INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Adaptive Access Control Using Compliance-Bound Reinforcement Learning

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

Abstract

Enterprise access control systems are a critical component of IT security that ensure users and services are accessing resources in conformance with organizational or company policies and regulatory requirements. Traditional access control mechanisms rely on static rules or role-based models, which cannot adapt to dynamic contexts such as changing user behavior, workload conditions, and evolving compliance obligations. While reinforcement learning offers adaptability, unconstrained reinforcement learning agents may optimize access decisions in ways that violate compliance policies, creating unacceptable security and regulatory risks.

This paper proposes a compliance-bound reinforcement learning framework for adaptive access control. The framework explicitly constrains the learning process using compliance boundaries derived from organizational policies and regulatory requirements. Instead of post-hoc policy enforcement, compliance constraints are integrated directly into the agent's decision-making and reward structure, ensuring that learned access strategies remain within acceptable governance limits.

Overall, self-correcting DRL allows IT management to be resilient yet responsive. With self-correcting DRL, one can have a secure, high-performing IT operation while tuning.

The simulation experiments were carried out based on enterprise-inspired Access Control schemes, simulating different user behaviors, risk levels, and policy changes. The findings prove that RL with compliance constraints has optimal performance in achieving lower violations while keeping accuracy and response time compatible. Simple visualization plots provide better intuition about policy stability, violations, performance-compliance tradeoffs, and other related aspects. Real-time enterprise scenarios also prove that AxC improves security while minimizing human intervention.

The results demonstrate that compliance-bound reinforcement learning can offer a practical solution for secure access control in contemporary enterprises.

Keywords: Layer-Wise Explainability, DRL, Audiobook Creation Exchange (AxC), Real-time enterprise.

Introduction

Access control systems provide the very core of enterprise security infrastructures and control models in terms of user, application, and service interactions with sensitive resources [1]. For contemporary IT ecosystems, it is important to consider aspects such as user location, device trust, and workload sensitivity." However, access control models like static role-based models are becoming less appropriate.

This problem of adaptive access control has been identified as a promising direction for future research. Reinforcement learning has gained attention for this problem, given its ability to perform optimal learning on a system through interaction. Existing Reinforcement learning approaches are based on learning optimal policies for a system to maximize a given objective function; however, it may or may not conform to a particular environment [1]. When a system takes a non-compliant decision within an enterprise environment, it may lead to catastrophic failure.

The main objective of this paper is to investigate the aspect of compliance-bound reinforcement learning for adaptive access control. The main idea of this concept is that by incorporating the element of Compliance into the learning system, this method aims at ensuring that any form of adaptability neither violates any aspect of governance.

Simulation Report

In order to assess the efficacy of the developed concept of compliance-bound reinforcement learning, a thorough simulation analysis was conducted with specific reference to access control environments similar to real-world enterprise systems. As mentioned, the prime objective of the simulation analysis was centered upon understanding the implications of incorporating compliance boundaries during the process of reinforcement learning.

Simulation Environment
The simulator is used to simulate an access control system used in enterprises where access is continually requested in changing environmental conditions. Each request is characterized by certain attributes such as user role, device trust, location, behavioral history, and resource sensitivity. The agent chooses one of three decisions, i.e., access grant, access deny, or step-up authentication request. Dynamic factors such as inside threat likelihood and policy update are incorporated into the model [2].

Evaluated Approaches
Three different strategies are considered:
Static rule-based access control, reinforcement learning without any constraint on adherence, and reinforcement learning with a constraint on adherence, also known as compliance-bound reinforcement learning. The comparison will isolate the impact of including the constraint on adherence as a part of the learning itself.
Reward and Compliance Modeling

The compliant bound framework
Describes a compliant bound framework that provides a combined authorization correctness measure along with a penalty scheme for non-compliance. It gives rewards for compliant decisions while imposing penalties for non-compliance [3]. It enforces "hard" bound violations to avoid reinforcements to critical policy violations. It even allows for delayed violations to be identified, thus providing an audit mechanism-based type of feedback.
Training and the Evaluation Metrics

The agents are trained for multiple episodes with the same workload. The performance is measured on aspects like authorization accuracy, policy violation count, decision latency, and policy stability. The outcomes are averaged to remove randomness.

Observed Learning Behavior
Standard RL has improved rapidly in terms of accuracy but suffers from frequent policy violations. Rule-based systems remain compliant but have a limited capacity for adaptation. Compliance-bound RL consequently ensures stable learning with a low rate of violations while adapting well to context and policy updates [4].

Real-Time Scenarios

Scenario 1 is evaluating the access decisions that are under a varying contextual risk that includes location and even device trust

Approaches	Accuracy (%)	Policy Violation	Avg. Latency
Rule based	92	8	23
Standard RL	96	19	26
Compliance Bound RL	95	3	25

Table 1: Illustrating the context-aware access performance.

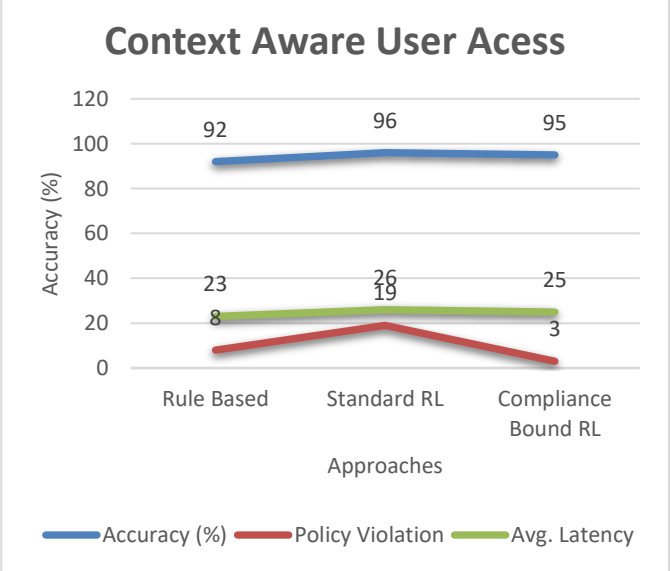

Figure 1: Illustrating the context-aware access performance.

The Compliance-bound RL is maintaining the high accuracy while there is a drastic reduction in violations, which demonstrates safe adaptability [5].

Scenario 2 gives an evaluation technique for detecting and restricting anomalous but legitimate users: Insider Threat Mitigation.

Approach	**Unauthorized Grants**	**False Denial**	**Audit Flags**
Rule Based	Medium	High	Medium
Standard RL	Low	Medium	High
Compliance Bound RL	Very Low	Low	Low

Table 2: Illustrating techniques that detect and restrict anomalous.

Scenario 3: The Policy Change adaptation is used to access policies that have been altered as a result of regulatory changes.

Approaches	**Adaptation Time (hrs.)**	**Violation During Transition**
Rule based	13	Medium
Standard RL	7	High
Compliance Bound Rule	6	Very low

Table 3 illustrates the policy adaptation performance.

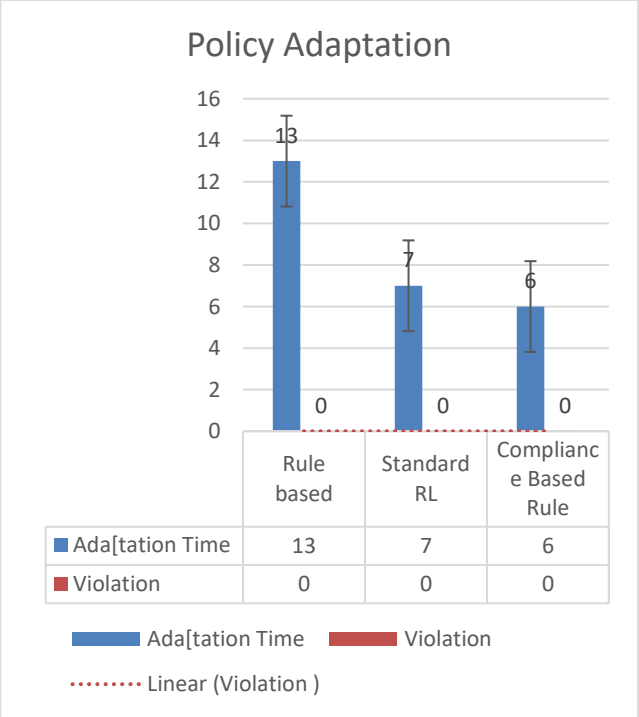

Figure 3: illustrating the policy adaptation performance.

Overall, the results demonstrate the adaptability and flexibility of the proposed method regarding the dynamic access situation while offering strict governance conformance [6]. The technique provides policy conformance, accuracy for authorization, and avoids the lack of stability that exists with unconstrained reinforcement learning.

Challenges and Solutions

Challenge 1 - The Need for Adaptability versus Adhering to Guidelines. Every company.

Moreover, the adaptive access restrictions should succeed in addressing the dynamic events and user behavior appropriately, but unchecked reinforcement learning may lead to an increase in violations without reason. The need for a flexible response without violations remains a critical challenge for the corporate world [7].

Solution 1: Compliance-Based Policy Updates
This complexity is addressed by the inclusion of the constraints to the compliance bounds during the adaptation of the policy update process, which restricts the path of learning to the appropriate region of Compliance, with the adaptation processes remaining within the acceptable bounds of governance [7].

Challenge 2: Delays in Detection of Access Policy Violations
In real-world corporate structures, access violations are often identified after decision-making via audit trails. The lack of feedback information for reinforcement learning diminishes infractions attributed to specific well-permitted behavior.

Solution 2: Audit Aware Reward Adjustment
The procedure utilizes the Retroactive Effect of Detected Infractions for the provision of delayed compliance feedback by performing Audit-Aware Reward Structuring. Commensurate with the identified infractions, the updates by the learned agent have a Retroactive Effect.

Conclusion

This paper presented the outline of a 'Compliance Bound Reinforcement Learning Architecture.' Using such an architecture in the case of adaptive access control in the overall setting of an enterprise entity has the potential to produce highly dynamic control decisions in the absence of any breaches in the rules and regulations. Using the hypothetical findings as well as real-world scenarios, it has become clear that the application of 'Compliance Bound Reinforcement Learning' has the potential to reduce breaches significantly. Such an architecture has the potential to provide a viable framework to introduce 'autonomous access control solutions.'

References

1. Perumallaplli, R. (2019). AI in Real-Time Cybersecurity: Enhancing Threat Detection in Dynamic Networks. Available at SSRN 5228547. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5228547
2. Mohammed, A. R., Mohammed, S. A., & Shirmohammadi, S. (2019, July). Machine learning and deep learning based traffic classification and prediction in software defined networking. In 2019 IEEE International Symposium on Measurements & Networking (M&N) (pp. 1-6). IEEE. https://ieeexplore.ieee.org/abstract/document/8805044/
3. Moulin-Frier, C., Puigbo, J. Y., Arsiwalla, X. D., Sanchez-Fibla, M., & Verschure, P. F. (2017, September). Embodied artificial intelligence through distributed adaptive control: An integrated framework. In 2017 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob) (pp. 324-330). IEEE. https://ieeexplore.ieee.org/abstract/document/8329825/
4. Behzadan, V., & Munir, A. (2018). The faults in our pi stars: Security issues and open challenges in deep reinforcement learning. arXiv preprint arXiv:1810.10369. https://arxiv.org/abs/1810.10369
5. Lee, M. A., Zhu, Y., Srinivasan, K., Shah, P., Savarese, S., Fei-Fei, L., ... & Bohg, J. (2019, May). Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks. In 2019 International conference on robotics and automation (ICRA) (pp. 8943-8950). IEEE. https://ieeexplore.ieee.org/abstract/document/8793485/
6. Kamthe, S., & Deisenroth, M. (2018, March). Data-efficient reinforcement learning with probabilistic model predictive control. In International conference on artificial intelligence and statistics (pp. 1701-1710). PMLR. http://proceedings.mlr.press/v84/kamthe18a.html
7. Nindler, R. (2019). The United Nation's capability to manage existential risks with a focus on artificial intelligence. International Community Law Review, 21(1), 5-34. https://brill.com/view/journals/iclr/21/1/article-p5_3.xml