M2-internship: Robust and Resource-Efficient Adaptive Key Refresh for IoT Networks using Deep Reinforcement Learning

Duration: 5 months
Supervision: Hamida SEBA (LIRIS, Lyon 1) and Lamine Messai (ERIC, Lyon 2).
Location: LIRIS, campus de la Doua Villeurbanne
Application should be sent to: Hamida.seba@univ-lyon1.fr and mohamed-lamine.messai@univ-lyon2.fr (CV+Transcripts)

Context

In resource-constrained Internet of Things (IoT) networks, maintaining secure communications requires periodic key refreshing to mitigate the impact of node compromises—a process known as self-healing. However, static refresh intervals create a critical trade-off: excessive updates deplete limited battery reserves, while infrequent updates extend the window of vulnerability for attackers.

The previous **I-Refresh** [10] framework addressed this by implementing a Deep Q-Network (DQN) agent at the Base Station to dynamically trigger key updates based on network security indicators. Despite its potential, the current prototype faces challenges regarding high communication overhead due to continuous metric reporting, a dependence on empirically tuned reward weights, and an assumption of honest metric reporting from potentially compromised nodes. This project aims to evolve I-Refresh into a production-ready framework by optimizing monitoring efficiency, hardening the system against malicious metrics (Byzantine resilience), and validating its performance against state-of-the-art adaptive security schemes.

Bibliography & Research Directions

The project will be anchored in the following theoretical pillars:

- Symmetric Key Management Self-Healing: Study of Random Key Predistribution (RKP) [4], RoK [3], and distributed schemes such as POSH [8] and POLISH [5].
- Adaptive Security Frameworks: Analysis of context-aware protocols including SKWN [6] and vector-based secret derivation (EVKMS) [1].
- Deep Reinforcement Learning (DRL): Implementation of DQN architectures [7], Bellman equation optimization, and -greedy exploration strategies for autonomous decision-making based on RL fundamentals [9].

- IoT Threat Modeling: Comprehensive analysis of node compromise patterns and attack arrival distributions (Exponential, Log-Normal, and Gamma) as described in general IoT security surveys [2].

[1] S. Bettayeb, M.-L. Messai, and S.M. Hemam. (2023). A robust and efficient vectorbased key management scheme for IoT networks. Ad Hoc Networks, 149, 103250.
[2] I. Butun, P. Österberg, and H. Song. (2020). Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures. IEEE Communications Surveys Tutorials, 22(1), 616–644.
[3] C. Castelluccia and A. Spognardi. (2007). RoK: A Robust Key Pre-Distribution Protocol for Multi-Phase Wireless Sensor Networks. SecureComm 2007, 351–360.
[4] L. Eschenauer and V. D. Gligor. (2002). A Key-Management Scheme for Distributed Sensor Networks. ACM CCS, 41–47.
[5] T. Iida, A. Miyaji, and K. Omote. (2011). Polish: proactive co-operative link selfhealing for wireless sensor networks. Symposium on Self-Stabilizing Systems, 235–267.
[6] S. Mesmoudi, B. Benadda, and A. Mesmoudi. (2019). SKWN: Smart and dynamic key management scheme for wireless sensor networks. International Journal of Communication Systems, 32(7), e3930.
[7] V. Mnih et al. (2015). Human-level control through deep reinforcement learning. Nature, 518, 529–533.
[8] R. Di Pietro, Di Ma, C. Soriente, and G. Tsudik. (2008). POSH: Proactive co- Operative Self-Healing in Unattended Wireless Sensor Networks. Symposium on Reliable Distributed Systems, 185–194.
[9] R. S. Sutton and A. G. Barto. (2020). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.
[10] Mohamed-Lamine Messai and Hamida Seba. 2016. A survey of key management schemes in multi-phase wireless sensor networks. Computer Networks 105 (2016), 60–74. https://doi.org/10.1016/j.comnet.2016.05.005

Objectives

The main goal is to transform the I-Refresh prototype into a robust industrial-grade framework through the following objectives:

1. **Reduce Communication Overhead:** Design and implement a lightweight monitoring mechanism (e.g., event-driven reporting or edge aggregation) to minimize the traffic generated by security vector transmissions.

2. **Dynamic Threat Weighting:** Replace empirical linear weights in the reward function with an automated, data-driven weighting system to ensure generalizability across heterogeneous IoT deployments.

3. **Byzantine-Resilient Monitoring:** Develop a trust-aware aggregation layer capable of filtering out false security metrics sent by compromised nodes intended to mislead the RL agent.

4. **Cryptographic Generalization:** Extend and validate the RL trigger mechanism for key management schemes beyond RKP, ensuring compatibility with hierarchical or vector-based architectures.

5. **Rigorous Benchmarking:** Establish a comprehensive evaluation suite comparing the proposed solution against periodic refresh, threshold-based heuristics, and advanced adaptive schemes like SKWN.