International Journal of Core Engineering & Management
Volume-7, Issue-04, 2022 ISSN No: 2348-9510

ML-BASED DETECTION OF CYBER ESPIONAGE ACTIVITIES TARGETING FEDERAL INSTITUTIONS

Hariprasad Sivaraman
Shiv.hariprasad@gmail.com

Abstract

Destructive or damaging cyber activity (such as denial-of-service attacks, or data destruction) was noted also, as were espionage campaigns that combined advanced tradecraft, including stealthy data extraction, social engineering and lateral movement. These kinds of threats are developing in nature, thus, traditional security measures such as rule-based intrusion detection systems (IDS) and signature-based methods provide limited protection against them. A Machine Learning (ML) based framework is proposed in this paper, for identifying and reducing the consequence of cyber espionage incidents, uniquely designed for federal institutions requirements. Supervised and unsupervised learning has been utilized for dynamic profiling of behaviors and to detect outliers which were marked as possible espionage activities in real-time. This framework can serve as the basis for sophisticated cybersecurity systems across agencies through detailed examination of the architecture including the training data requirements at each stage, and feedback loop mechanisms tied to continuous improvement.

Keywords: Cyber espionage, machine learning, anomaly detection, federal institutions, cybersecurity, network security, adversarial tactics, real-time monitoring, supervised learning, unsupervised learning, behavioural profiling.

I. INTRODUCTION

A. Overview of Cyber Espionage

Federal agencies are stewards of critical security-sensitive information. It is for this reason that they are enticing targets of cyber espionage, where attackers who may be state-sponsored attempt to secretly pull intelligence or paralyze operations. The espionage tactic may incorporate advanced social engineering, credential access and abuse, data exfiltration and other tactics that tend to be silent in nature and maybe lateral in movement making them hard to spot. Espionage, unlike many traditional cyber-attacks that are seen today, with an apparent immediate goal of damages; is sneaky and low-profile, hiding from the basic detection systems.

B. Challenges in Detecting Espionage Activities

Espionage is too sophisticated and adaptive for rule-based IDS and signature-based methods. Because espionage uses previously unknown tactics, traditional systems relying on known signatures or defined rules do not work. This results in numerous false positives, failing against new attack vectors, and analyst burnout. In addition, static defences are not strong enough against competitive adaptations. A robust defence demands a data-driven method that adjusts based on context and can identify familiar and fresh attack threats.

C. Proposed Solution

In this paper, a specific ML-based detection framework is suggested, tailored for federal institutions. The framework combines supervised learning to classify known threats with unsupervised learning to identify abnormal patterns, supporting real-time detection and ongoing learning from new data. Dynamic profiling of behaviour, dynamic threshold adjustment using a reinforcement learning feedback loop and compliance with privacy protocols through data anonymization are all part of the system. Such a theoretical solution forms the basis for introducing an adaptive defence mechanism capable of adapting to dynamic behaviour of advanced attackers in federal environment.

II. PROBLEM STATEMENT

A. Nature of Cyber Espionage Threats

Cyber espionage is the act of secretly conducting operations, such as gathering intelligence or accessing sensitive information. Common techniques include:

- Social Engineering and Phishing: The entry point in many breaches, phishing and social engineering tactics trick users into sharing credentials or clicking on malicious links.
- Lateral Movement: After gaining entry, attackers will then move laterally to escalate their privileges and access sensitive data.
- Data Exfiltration: This is an outcome in which information gets extracted secretly, often as part of encrypted outgoing traffic to prevent detection.
- Lingering Presence: Espionage groups attempt to embed themselves within the network for long-term monitoring, even operating silently and intermittently, which makes detection extremely difficult.

B. Challenges with Traditional Methods

- High False Positives: The drone of alerts from rule-based systems often raises thousands based on non-threatening anomalies, exhausting analysts.
- Insufficient Flexibility: Static defences are ineffective against espionage methods that are changing so quickly, particularly as the attackers use innovative ways to avoid detection.
- Scale and Complexity of Data: A lot of data from multiple sources flows through federal networks which makes timely detection a challenge.

C. Proposed Framework's Objectives

The ML framework aims to address these shortcomings with effective dynamic learning of new threat patterns that minimizes false positives and combines the strengths of both behavior profiling and anomaly detection. It employs a feedback loop to constantly learn and improve mass detection.

III. ML FRAMEWORK FOR DETECTION

This ML-based framework is scalable and designed as a pipeline to ingest, process, train models, detect anomalies, and integrate feedback based on large-scale data. Every part of this framework aid in the real-time detection, understanding of and management to mitigate espionage threats.

A. Data Collection and Pre processing

Data Sources: The framework aggregates data from:

- Network Logs: Access points, packet flow, initiation and termination of a session.
- Access Control Logs (o Authentication attempts, login duration, and access permission)
- Endpoint data collection: you can trace behavior across workstations, servers as well as connected devices.
- Application Logs: Tools that capture data based on interaction with mission-critical applications in the sensitive areas

Data Pre-processing:

- Normalization and Standardization: It converts aggregate data into common values from diverse sources which helps in minimizing the model complexity.
- Feature Extraction: Extracted features such as login frequency, data transfer amounts, geolocation anomalies, and a typical access pattern.
- Anonymization: Stakeholder identifiers and IP addresses are anonymized to address data governance and regulatory needs.

B. Feature Engineering

- Behaviour-Based Features: This includes the parameter through which a behavior is viewed with respect to past activity to develop a baseline known as behavioral profiling (e.g., typical login hours, session times and user-resources accessed).
- Derived Indicators: Either File Access Frequency, Unexpected Volume on Network Traffic and Geolocation deviation: These IOCs are created to detect involve Sniffing or Wiretapping.
- Contextual Features: These features are more equipped with department-based rules, and access rights provided to the model so that it can flag anomalies with more relevant

C. Model Selection and Training

- Supervised Learning Models: These models are trained on the labeled datasets having known espionage patterns and attack behaviours. Common models include:
  - Random Forests and Decision Trees: Can be beneficial in recognizing patterns within structured data
  - Support Vector Machines (SVM): Works well if there are separate hyperplane for normal and anomalous activities, especially in high dimensional space
- Unsupervised Learning Models: Used to identify new Patterns in data, without the availability of pre-labelled data. Includes clustering algorithms, K-means and the auto encoder type. By analysing deviations from expected behaviour profiles, these models detect anomalies.
- Hybrid Model Fusion: Since it uses both supervised and unsupervised, the outputs are fused into a unified anomaly score which provides balanced coverage of known and unknown threats.

D. Real-Time Anomaly Detection and Alerting

- Anomaly Scoring: A score is assigned to each event or behaviour assessed, and higher the score; higher the deviation from baseline behaviour.
- Thresholding and Prioritization: This anomaly score, when above a dynamic threshold, raises an alert; the more severe the event, the higher priority it is for analysts in order to get them to react first to high-risk activities.
- Integration with Security Information and Event Management (SIEM) Systems: Security

Orchestration Automation and Response (SOAR) feeds alerts into existing SIEM tools so that organizations can have a centralized view of their network security and effectively streamline the incident response process.

IV. TRAINING DATA REQUIREMENTS

A. Supervised Training Data:

- Signal Incidents: Past espionage-related incidents identified by providing information on tactics, user behaviours, and network anomalies.
- Simulation-based Transfers: Data that is synthetically produced using controlled simulations of espionage, including credential theft and password breach.

B. Unsupervised Training Data:

- Behavioural-data Normal: Data representing the usual activity on a network, aggregate for long periods of time.
- Network Logs: Data from network sources (unlabelled) that serves as the basis for novel espionage detection.

C. Synthetic Data:

- Rare Event Simulation: Data augmentation are then employed to create synthetic data, essentially adding a digitally born rival, aimed to boost the ability of the model or system to capture espionage behaviours that may not exist in enough frequency within ground truth data.

V. FEEDBACK LOOP FOR CONTINUOUS IMPROVEMENT

- Analyst Feedback: Flagged anomalies are verified by security analysts, tagging true positives (confirmed espionage events) and false positives (benign events), which are then fed back into the model for further training.
- Reinforcement Learning: The feedback is fed into a reinforcement learning component to update the thresholds for anomaly detection. Thresholds are raised by repeated false positives from certain patterns and reduced when espionage behaviour has been verified.
- Model Automated Retraining: Once model performance tracking starts to show drift, retraining is triggered utilizing new labelled data to adapt the espionage tactics seen in the wild.

Continuous Monitoring for Model Drift

- Drift Detection Algorithms: Drift detectors keep an eye on data distribution and model performance shifts. In simple words, if there is a drift detected then the model gets retrained based on the latest data so that it retains its accuracy.
- Adaptive learning: The feedback loop can lead to the model's improvement with new data and user feedback, meaning it would be able to adapt quickly in the face of a changing threat landscape.

VI. SOLUTION BENEFITS AND LIMITATIONS

Benefits:

- Dynamic and Accurate: The framework provides real-time learning capabilities for evolving snooping techniques while achieving scalability and accuracy with lower false-positive rates.
- Automation Minimizes Human Lab Work: Automated alerts and prioritization are resource-restrictive, as opposed to labour-intensive human verification.
- Scalability: The architecture is scalable, enabling federal institutions to add more data stores if required.

Limitations:

- Demanding Resources: Training ML and constant monitoring take a lot of computational resources.
- Compliance with privacy: Robust anonymization protocols need to be applied ensuring a balance between privacy and effective tracking.
- Model Drift: To maintain performance over time, the framework should be updated and retrained.

VII. FUTURE WORK AND RESEARCH DIRECTIONS

- Improved Interpretability: SHAP/LIME like techniques can be used authentication analyst with better understanding of decision processes of ML.
- Adversarial training: In the future, we can build more robust against evasion with adversarial training techniques for the framework.
- Cross-Institutional Collaboration: A centralized, federal repository of anonymized threat intelligence data may allow for better cross-agency detection.
- Quantum ML exploration: Quantum computing-based algorithms may be investigated for implementation of such machine learning applications to advance large-scale capabilities related to real-time threat detection across the federal enterprise.

VIII. CONCLUSION

A. Key Threats and Challenges

- **Cyber Espionage:** The most pernicious of threats posed to federal institutions, cyber espionage involves adversaries utilizing sophisticated methods to infiltrate critical information.
- **Traditional Defences Insufficient**: The stealth, sophistication, and evasiveness of espionage attacks have rendered traditional cybersecurity defenses, which are rule-based and signature-based in nature, radically insufficient.

B. Proposed Framework

- **Comprehensive Solution:** The paper proposes a comprehensive adaptive ML-based detection and response framework to identify cyber espionage targeting federal sites and effectively mitigate them.
- **Model Combination**: This framework, which combines supervised and unsupervised ML models, learns in real-time to identify known attack patterns as well as new behaviors that other detection systems might miss.

C. Framework Architecture

- **Layers of Architecture:**
  1. Data collection
  2. Pre-processing
  3. Feature engineering
  4. Model selection
- **Real-Time Monitoring:** Includes real-time monitoring with a feedback loop for continuous improvement.

D. Benefits of the Approach

- **Supervised Learning Techniques:** Capable of identifying known espionage tactics from historical and simulated data.
- **Unsupervised Learning Techniques**: Add an important dimension by detecting novel forms of espionage-related behaviour.
- **Multi-Pronged Approach**: These approaches form a low-touch, high-confidence detection framework that profiles behaviour, identifies anomalies, and escalates alerts while minimizing false positives, thereby reducing alert fatigue among security analysts.

E. Novel Features

- **Feedback Loop:** A novel component of the framework is the feedback loop, which utilizes real-time analyst feedback and reinforcement learning to adapt detection thresholds in an online fashion.
- **Dynamic Adaptation**: The adaptive feedback loop is imperative for federal institutions, where the cyber threat landscape is dynamic, and the framework should adapt rapidly to newly fashioned tactics used by opponents.

F. Flexibility and Relevance

- **Model Retention:** Retaining the model through this process allows it to remain relevant as behaviours and attack vectors shift and change.
- **Mitigating Model Drift**: Such flexibility is especially important to mitigate model drift, a well-known problem in security contexts involving dynamic environments, and highlights the long-term relevance of the framework.

G. Limitations and Future Research

- **Limitations:**
  - High computational demands and other resource requirements are a consideration.
  - Supervised models require huge datasets to be trained well.
  - Privacy compliance and data anonymization remain critical for handling highly sensitive information.
- **Future Directions:**
  - Deliver model interpretability improvements, with consideration of complex ML decisions needing to be actionable by analysts.
  - Advance adversarial training to make the framework harder to trick or bypass by attackers.
  - Foster cross-agency collaboration, establishing a joint database of anonymized threat intelligence data to improve accuracy and speed of threat detection.
  - Explore quantum-advanced ML models for real-time data ingestion on a large scale, potentially transforming national-level cyber threat detection.

H. Final Insights

- **Theoretical Solution:** This ML-based framework is a theoretical but practical solution for federal networks to prevent unwanted intrusions for the purpose of cyber espionage.
- **Cybersecurity Necessity**: It highlights the necessity for nimble, analytics-based, and scalable cybersecurity solutions that can evolve with threats.
- **Proactive Stance**: Using ML in conjunction with constant feedback allows federal institutions to take a proactive and resilient stance against cyber espionage threats.
- **Potential for Revolutionizing Capabilities**: This framework can serve as a basis for ML deployments in cybersecurity, with the potential to revolutionize federal capabilities against the constantly morphing landscape of cyber espionage.

REFERENCES

1. J. Ma, Y. Li, and D. Gao, "Application of Machine Learning Algorithms in Cybersecurity," IEEE Access, vol. 8, pp. 26131-26140, 2020. doi: 10.1109/ACCESS.2020.2967959.
2. D. E. Denning, "An Intrusion-Detection Model," IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222-232, 1987. doi: 10.1109/TSE.1987.232894.
3. N. A. Mousavi, M. N. Marsono, and R. Wong, "Behavioural Anomaly Detection of Cyber Attacks in Network Traffic Using Unsupervised Machine Learning," IEEE Transactions on Network and Service Management, vol. 18, no. 1, pp. 131-143, 2021. doi: 10.1109/TNSM.2020.3039358.
4. A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A Deep Learning Approach for Network Intrusion Detection System," Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies, New York, NY, USA, pp. 21-26, 2015. doi: 10.4108/eai.3-12-2015.2262516.
5. L. Xu, X. Wang, Z. Jin, and H. Xu, "Machine Learning-Based Anomaly Detection in Real-Time Network Traffic," IEEE Transactions on Industrial Informatics, vol. 17, no. 8, pp. 5610-5618, 2021. doi: 10.1109/TII.2021.3066783.
6. S. Dua and X. Du, Data Mining and Machine Learning in Cybersecurity. Boca Raton, FL, USA: CRC Press, 2016.
7. S. S. Shams, H. Xu, and D. I. Dawoud, "A Comprehensive Survey on Intrusion Detection Based on Machine Learning," IEEE Access, vol. 9, pp. 44470-44490, 2021. doi: 10.1109/ACCESS.2021.3065878.
8. C. Yin, Y. Zhu, S. Fei, and H. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," IEEE Access, vol. 5, pp. 21954-21961, 2017. doi: 10.1109/ACCESS.2017.2762418.
9. G. Liu, S. Teng, and J. Zhang, "Survey of Machine Learning Techniques for Intrusion Detection," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 1-20, 2021. doi: 10.1109/TNNLS.2020.2966456.
10. A. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016. doi: 10.1109/COMST.2015.2494502.
11. Y. Y. Liang, B. C. Lin, and P. Y. Chen, "Adversarial Machine Learning in Network Intrusion Detection: Recent Advances and Challenges," IEEE Network, vol. 35, no. 2, pp. 67-73, 2021. doi: 10.1109/MNET.2020.3031897.

12. B. Biggio and F. Roli, "Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning," Pattern Recognition, vol. 84, pp. 317-331, 2018. doi: 10.1016/j.patcog.2018.07.023.
13. Y. K. Sharma and D. Sharma, "A Hybrid Machine Learning Model for Anomaly Detection in Network Traffic," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 12, pp. 8011-8022, 2021. doi: 10.1109/TSMC.2020.2991657.
14. Z. Wang, T. Zhang, and J. Wang, "Adaptive Intrusion Detection of Cyber-Physical Systems with Concept Drift and Distribution Shift Using Reinforcement Learning," IEEE Access, vol. 8, pp. 2021-2032, 2021. doi: 10.1109/ACCESS.2020.3045251.
15. S. Yu, "Big Privacy: Challenges and Opportunities of Privacy Study in the Age of Big Data," IEEE Access, vol. 4, pp. 2751-2763, 2016. doi: 10.1109/ACCESS.2016.2577036.
16. P. Ferrag, L. Maglaras, and H. Janicke, "Anomaly Detection in Next-Generation Cyber Physical Systems Using Distributed Artificial Intelligence," IEEE Internet of Things Journal, vol. 8, no. 2, pp. 1146-1158, 2021. doi: 10.1109/JIOT.2020.3014613.
17. A. Zolanvari, M. A. Teixeira, and R. Jain, "Machine Learning-Based Network Intrusion Detection for Cyber-Physical Systems: A Review," IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6541-6554, 2019. doi: 10.1109/JIOT.2019.2899262.
18. N. Idika and A. P. Mathur, "A Survey of Malware Detection Techniques," IEEE Communications Surveys & Tutorials, vol. 10, no. 3, pp. 167-179, 2008. doi: 10.1109/COMST.2008.4625802.
19. P. Velarde, J. Lopez, and E. A. Ruano, "A Framework for the Evaluation of Anomaly Detection Techniques in Cyber-Physical Systems," IEEE Transactions on Cybernetics, vol. 51, no. 5, pp. 2464-2475, 2021. doi: 10.1109/TCYB.2020.2965967.
20. Y. Hu, G. Zheng, and X. Wang, "Anomaly Detection Based on Generative Adversarial Networks for Imbalanced Industrial Data," IEEE Transactions on Industrial Informatics, vol. 16, no. 4, pp. 2714-2723, 2020. doi: 10.1109/TII.2019.2943896.
21. D. Lee and W. Kim, "Adversarial Learning for Cybersecurity and Privacy Protection: Methods, Challenges, and Research Directions," IEEE Access, vol. 9, pp. 13673-13693, 2021. doi: 10.1109/ACCESS.2021.3052459.
22. M. Ring, D. Schlör, D. Landes, and A. Hotho, "Flow-Based Network Traffic Generation Using Generative Adversarial Networks," IEEE Access, vol. 7, pp. 170139-170153, 2019. doi: 10.1109/ACCESS.2019.2955461.
23. B. Liu, Y. Liu, and L. Zhao, "Interpretable Machine Learning Models for Cybersecurity in the Internet of Things," IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6275-6284, 2020. doi: 10.1109/JIOT.2019.2960846.
24. D. Wang, W. Li, and C. Li, "A Distributed and Privacy-Preserving Anomaly Detection Framework Based on Machine Learning in Fog Computing Environments," IEEE Access, vol. 8, pp. 209573-209583, 2020. doi: 10.1109/ACCESS.2020.3037385.
25. F. S. Gharehbaghi and M. Safa, "Machine Learning and Cybersecurity Threat Intelligence," IEEE Access, vol. 9, pp. 11415-11429, 2021. doi: 10.1109/ACCESS.2021.3051119.