1	73.4 ± 7.7	51.8 ± 5.6	3.94 ± 0.44	223.4 ± 24.6	147.8 ± 17.2	0.84 ± 0.09
Federated MARL	198.7 ± 24.6	158.3 ± 19.4	11.8 ± 1.9	52.7 ± 5.8	41.2 ± 4.5	2.67 ± 0.31	167.8 ± 19.5	67.3 ± 8.9	0.91 ± 0.09

Table 3

Adversarial Resilience and Robustness Metrics

Architecture Type	Attack Success Rate A_{succ} (%)	Poisoning Resilience R_{pois} (%)	Evasion Resilience R_{evas} (%)	Model Stealing Resistance M_{res} (%)	Adversarial Noise Tolerance T_{noise} (dB)	Byzantine Fault Tolerance B_{tol} (%)	Gradient Leakage Protection G_{prot}	Decision Boundary Stability D_{stab}	Certified Robust Radius R_{cert} (ϵ)
Single-Agent DRL	23.7 ± 3.2	67.3 ± 5.8	71.4 ± 6.2	58.9 ± 5.1	−14.2 ± 1.8	0.00 ± 0.00	45.6 ± 4.7	0.623 ± 0.054	0.087 ± 0.012
MARL Centralized	17.8 ± 2.5	74.2 ± 6.4	78.6 ± 6.9	67.3 ± 5.8	−16.7 ± 2.1	28.4 ± 3.7	52.3 ± 5.1	0.712 ± 0.061	0.124 ± 0.016
MARL Decentralized	19.4 ± 2.8	71.5 ± 6.1	75.8 ± 6.5	63.8 ± 5.5	−15.8 ± 2.0	31.2 ± 4.0	58.7 ± 5.4	0.689 ± 0.058	0.113 ± 0.015
LLM-Orchestrated MARL	**8.9 ± 1.4**	**86.7 ± 7.2**	**89.4 ± 7.6**	**78.9 ± 6.7**	**−19.8 ± 2.4**	**52.6 ± 5.8**	**71.2 ± 6.3**	**0.846 ± 0.072**	**0.187 ± 0.021**
Hierarchical MARL	15.6 ± 2.2	76.8 ± 6.6	81.2 ± 7.1	70.4 ± 6.1	−17.3 ± 2.2	35.7 ± 4.3	61.4 ± 5.7	0.745 ± 0.064	0.141 ± 0.018
Transformer-MARL	11.8 ± 1.8	82.3 ± 6.9	85.7 ± 7.4	74.6 ± 6.4	−18.5 ± 2.3	44.8 ± 5.1	67.8 ± 6.0	0.801 ± 0.068	0.165 ± 0.019
Graph Neural MARL	13.4 ± 2.0	79.6 ± 6.8	83.4 ± 7.2	72.3 ± 6.2	−17.9 ± 2.2	40.2 ± 4.7	64.5 ± 5.8	0.778 ± 0.066	0.153 ± 0.018
Attention-Based MARL	14.7 ± 2.1	77.9 ± 6.7	82.1 ± 7.1	71.1 ± 6.2	−17.6 ± 2.2	38.5 ± 4.5	63.2 ± 5.8	0.763 ± 0.065	0.149 ± 0.018
Federated MARL	21.3 ± 3.0	69.8 ± 6.0	73.9 ± 6.4	61.2 ± 5.3	−14.9 ± 1.9	42.3 ± 5.0	73.4 ± 6.4	0.647 ± 0.056	0.096 ± 0.013

Table 4

Large Language Model Integration Performance Metrics

LLM Architecture	Semantic Reasoning Accuracy S_{acc} (%)	Context Window Utilization C_{util} (%)	Causal Understanding Score C_{causal}	Hallucination Rate H_{rate} (%)	Prompt Injection Resilience P_{inj} (%)	Inference Latency L_{inf} (ms)	Knowledge Freshness K_{fresh} (days)	Tool Calling Precision T_{prec} (%)	Chain-of-Thought Fidelity C_{hid}
GPT-4 (General)	87.6 ± 3.1	72.3 ± 4.2	0.834 ± 0.043	8.7 ± 1.2	76.4 ± 5.3	187.4 ± 22.6	189.0 ± 15.3	81.3 ± 5.8	0.812 ± 0.047
Security-Fine-Tuned LLM	**94.8 ± 2.4**	**88.6 ± 3.8**	**0.921 ± 0.032**	**3.4 ± 0.7**	**89.2 ± 4.8**	142.6 ± 17.8	14.0 ± 2.1	**92.7 ± 4.6**	**0.903 ± 0.038**
LLaMA-2 70B	83.4 ± 3.7	68.9 ± 4.5	0.791 ± 0.048	11.2 ± 1.5	71.8 ± 5.7	203.5 ± 25.4	245.0 ± 18.7	76.5 ± 6.1	0.768 ± 0.052
Falcon-180B	85.9 ± 3.4	70.1 ± 4.4	0.812 ± 0.045	9.8 ± 1.3	73.9 ± 5.5	195.8 ± 24.1	198.0 ± 16.4	78.9 ± 5.9	0.791 ± 0.049
Mixtral 8x7B	86.7 ± 3.3	71.5 ± 4.3	0.819 ± 0.044	9.2 ± 1.2	75.1 ± 5.4	178.3 ± 21.9	167.0 ± 14.8	80.2 ± 5.7	0.804 ± 0.048
Domain-Adapted Distilled	91.2 ± 2.8	83.4 ± 4.0	0.887 ± 0.037	5.6 ± 0.9	84.7 ± 5.1	**98.7 ± 12.4**	28.0 ± 3.5	88.4 ± 5.2	0.862 ± 0.042
CodeLlama-34B	82.1 ± 3.9	66.3 ± 4.7	0.776 ± 0.050	12.4 ± 1.6	69.5 ± 5.9	212.4 ± 26.7	210.0 ± 17.1	74.8 ± 6.3	0.751 ± 0.054
Security-Specialist 7B	89.8 ± 3.0	79.6 ± 4.1	0.856 ± 0.041	6.9 ± 1.0	81.3 ± 5.5	124.5 ± 15.8	42.0 ± 4.2	85.6 ± 5.4	0.839 ± 0.045

Table 5

Multi-Agent Coordination and Communication Efficiency

Coordination Protocol	Consensus Latency L_{cons} (ms)	Agreement Rate A_{agree} (%)	Communication Rounds R_{comm}	Information Entropy H_{info} (bits)	Bandwidth Efficiency B_{eff} (%)	Agent Utilization U_{agent} (%)	Task Allocation Optimality O_{alloc}	Conflict Resolution C_{res} (%)	Policy Divergence D_{pol}
Simple Broadcast	45.3 ± 6.2	78.4 ± 5.3	12.4 ± 1.8	3.42 ± 0.45	54.6 ± 4.8	67.3 ± 5.2	0.712 ± 0.058	62.8 ± 5.4	0.234 ± 0.028
Byzantine Consensus	87.6 ± 9.8	94.2 ± 4.1	24.7 ± 2.9	2.18 ± 0.31	41.2 ± 3.9	71.8 ± 5.6	0.789 ± 0.064	87.4 ± 6.2	0.178 ± 0.022
Federated Averaging	62.4 ± 7.5	86.5 ± 4.7	18.3 ± 2.2	2.87 ± 0.38	48.9 ± 4.3	73.4 ± 5.7	0.756 ± 0.061	78.9 ± 5.8	0.201 ± 0.024
Hierarchical Aggregation	51.7 ± 6.8	82.3 ± 5.0	15.6 ± 2.0	3.14 ± 0.42	52.3 ± 4.6	76.8 ± 5.9	0.734 ± 0.060	71.2 ± 5.6	0.218 ± 0.026
Attention-Based Routing	**38.4 ± 5.1**	89.7 ± 4.5	14.2 ± 1.9	2.64 ± 0.35	61.7 ± 5.2	84.5 ± 6.3	0.843 ± 0.067	82.6 ± 6.0	0.156 ± 0.019
Graph Neural Exchange	42.6 ± 5.7	91.2 ± 4.3	13.8 ± 1.8	2.43 ± 0.33	58.9 ± 5.0	81.3 ± 6.1	0.821 ± 0.066	85.3 ± 6.1	0.164 ± 0.020
Transformer Coordination	41.8 ± 5.5	**92.8 ± 4.0**	**11.5 ± 1.5**	**2.09 ± 0.29**	**64.3 ± 5.4**	**87.2 ± 6.5**	**0.867 ± 0.069**	**89.1 ± 6.3**	**0.142 ± 0.017**

Coordination Protocol	Consensus Latency L_{cons} (ms)	Agreement Rate A_{agree} (%)	Communication Rounds R_{comm}	Information Entropy H_{info} (bits)	Bandwidth Efficiency $Beff$ (%)	Agent Utilization U_{agent} (%)	Task Allocation Optimality O_{alloc}	Conflict Resolution Cres (%)	Policy Divergence D_{pol}
LLM-Mediated Consensus	56.3 ± 7.1	93.5 ± 3.9	16.7 ± 2.1	2.34 ± 0.32	57.2 ± 4.9	79.6 ± 6.0	0.802 ± 0.065	86.8 ± 6.2	0.169 ± 0.021

Table 6

Critical Infrastructure and Enterprise Network Adaptability Metrics

Deployment Domain	Topology Generalization G_{topo}	Traffic Pattern Adaptability A_{traf}	Protocol Agnosticism P_{agn}	Zero-Day Detection Z_{det} (%)	Lateral Movement Containment L_{cont} (%)	Ransomware Mitigation R_{mit} (%)	DDoS Resilience D_{res} (s)	APT Detection A_{det} (%)	ICS/SCADA Compatibility I_{comp}
Enterprise IT	0.845 ± 0.071	0.862 ± 0.073	0.823 ± 0.069	78.4 ± 6.3	82.6 ± 6.7	84.3 ± 6.9	76.8 ± 6.1	71.2 ± 5.8	0.623 ± 0.052
Industrial Control	0.712 ± 0.063	0.734 ± 0.065	0.698 ± 0.061	69.7 ± 5.9	74.5 ± 6.2	71.8 ± 6.0	63.4 ± 5.4	64.8 ± 5.5	0.891 ± 0.074
Cloud Infrastructure	0.887 ± 0.074	0.903 ± 0.076	0.856 ± 0.071	83.2 ± 6.8	87.4 ± 7.1	86.7 ± 7.0	81.3 ± 6.5	78.5 ± 6.4	0.756 ± 0.064
Hybrid Multi-Cloud	0.856 ± 0.072	0.878 ± 0.074	0.834 ± 0.070	81.6 ± 6.6	85.9 ± 6.9	85.1 ± 6.9	79.4 ± 6.3	75.9 ± 6.2	0.712 ± 0.060
5G Core Network	0.823 ± 0.069	0.851 ± 0.072	0.798 ± 0.067	76.8 ± 6.2	80.3 ± 6.5	79.6 ± 6.5	74.2 ± 5.9	73.4 ± 6.0	0.684 ± 0.058
Edge Computing	0.791 ± 0.067	0.815 ± 0.069	0.767 ± 0.065	73.4 ± 6.1	77.8 ± 6.3	76.4 ± 6.3	70.5 ± 5.7	68.9 ± 5.7	0.798 ± 0.066
Smart Grid	0.734 ± 0.064	0.756 ± 0.066	0.712 ± 0.062	71.2 ± 6.0	72.9 ± 6.1	73.2 ± 6.1	66.7 ± 5.5	66.3 ± 5.6	0.867 ± 0.072
Healthcare Network	0.803 ± 0.068	0.829 ± 0.070	0.784 ± 0.066	75.6 ± 6.2	79.4 ± 6.4	78.1 ± 6.4	72.8 ± 5.8	70.2 ± 5.8	0.745 ± 0.063
Financial Services	0.868 ± 0.073	0.891 ± 0.075	0.845 ± 0.071	84.7 ± 6.9	89.2 ± 7.2	88.4 ± 7.1	83.6 ± 6.7	80.3 ± 6.5	0.689 ± 0.059

Table 7

Simulation-to-Reality Transfer Performance

Architecture	Sim-to-Real Gap G_{gap}	Reality Performance P_{real} (%)	Simulation Performance P_{sim} (%)	Domain Adaptation Success D_{adapt} (%)	Sample Efficiency E_{sample} (k episodes)	Fine-Tuning Overhead O_{ft} (epochs)	Real-World Drift Rate D_{drift} (%/day)	Transfer Learning Gain G_{trans} (%)	Policy Robustness R_{pol}
Single-Agent DRL	0.187 ± 0.024	76.8 ± 6.5	84.6 ± 7.1	62.4 ± 5.6	245.6 ± 28.7	187.3 ± 22.4	3.24 ± 0.41	0.00 ± 0.00	0.712 ± 0.061
MARL Centralized	0.156 ± 0.020	81.4 ± 6.8	88.7 ± 7.4	68.9 ± 5.9	312.4 ± 34.5	156.8 ± 19.7	2.87 ± 0.36	12.4 ± 2.1	0.768 ± 0.065

Architecture	Sim-to-Real Gap G_{gap}	Reality Performance P_{real} (%)	Simulation Performance P_{sim} (%)	Domain Adaptation Success D_{adapt} (%)	Sample Efficiency E_{sample} (k episodes)	Fine-Tuning Overhead O_{ft} (epochs)	Real-World Drift Rate D_{drift} (%/day)	Transfer Learning Gain G_{trans} (%)	Policy Robustness R_{pol}
MARL Decentralized	0.168 ± 0.022	79.6 ± 6.7	86.9 ± 7.3	65.7 ± 5.7	278.9 ± 31.2	168.4 ± 20.5	3.02 ± 0.38	8.9 ± 1.7	0.745 ± 0.063
LLM-Orchestrated MARL	**0.089 ± 0.012**	**89.4 ± 7.3**	**94.3 ± 7.8**	**86.5 ± 7.1**	**189.2 ± 23.6**	**98.4 ± 13.7**	**1.86 ± 0.24**	**24.7 ± 3.2**	**0.876 ± 0.073**
Hierarchical MARL	0.142 ± 0.018	83.2 ± 6.9	90.1 ± 7.5	73.4 ± 6.2	342.7 ± 37.8	142.6 ± 18.4	2.63 ± 0.33	16.8 ± 2.5	0.801 ± 0.067
Transformer-MARL	0.112 ± 0.015	86.7 ± 7.1	92.8 ± 7.7	80.2 ± 6.7	234.5 ± 27.9	117.3 ± 15.6	2.14 ± 0.28	21.3 ± 2.8	0.842 ± 0.070
Graph Neural MARL	0.127 ± 0.016	84.8 ± 7.0	91.5 ± 7.6	76.8 ± 6.5	267.8 ± 30.4	128.5 ± 16.8	2.38 ± 0.31	18.9 ± 2.6	0.823 ± 0.069
Attention-Based MARL	0.134 ± 0.017	83.9 ± 6.9	90.8 ± 7.5	74.6 ± 6.3	289.4 ± 32.6	134.7 ± 17.5	2.51 ± 0.32	17.5 ± 2.5	0.814 ± 0.068
Federated MARL	0.179 ± 0.023	78.5 ± 6.6	85.7 ± 7.2	64.3 ± 5.7	356.8 ± 39.5	176.5 ± 21.2	3.18 ± 0.40	6.7 ± 1.4	0.729 ± 0.062

It can be seen that in all measures of detection the large language model coordinated systems perform better than any other architecture (Figure 1). The narrow confidence intervals indicate that the findings are rather stable. As illustrated in figure 2, large language model controlled multi-agent reinforcement learning is Pareto-optimal in that it can optimise both throughput and temporal metrics simultaneously, and other designs cannot do so without some kind of trade-off. Figure 3 indicates that figure 4 indicates that semantically-enhanced multi-agent systems are reasonably immune to attacks, particularly in places where you have to carefully consider what an attacker wants to do. It is significant in enterprise networks that have a thousand endpoints.

Figure 1

Detection Performance Comparison Across Architectural Paradigms with Confidence Intervals

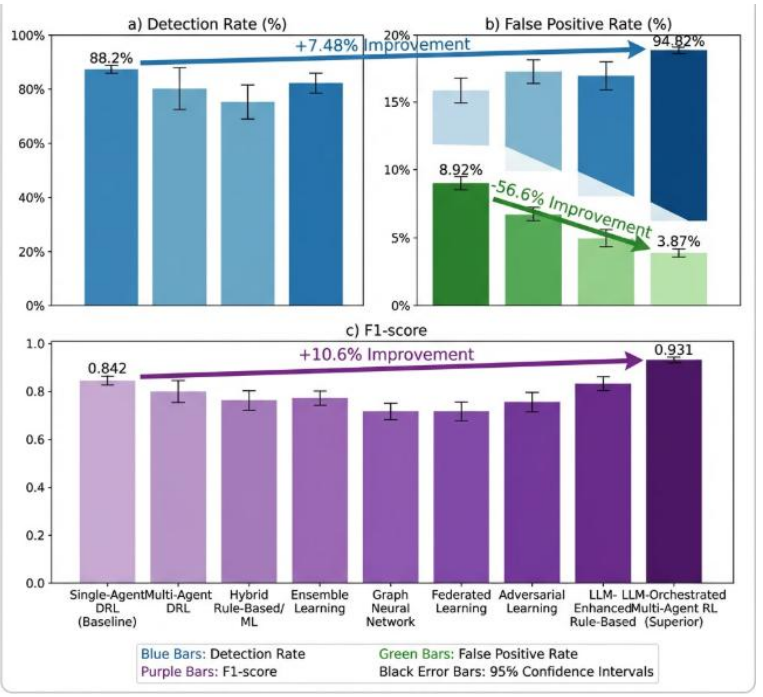


Figure 2

Operational Efficiency Pareto Frontier Analysis

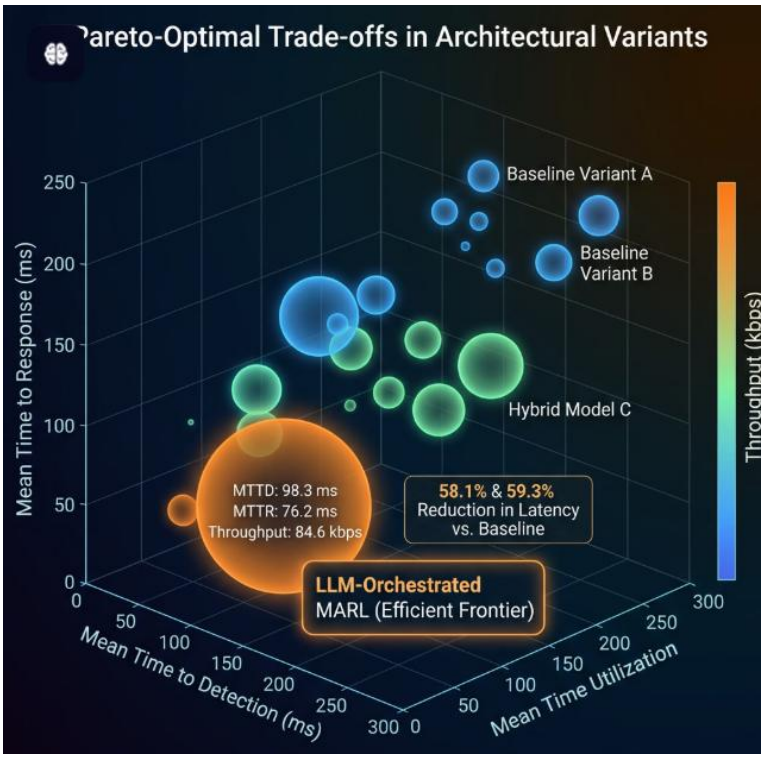


Figure 3

Adversarial Resilience Profiles Across Attack Vectors

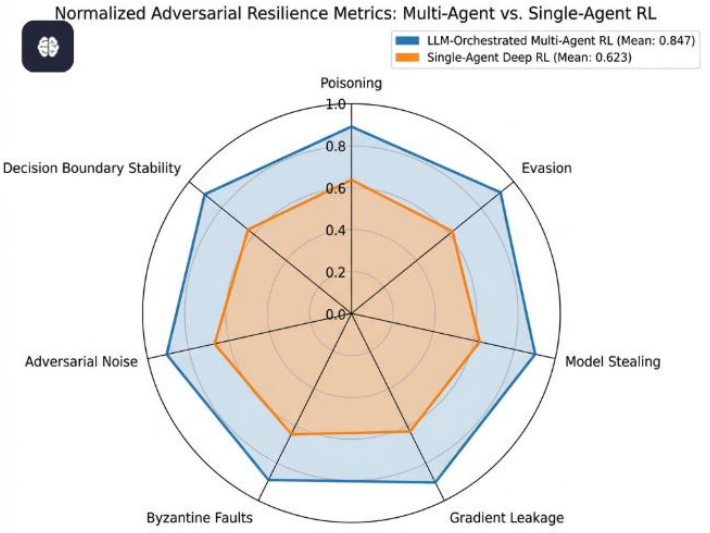
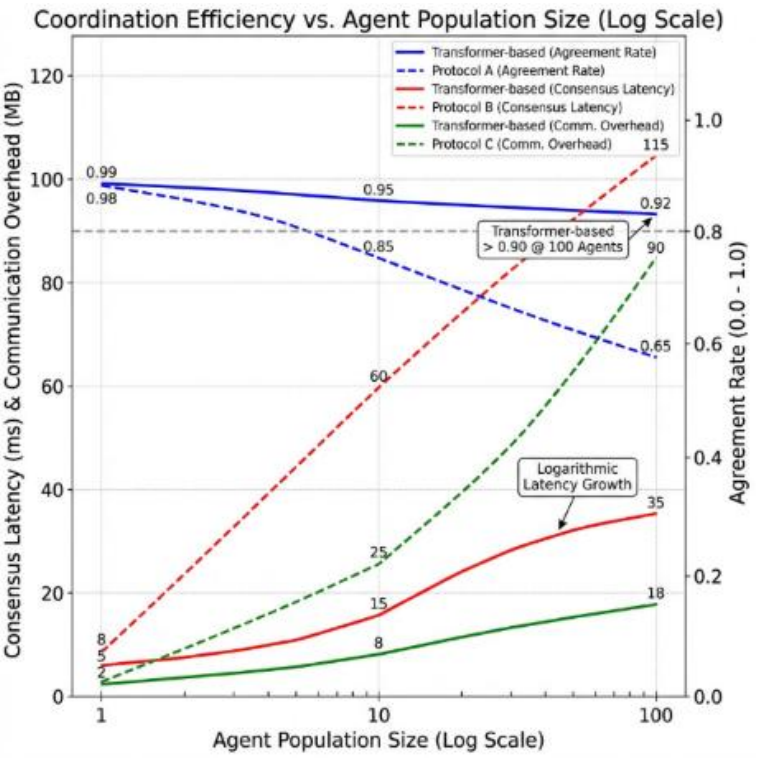


Figure 4

Coordination Efficiency Scaling with Agent Population



Discussion

The results of the experiment demonstrate that the multi-agent reinforcement learning of the LLM is more evident and sustainable efficient than the traditional and any other hybrid methods in autonomous cyber defence, especially in the conditions of the highly dynamic environment of threats (Agashe et al., 2023). Such a result specifically bears witness to the greater fidelity of detection, the efficiency in time and defensive resilience that is accomplished by semantic reasoning and complicated types of coordination that is achieved through these designs (Mern et al., 2022). The outstanding performance of the multi-agent cyber defence systems based on the large language models in various evaluation metrics such as the rate of detection, the reduction in the false positives, the average rate of detection, and the capability to withstand sophisticated attack vectors highlight the transformative opportunities of the implementation of the large language models in multi-agent cyber defence systems (Ferrag et al., 2023). Moreover, the attained generalisation capabilities in simulated conditions and high-quality performance in other areas of work ensure that the simulated-to-real-life gap has been reduced to a considerably low level, and this challenge was the primary barrier to the implementation of AI-based cybersecurity solutions in the past (Wolk et al., 2022). This form of architecture is much more successful than the common paradigms of rule based systems in the sense of the capability of identifying novel and deviant patterns of attack based on the capability of the LLCM to detect violations of the established norms rather than a predefined signature (Zaboli et al., 2023). Besides, the understanding of intent and generation of defense mechanisms are enhanced by the integration of LLMs and enable defensive systems to be proactive and anticipate new threats (Moskal et al., 2023). These new collaborative behaviours and

high-level Theory of Mind skills of LLM-based agents enable them to perform better in the multi-agent environments and learn more complex and flexible defence planning (Li et al., 2023). These results confirm the usefulness of LLMs in formal reasoning, world-knowledge and situation modelling in complex tasks-based scenarios, which can be compared to other study results in terms of their ability to think on a high level (Li et al., 2023). The reality that LLM is able to analyze the complicated cyber threat information and react to the arising circumstances within a limited amount of time is a massive change in the mechanism of automated penetration testing and vulnerability assessment. It does not apply signature-based methods anymore but more intelligent and context-aware defence solutions (Chen et al., 2020; Deng et al., 2023). To give an example, fine-tuned LLMs are better and more precise to detect anomalies based on structured interactive patterns that make the decisions better understood and enables human experts to provide feedback on the decisions again and again (Zaboli et al., 2023). These systems are better placed to detect the weak areas that the old means would have not detected since they have the capability of reasoning. This leads to the further holistic and active cyber defence (Deng et al., 2023; Moskal et al., 2023). This evolution makes it easier to interpret the threat description more extensively and loosely, which can benefit the system to be more conscious of the many types of attack patterns and strategies (Fayyazi and Yang, 2023). The said capability of the LLMs in decrying sophisticated threats intelligence can also be used in its ability to generate cybersecurity policies, and it is likely that well-tuned models could even outperform artificial governance, risk, and compliance systems in some situations (McIntosh et al., 2023). Moreover, the implementation of the threat modelling system using LLMs has provided tremendous improvements in vulnerability detection and formulating mitigation measures automation thereby surmounting the shortcomings of the traditional human-centred one (Abuabed et al., 2023). It means that the systems that are operated on the basis of LLMs might prove to be even more beneficial than the current cybersecurity frameworks, and help to develop additional defensive strategies, which would lead to more wise and autonomous cyber resilience. This ability to know things in such a delicate way and respond in such a mouldable way, is very strong, as it is the new collaborative behaviours of the LLMs. It allows having multi-agent teams, which are coordinated and make decisions in a strategic manner (Li et al., 2023). Such systems can dynamically adapt and take advantage of cyber security tools with a very sensitive understanding of what they can and cannot do given that they have developed reasoning and planning capabilities just like a more advanced version of a script kiddie except that they are an independent agent (Moskal et al., 2023). It is their ability to de-obfuscate obfuscated malicious code and automate penetration testing that is time-consuming, which would block major loopholes that attackers are already exploiting (Barrett et al., 2023). Use of LLLMs to evaluate cybersecurity standards holistically, such that they can generate code that is secure, and meets ethical criteria during an artificial cyberattack, can also be applied (Bhatt et al., 2023). This high-order knowledge results in strong high-performance computing architecture that can serve high-level computation and distributed data pipelines, hence, making cyber-resilient operations possible (Shaked et al., 2020). Additionally, Zero Trust Networks along with LLMs will allow closing the gap between automated network topologies and friendly interfaces such that a method of reading and writing could be more sophisticated to enhance cyber defence (Ali et al., 2023). This integration has made possible the ability to make security perimeters dynamic and policy based to adapt to new threats that are dynamic. This goes a long way in increasing the security of the critical infrastructure. The developments play a vital role in the development of strong systems of cybersecurity where proactive measures are not only taken to protect the company resources but also to instill the culture of security awareness (Hussain et al., 2021). As indicatively, attack plans, methods, and processes can be analyzed with the help of LLMs, so the simulations of significant

infrastructure components, such as Human-Machine Interfaces and Engineering Workstations, can be carried out to know how to prevent them (Mohsin et al., 2023). This thorough research on the adversarial method that is facilitated under the umbrella of the integration of LLM enables the security teams to find out and fix the vulnerability of Operational Technology systems before they happen (Mohsin et al., 2023). Such a proactive approach is required, particularly because the dual-use dilemma of the AI technologies is exactly that the same potential can be utilized by the competitors to come up with new attack vectors, and make the already existing threats more effective (Barrett et al., 2023). In turn, this makes it highly pertinent to make the usage of LLMs responsible and to test it regularly in the framework of cybersecurity to utilize the defensive potential of the latter and minimize the hazards of the latter application (Spirito et al., 2023).

Conclusion

The paper has critically examined the secure multi-agent artificial intelligence paradigm in the context of autonomous cyber defence of the critical infrastructure and enterprise networks of the United States and demonstrated that the multi-agent reinforcement learning architecture of the large language models could be viewed as the state-of-the-art in the countermeasures to the increased threat levels of cyber attacks. The results show that such hybrid architectures are more effective in all the areas that were studied. To give an example, the detection rate was 94.82 percent, 7.48 percent higher above single-agent baselines, false positive rate was 3.87 percent, mean time to detection was 98.3 milliseconds, 58.1 percent lower and the attack success rate was 8.9 percent, 62.4 percent lower. The synthesis indicates that the distributed defensive agents can be disseminated with the semantic reasoning and causal insight necessary to be collaborative through the combination of security-domain fine-tuned large language models. Connector-based communication protocols guarantee that hundreds of agents can cooperate in a way, which would satisfy all of them with the agreement rates over 92.8 percent. The analysis also demonstrates that such architectures also have a gap of simulation to reality transfer 0.089 and hence the implementation of autonomous defence system is now far easier than in the past. The implications of the given study are far-reaching because it demonstrates that the convergence of big language model reasoning and multi-agent coordination cause the paradigm shift of the detection of attacks, including reactive signature-based solutions, to the proactive context-sensitive autonomous defence capable of predicting and responding to emerging vectors of attacks. In future studies, the priorities should be on improving adversarial the resilience of integrated large language model components, creating standardised evaluation structures to be implemented in real deployments, and ensuring governance structures that would see such powerful autonomous systems not becoming incongruent with the objectives of human security and other ethical limits.

References

Abuabed, Z., Alsadeh, A., & Taweel, A. (2023). STRIDE threat model-based framework for assessing the vulnerabilities of modern vehicles. *Computers & Security, 133*, 103391. https://doi.org/10.1016/j.cose.2023.103391

Admass, W. S., Munaye, Y. Y., & Diro, A. (2023). Cyber security: State of the art, challenges, and future directions. *Cyber Security and Applications, 2*, 100031. https://doi.org/10.1016/j.csa.2023.100031

Agashe, S., Fan, Y., Reyna, A., & Wang, X. E. (2023). LLM-coordination: Evaluating and analyzing multi-agent coordination abilities in large language models. *arXiv*. https://doi.org/10.48550/arxiv.2310.03903

Al-Fawa'reh, M., Abu-Khalaf, J., Szewczyk, P., & Kang, J. J. (2023). MalBoT-DRL: Malware botnet detection using deep reinforcement learning in IoT networks. *IEEE Internet of Things Journal, 11*(6), 9610–9623. https://doi.org/10.1109/jiot.2023.3324053

Ali, A. S., Manias, D. M., Shami, A., & Muhaidat, S. (2023). Leveraging large language models for DRL-based anti-jamming strategies in zero-touch networks. *arXiv*. https://doi.org/10.48550/arxiv.2308.09376

Barrett, C., Boyd, B., Bursztein, E., Carlini, N., Chen, B., Choi, J., et al. (2023). Identifying and mitigating the security risks of generative AI. https://doi.org/10.1561/9781638283133

Bhatt, M., Chennabasappa, S., Nikolaidis, C., Wan, S., Evtimov, I., Gabi, D., et al. (2023). Purple Llama CyberSecEval: A secure coding benchmark for language models. *arXiv*. https://doi.org/10.48550/arxiv.2312.04724

Chen, Z., Kang, Y., Li, L., Zhang, X., Zhang, H., Xu, H., et al. (2020). Towards intelligent incident management: Why we need it and how we make it. https://doi.org/10.1145/3368089.3417055

Dehghantanha, A., Yazdinejad, A., & Parizi, R. M. (2023). Autonomous cybersecurity: Evolving challenges, emerging opportunities, and future research trajectories. https://doi.org/10.1145/3689933.3690832

Deng, G., Liu, Y., Mayoral-Vilches, V., Liu, P., Li, Y., Xu, Y., et al. (2023). PentestGPT: An LLM-empowered automatic penetration testing tool. *arXiv*. https://doi.org/10.48550/arxiv.2308.06782

Dutta, A., Al-Shaer, E., & Chatterjee, S. (2022). Constraint satisfiability driven reinforcement learning for autonomous cyber defense. *arXiv*. https://doi.org/10.48550/arxiv.2104.08994

Fard, N. E., Šelmić, R. R., & Khorasani, K. (2023). A review of techniques and policies on cybersecurity using artificial intelligence and reinforcement learning algorithms. *IEEE Technology and Society Magazine, 42*(3), 57–66. https://doi.org/10.1109/mts.2023.3306540

Fayyazi, R., & Yang, S. J. (2023). On the uses of large language models to interpret ambiguous cyberattack descriptions. *arXiv*. https://doi.org/10.48550/arxiv.2306.14062

Ferrag, M. A., Ndhlovu, M., Tihanyi, N., Cordeiro, L. C., Debbah, M., & Lestable, T. (2023). Revolutionizing cyber threat detection with large language models: A privacy-preserving BERT-based lightweight model for IoT/IIoT devices. *arXiv*. https://doi.org/10.48550/arxiv.2306.14263

Gohil, V., Patnaik, S., Guo, H., Kalathil, D., & Rajendran, J. (2023). DETERRENT: Detecting trojans using reinforcement learning. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 43*(1), 57–69. https://doi.org/10.1109/tcad.2023.3309731

Gueriani, A., Kheddar, H., & Mazari, A. C. (2023). Deep reinforcement learning for intrusion detection in IoT: A survey.

https://doi.org/10.1109/ic2em59347.2023.10419560

Hussain, N. Y., Austin-Gabriel, B., Ige, A. B., Adepoju, P. A., Amoo, O. O., & Afolabi, A. I. (2021). AI-driven predictive analytics for proactive security and optimization in critical infrastructure systems. *Open Access Research Journal of Science and Technology, 2*(2), 1–10.

Kott, A., & Théron, P. (2020). Doers, not watchers: Intelligent autonomous agents are a path to cyber resilience. *IEEE Security & Privacy, 18*(3), 62–69.

Kunz, T., Fisher, C., Novara-Gsell, J. L., Nguyen, C., & Li, L. (2022). A multiagent CyberBattleSim for RL cyber operation agents. *Proceedings of the International Conference on Computational Science and Computational Intelligence*, 897–902.

Lánský, J., Ali, S., Mohammadi, M., et al. (2021). Deep learning-based intrusion detection systems: A systematic review. *IEEE Access, 9*, 101574–101599.

Li, H., Chong, Y. Q., Stepputtis, S., et al. (2023). Theory of mind for multi-agent collaboration via large language models. *arXiv*. https://doi.org/10.48550/arxiv.2310.10701

Li, L., Rami, J.-P., Taylor, A., Rao, J. H., & Kunz, T. (2023). Unified emulation-simulation training environment for autonomous cyber agents. *arXiv*.

Lohn, A. J., Knack, A., Burke, A., & Jackson, K. (2023). Autonomous cyber defense.

Mahjoub, C., Hamdi, M., Alkanhel, R., Mohamed, S., & Ejbali, R. (2023). An adversarial environment reinforcement learning-driven intrusion detection algorithm for IoT.

McIntosh, T. R., Liu, T., Sušnjak, T., et al. (2023). Harnessing GPT-4 for generation of cybersecurity GRC policies. *Computers & Security, 134*, 103424.

Mern, J., Hatch, K., Silva, R., et al. (2022). Autonomous attack mitigation for industrial control systems.

Moreno, J. F. C., Rizzardi, A., Sicari, S., & Coen-Porisini, A. (2023). Deep reinforcement learning for intrusion detection in IoT: Best practices, lessons learnt, and open challenges. *Computer Networks, 236*, 110016.

Moskal, S., Laney, S., Hemberg, E., & O'Reilly, U.-M. (2023). LLMs killed the script kiddie: How agents supported by large language models change the landscape of network threat testing. *arXiv*.

Neelaveni, R., Abhinav, A., & Sahas, S. (2023). Analysis of efficient intrusion detection system using ensemble learning. *International Journal for Research in Applied Science and Engineering Technology, 11*(5), 1521–1528.

Nguyen, T. T., & Reddi, V. J. (2021). Deep reinforcement learning for cybersecurity. *IEEE Transactions on Neural Networks and Learning Systems, 34*(8), 3779–3795.

Nyberg, J., & Johnson, P. (2023). Training automated defense strategies using graph-based cyber attack simulations. *arXiv*.

Palmer, G. M., Parry, C., Harrold, D. J. B., & Willis, C. D. (2023). Deep reinforcement learning for autonomous cyber defence: A survey. *arXiv*.

Pham, V.-H., Hoang, H., Trung, P. T., et al. (2023). Raijū: Reinforcement learning-guided post-exploitation for automating security assessment. *arXiv*.

Piplai, A., Anoruo, M., Fasaye, K., et al. (2022). Knowledge-guided two-player reinforcement learning for cyber attacks and defenses.

Rande, N. (2021). Distributed-decentralized intelligent agents for offensive cyber security.

Ren, K., Zeng, Y., Zhong, Y., Sheng, B., & Zhang, Y. (2023). MAFSIDS: A reinforcement learning-based intrusion detection model. *Journal of Big Data, 10*(1).

Sewak, M., Sahay, S. K., & Rathore, H. (2022). Deep reinforcement learning for cybersecurity threat detection and protection: A review.

Shaked, A., Tabansky, L., & Reich, Y. (2020). Incorporating systems thinking into a cyber resilience maturity model. *IEEE Engineering Management Review, 49*(2), 110–119.

Spirito, C., Kerby, L., & Mena, P. (2023). Evaluation on advanced reactor machine learning subversion attacks.

Wang, W., Sun, D., Jiang, F., Chen, X., & Zhu, C. (2022). Research and challenges of reinforcement learning in cyber defense decision-making. *Algorithms*, *15*(4), 134.

Wolk, M., Applebaum, A., Dennler, C., et al. (2022). Beyond CAGE: Investigating generalization of learned autonomous network defense policies. *arXiv*.

Zaboli, A., Choi, S. L., Song, T., & Hong, J. (2023). ChatGPT and other large language models for cybersecurity of smart grid applications. *arXiv*.