Concept for Summer Research Focus: An Adaptive AI Framework for Ransomware Defense

Alex Towell
atowell@siue.edu
Department of Computer Science
Southern Illinois University-Edwardsville

May 10, 2025

Introduction

I've been developing an advanced AI framework to tackle ransomware, combining reinforcement learning principles with large language models. This proposal outlines a focused research direction that aligns with our departmental expertise in ransomware, AI, and machine learning.

I'm requesting approval for a 3-credit independent study/research course this summer under Dr. Fujinoki's primary guidance and with Dr. Gultepe's collaborative input, with optional co-mentorship or collaborative input from Dr. Ahmed Imteaj (SIUC), whose research in AI-based security may align with this proposal.

This focused period would allow me to make progress on what could become a significant component of my dissertation research. This research direction connects to several active areas in the literature, including interpretable AI for security applications, reinforcement learning for adaptive defense, and LLM-augmented decision systems. Our approach is distinct in combining all three elements into a unified framework specifically optimized for ransomware defense.

The Big Idea: Intelligent, Adaptive Ransomware Defense as an MDP

The core concept is an AI system that learns to proactively and intelligently handle ransomware threats by modeling the defense process as a Markov Decision Process (MDP). Instead of static defenses, this system would feature:

- **A Cognitive Core**: An RL-trained policy (based on a DNN) that learns optimal strategies for detection, response, and mitigation through Q-learning.
- **State Representation**: States represented by embeddings of the current context (system logs, network activity, file behaviors, etc.), capturing the essential semantics of the environment.
- **Discrete Action Space**: The policy selects from a set of specific actions including:
  - *Threat Decomposition*: Breaking down complex attack patterns into identifiable components
  - *Context Retrieval*: Querying a knowledge base of known ransomware patterns
  - *Environment Modification*: Implementing specific defensive measures
  - *Analysis Tool Execution*: Running specialized detection tools and integrating results

  - *Alert Generation*: Determining when and how to escalate to human operators

- **An LLM as a Dynamic Tool**: The RL policy would learn to "prompt engineer" an LLM to:
  - Dynamically rewrite and augment the context to highlight suspicious patterns
  - Analyze system behavior to generate useful metadata, enriching the policy's state awareness
  - Retrieve and apply relevant security knowledge from its training corpus
  - Generate explanations of detected threats for human security teams

For example, when detecting unusual file encryption activity, the policy might:

1. Query similar known ransomware behavior patterns from a vector database
2. Prompt the LLM to analyze similarities between current activity and known patterns
3. Execute specific monitoring tools to gather additional system state information
4. Decide whether to isolate affected systems, block specific processes, or alert operators

This approach offers several advantages over conventional signature-based detection methods:

- **Adaptive Detection:** Learning to identify novel attack patterns based on behavioral similarities rather than exact signatures
- **Explainable Responses:** Generating clear justifications for security actions taken, critical for incident response teams
- **Balanced Decision-Making:** Optimizing the trade-off between false positives (which cause business disruption) and missed detections (which lead to data loss)

This approach offers advantages over conventional static detection methods by adapting to novel threats and learning from each encounter.

Theoretical Framework: MDP Formulation

We can formalize the ransomware defense problem as an MDP defined by:

- **States (s)**: Each state is an embedding of the current system context, including active processes, network connections, file system activity, user behavior patterns, and augmentations from the LLM. The embedding captures the essential semantics of the environment, allowing the policy to learn from both current and historical data.
- **Actions ($a \in A$)**: The discrete action set includes defensive operations that modify the environment or analysis approach:
  - **Pattern Decomposition**: Breaking complex system behavior into analyzable components
  - **Knowledge Retrieval**: Querying a vector database of known ransomware patterns and tactics
  - **Defensive Tool Execution**: Deploying specific security measures or analysis tools
  - **Context Enhancement**: Generating additional metadata about suspicious activities
  - **Terminal Action**: Implementing a final defensive response (block, isolate, alert)

- **Transition Function**: The system evolves based on both the ransomware's behavior and the defensive actions taken
- **Reward Structure**: A combination of:
  - **Primary reward**: Successfully preventing data encryption/exfiltration
  - **Secondary rewards**: Early detection, minimal system disruption
  - **Penalties**: False positives, excessive resource utilization, business disruption

By applying Q-learning techniques to this formulation, we can derive a policy that optimizes:

$$\pi(s) = \arg\max_a Q(s, a)$$

This approach allows the system to learn increasingly sophisticated defense strategies through experience.

Theoretical Considerations and Challenges

While implementing this framework, we'll need to address several key challenges:

- **State Representation:** Ensuring embeddings capture sufficient information from system logs and network activity
- **Reward Design:** Balancing immediate detection rewards against penalties for disruption
- **Training Efficiency:** Developing methods to train the policy with limited examples of actual ransomware

These challenges present valuable research opportunities that align well with both Dr. Fujinoki's cybersecurity expertise and Dr. Gultepe's machine learning background.

Promising Extensions to Explore

Two particularly exciting avenues we could investigate as part of this research include:

- **Ransomware-Specific LLM**: Pretraining/fine-tuning the LLM on cybersecurity data to make it a "ransomware expert," significantly boosting its effectiveness within the framework.
- **Adversarial Training**: Developing the defense policy by pitting it against an AI-driven attacker in self-play scenarios. This would:
  - Push the defender to become exceptionally robust and adaptive to novel threats
  - Automatically discover vulnerabilities in the defense system
  - Generate a diverse corpus of synthetic attack scenarios for training
  - Create a red-team/blue-team dynamic that mimics real-world security operations

Summer Research Focus (3-Credit Independent Study)

I propose a 3-credit independent study course this summer that would allow me to explore this research direction under Dr. Fujinoki's guidance and with Dr. Gultepe's input. Rather than following a rigid schedule, I envision a flexible approach with the following components:

- **Literature Exploration**: Reading key papers in ransomware detection, reinforcement learning for security applications, and LLM-based decision support, as recommended by Dr. Fujinoki and Dr. Gultepe
- **Theoretical Framework Development**: Working with both professors to refine the MDP formulation for ransomware defense in ways that align with their expertise
- **Initial Experimentation**: Exploring the feasibility of core components through proof-of-concept implementations or simulations
- **Documentation and Discussion**: Regular meetings to discuss findings, challenges, and potential directions

The specific focus areas and depth of investigation would be determined collaboratively with Dr. Fujinoki and Dr. Gultepe, with regular online (Zoom) check-in meetings (e.g., biweekly), allowing us to adapt the research direction based on emerging insights and their input throughout the summer.

Expected Outcomes

By the end of the summer, I anticipate delivering:

- A well-grounded understanding of the current literature at the intersection of reinforcement learning and cybersecurity
- A refined theoretical framework for the ransomware defense MDP that incorporates feedback from both professors
- Initial proof-of-concept work exploring at least one key component of the system
- A written report documenting findings, challenges, and potential future directions, including preliminary evaluation of the approach's feasibility using criteria such as detection accuracy, explainability, and adaptability to novel attacks

These outcomes would establish a foundation upon which more substantial implementation and experimentation could be built in subsequent semesters, while generating empirical insights about the feasibility and potential of this approach.

Next Steps

I propose to meet with Dr. Fujinoki and Dr. Gultepe to discuss this research direction, obtain their input on the most promising aspects to focus on initially, and determine if this aligns with their interests and availability for summer supervision. Their expertise would be valuable in refining this concept into a productive summer research endeavor.

Conclusion

This proposed research presents a significant opportunity to integrate advanced AI techniques with practical cybersecurity challenges. By examining how reinforcement learning and large language models can enhance ransomware defense, this summer research could establish groundwork for novel approaches to adaptive security systems while contributing potentially valuable findings to my dissertation research. The interdisciplinary nature of this work, combining Dr. Fujinoki's expertise in cybersecurity with Dr. Gultepe's knowledge in machine learning, provides a solid foundation for productive investigation.

Contact and Follow-Up

I welcome any feedback or suggestions from Dr. Fujinoki, Dr. Gultepe, and Dr. Imteaj, and I'm happy to revise this plan accordingly.