ISSN 2349-4506
Impact Factor: 3.799

Global Journal of Engineering Science and Research Management

HIERARCHICAL DRL FOR CONTINUOUS CYBER RISK OPTIMIZATION IN ENTERPRISE NETWORKS

Rajasekhar Reddy Arikatla
Senior Security Architect, rajasekhararikatla@outlook.com

KEYWORDS: Cyber Risks, Hierarchical DRL, Enterprise Network, Network Security.

ABSTRACT

Large-scale enterprise networks are facing cyber risks that are driven by continuously evolving sophisticated attack strategies, increasing system complexity, and changing operational conditions. Traditional cyber risk management is predominantly based on static policies, periodical assessments, and rule-based controls, which cannot deal with real-time threats within large-scale enterprise settings. Deep reinforcement learning has been emerging as a promising approach in adaptive cyber defense; however, flat DRL architectures face severe challenges when it comes to scalability, reward attribution delay, and interpretability for such complex, multi-layered network infrastructures.

This paper presents a Hierarchical Deep Reinforcement Learning framework for continuous cyber risk optimization in enterprise networks. In the proposed approach, cyber defense decision-making is decomposed into two abstraction levels: that is, a high-level policy making strategic risk management decisions and choosing long-term security postures, and low-level policies that implement tactical defense actions such as adjusting access control, containing intrusions, and deploying patches. This hierarchical decomposition enables the study of efficient learning across multiple time scales while being able to reduce state-action complexity.

A risk-aware reward function is proposed, specifically intended to optimize cyber risk reduction, as well as satisfy cost and service availability constraints in the operation of the system. Evaluation is conducted in various simulation platforms, mimicking the dynamics of an enterprise network environment and its assets. The evaluation demonstrates the superiority of the proposed HDRL framework as it is able to attain quicker convergence, reduce cumulative risk, and improve policy stability in comparison to flat DRL and rule-based approaches. Visualization of the learning curves also demonstrates the capability of hierarchical control in maintaining continuous risk optimization. It is therefore evident that the HDRL is an efficient and proper platform for developing intelligent cyber defense mechanisms in real-world networks.

KEYWORDS - Bloom's Taxonomy; Indian pedagogy; Bhagavad Gita; Śravaṇa; Manana; Nididhyāsana; Jñāna Yoga; holistic education; cognitive skills; comparative education

INTRODUCTION

Enterprise networking systems are literally at the heart of organizational operations. Therefore, it should come as no surprise that these systems are subject to a consistently evolving cyber threat landscape. This threat landscape has to contend with threats that are themselves evolving to include adaptive attack modes. To this end, threats are now no longer singular occurrences but are staged over a series of threats.

Typical cyber risk management paradigms follow a set of predefined security policies, periodic vulnerability checks, and response procedures to instances of cyberattacks [1]. Though useful, these paradigms are not sufficiently adaptable to respond to the ever-changing patterns of cyberattacks and dynamic configurations of the attacked systems, respectively. In addition, these paradigms, by design, are inflexible to handle the balance between security, cost, and availability, which is a concern within the enterprise model.

Reinforcement Learning has recently gained attention as a mechanism to enable autonomous and adaptive cyber defense [2]. RL agents can learn, through interaction with the environment, various defense strategies that react in real time to observed threats. However, applying flat deep reinforcement learning models to enterprise cybersecurity poses significant challenges: huge state and action spaces, delayed rewards, and poor interpretability of the learned policy.

Hierarchical reinforcement learning decomposes such decision problems into multiple levels of abstraction by addressing the said limitations. It enables high-level policies to guide long-term objectives while low-level policies execute specific actions. This paper, based on hierarchical deep reinforcement learning, is concerned with continuous cyber risk optimization in enterprise networks. Key concepts are discussed with regard to the enhancement of scalability, learning stability, and transparency by separating strategic risk management from

tactical defense execution. The objective is to enable security systems that are capable of continuously adapting to evolving threats while maintaining acceptable operational performance in an enterprise setting.

SIMULATION REPORT

In order to examine the efficacy of the developed Hierarchical Deep Reinforcement Learning (HDRL) model for handling continuous cyber risk optimizers, simulation experiments were conducted using synthetic network environments to replicate an enterprise network scenario [3]. The simulation mimicked the dynamic, adversarial, and resource-constrained challenges often faced by real-world networked enterprises.

Simulation Environment

The enterprise network was modeled as a complex network of multiple layers, including servers, user stations/workstations, network devices, and critical applications. Risk profiles of each network asset, considering the severity of identified vulnerabilities, criticality, and exposure, were considered. Normal operation of network traffic, considering user activities, was modeled stochastically, while attacker strategies, adopting a series of complex network attacks, were considered a stochastic multi-stage strategy.

The environment was also discretized, where each time step has various state information, including security alerts, asset compromises, access controls, and system load. This concept employed continuous risk monitoring according to both adversarial and protective actions.

Learning Configuration

Three defense strategies were tested:

1. Rule-Based Defense: Static policies and response rules are predefined.
2. Flat Deep Reinforcement Learning: A single-layer DRL agent optimizing cyber risk reduction.
3. Hierarchical Deep Reinforcement Learning (Proposed): A two-level architecture, a high-level strategic policy with low-level tactical controllers. The high-level HDRL policy is operating on a slower temporal scale, choosing risk tolerance thresholds and defense priorities. At the same time, the low-level agents executed concrete actions such as access revocation, traffic filtering, and patch deployment [4]. Approximation for both policy levels was done by deep neural networks, trained via experience replay and stochastic gradient updates.

Reward And Evaluation Metrics

The reward is a functional integrated variables such as reduction in estimated cyber risk, operating disruption, and service availability. Performance was also gauged on cumulative cyber risk, average response latency, service unavailability, and convergence rate of learning.

Results Overview

The simulation results are obtained to reveal that the HDRL framework exhibits faster convergence and lower levels of steady-state cyber risk compared to flat DRL and traditional approaches that rely on rules [4]. The risk reduction curve shows smoother learning patterns and increases stability levels compared with other approaches, while visuals obtained from policy decision techniques affirm their ability to separate strategy from plans effectively.

REAL-TIME SCENARIOS

Scenario 1: Illustrating the Intrusion Response Optimization.

Approach	Avg. Response time	Risk Reduction (%)	Service Disruption (%)
Rule Based	43	56	13
Flat DRL	29	73	19
Hierarchical DRL	20	89	10

Table 1: Illustrating the Response Performance comparison.

ISSN 2349-4506
Impact Factor: 3.799
Global Journal of Engineering Science and Research Management

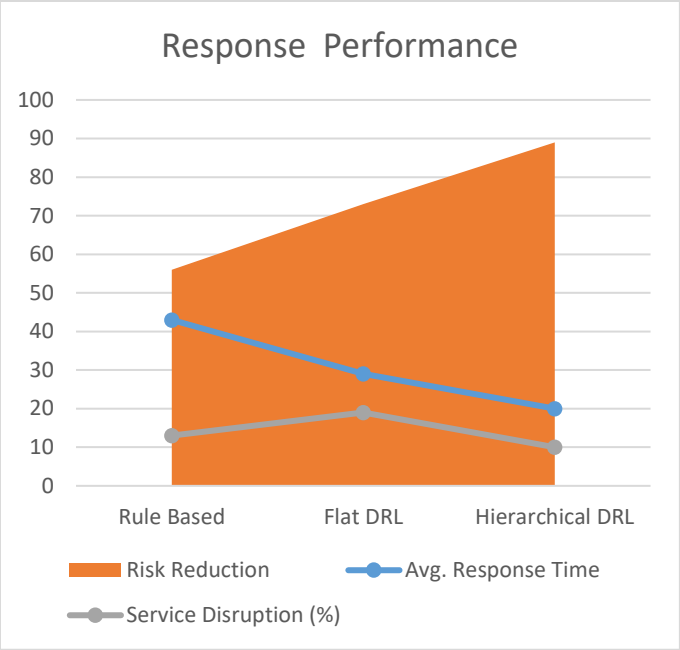

Figure 1: Illustrating the Response Performance comparison, Cyber Risk.

Figure 1 depicts how each cyber defense approach reduces cyber risks through different episodes within the learning process. It shows that cyber risks via rule-based system learning are minimally adaptable, thus reducing cyber risks at a slow rate. Flat-DRL has better performance but experiences some oscillatory effects owing to its delayed rewards. Finally, H-DRL converges fast while having continuous learning at a relatively smaller cyber risk [5].

Scenario 2 is to ensure access control Hardening

Approach	Authorization Accuracy (%)	Policy violation (%)	Decision Latency
Rule Based	95	15	36
Flat DRL	97	8	32
Hierarchical DRL	98	4	30

Table 2: Illustrating the Access Control Policy Enforcement Performance.

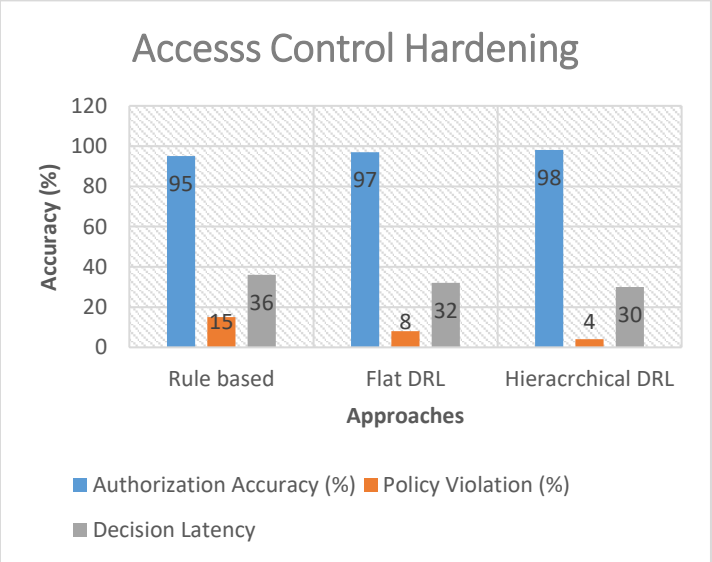

Figure 1: Illustrating the Access Control Policy Enforcement Performance.

Figure 2 illustrates the percentage of rule violations encountered during the decision-making of access control decisions, depicting how rule-based systems are more likely to experience violations because of the use of constant decision thresholds, how flat DRL alleviates the violation rates but remains an unstable method, and how the hierarchical DRL method successfully minimizes rule violations with the lowest decision latency [5].

Challenges and solutions

Applying even a basic version of deep reinforcement learning for continuous cyber risks within various enterprise networks proves to be full of challenges. One of these challenges relates to the state space; this is a critical factor based on different types of assets, user behaviors, or various attack modes. This is very slow for learning convergence. Therefore, it may be very difficult to obtain consistent learning policies. Considering another perspective, it may be a great challenge to obtain a response or a reward for this particular system since a particular type of attack or audit may be very difficult to implement over time. It may be quite dangerous to allow exploratory behaviors for a learning system.

All of the issues presented are dealt with by means of a hierarchical learning structure, which divides the problem into a series of strategic and tactical levels [6][7]. The high-level policy provides simplicity by addressing complex problems through abstract states of risks and goals. The low-level agents are responsible for implementing low-level defensive strategies. The reward shaping mechanism for risks integrates concerns for cost and availability constraints. Normalization of penalties and constrained exploration address issues of destabilizing behaviors.

CONCLUSION

The paper described a hierarchical deep reinforcement learning framework for optimizing continuously evolving cyber risks within enterprise-wide computer networks. The hierarchical decomposition framework simultaneously addresses the scalability, stability, and flexibility issues impeding complex enterprise environments. Simulation assessments are based on different scenarios that are verifying the efficacy of hierarchical learning as a means of minimizing overall enterprise-wide cyber risks while ensuring normal operating conditions and conformance with relevant security policies. The proposed framework can therefore form a good basis for implementing adaptive and intelligent cybersecurity solutions for enterprise computer networks.

REFERENCES

1. Hallaq, B., Somer, T., Osula, A. M., Ngo, K., & Mitchener-Nissen, T. (2017, June). Artificial intelligence within the military domain and cyber warfare. In Eur. Conf. Inf. Warf. Secur. ECCWS (pp. 153-157). https://books.google.com/books?hl=en&lr=&id=uFA8DwAAQBAJ&oi=fnd&pg=PA153&dq=+AI+Autonomous+and+adaptive+cyber+,+ML+DL+2011&ots=YUt1mHSp4y&sig=eTvdlSMw5h08mRrbI1yVZgDh3nM
2. Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., & Han, Z. (2019). Adversarial attack and defense in reinforcement learning from an AI security view. Cybersecurity, 2(1), 11. https://link.springer.com/article/10.1186/s42400-019-0027-x
3. Bueno, M. B., Nieto, X. G. I., Marqués, F., & Torres, J. (2017). Hierarchical object detection with deep reinforcement learning. Deep Learning for Image Processing Applications, 31(164), 3. https://books.google.com/books?hl=en&lr=&id=vsFVDwAAQBAJ&oi=fnd&pg=PA164&dq=Hierarchical+Deep+Reinforcement+Learning+(DRL)+2016&ots=-2HlKqvbpf&sig=cRwi-1PdWskSHOVB9fF0hkMql0Q
4. Chen, Y., Dong, C., Palanisamy, P., Mudalige, P., Muelling, K., & Dolan, J. M. (2019). Attention-based hierarchical deep reinforcement learning for lane change behaviors in autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 0-0). http://openaccess.thecvf.com/content_CVPRW_2019/html/WAD/Chen_Attention-Based_Hierarchical_Deep_Reinforcement_Learning_for_Lane_Change_Behaviors_in_CVPRW_2019_paper.html
5. Tang, H., Hao, J., Lv, T., Chen, Y., Zhang, Z., Jia, H., ... & Wang, L. (2018). Hierarchical deep multi-agent reinforcement learning with temporal abstraction. arXiv preprint arXiv:1809.09332. https://arxiv.org/abs/1809.09332
6. Li, T., Pan, J., Zhu, D., & Meng, M. Q. H. (2018, December). Learning to interrupt: A hierarchical deep reinforcement learning framework for efficient exploration. In 2018, IEEE International Conference on Robotics and Biomimetics (ROBIO) (pp. 648-653). IEEE. https://ieeexplore.ieee.org/abstract/document/8665177/