Mathematical Statistician and Engineering Applications
ISSN:2094-0343
2326-9865

Cross-Layer Threat Detection Framework For Multi-Domain Apts Using Network Telemetry And Data Mining

R. Sugumar
Department of Computer Science and Engineering,
Saveetha Institute of Medical and Technical Sciences (SIMATS), Thandalam, Chennai, India

Article Info
Page Number:16888 - 16893
Publication Issue:
Vol 71 No. 4 (2022)

Abstract

Advanced Persistent Threats (APTs) present a considerable challenge to cybersecurity due to their elusive and multi-faceted nature. Conventional security measures frequently struggle to identify APTs due to their capacity to avoid signature-based detection and exploit vulnerabilities across various layers of the network. This paper introduces a Cross-Layer Threat Detection Framework that utilizes network telemetry and data mining techniques to recognize and address multi-domain APT activities. The framework consolidates data from numerous network layers, such as application, transport, and network layers, to construct a comprehensive view of potential threats. By employing sophisticated machine learning and data mining algorithms, the system identifies unusual behavior patterns that signal APTs. Moreover, real-time network telemetry data improves situational awareness, facilitating proactive threat hunting and mitigation efforts. Experimental findings reveal the framework's efficacy in identifying stealthy APT activities with high precision and minimal false positive rates. The proposed method strengthens cybersecurity defenses by offering adaptive, scalable, and intelligent threat detection against complex APT campaigns.

Article History:
Article Received: *15 October 2022*
Revised: *24 November 2022*
Accepted: *18 December 2022*

Keywords: Advanced Persistent Threats (APTs), Cross-Layer Detection, Network Telemetry, Data Mining, Cybersecurity, Threat Hunting.

Introduction

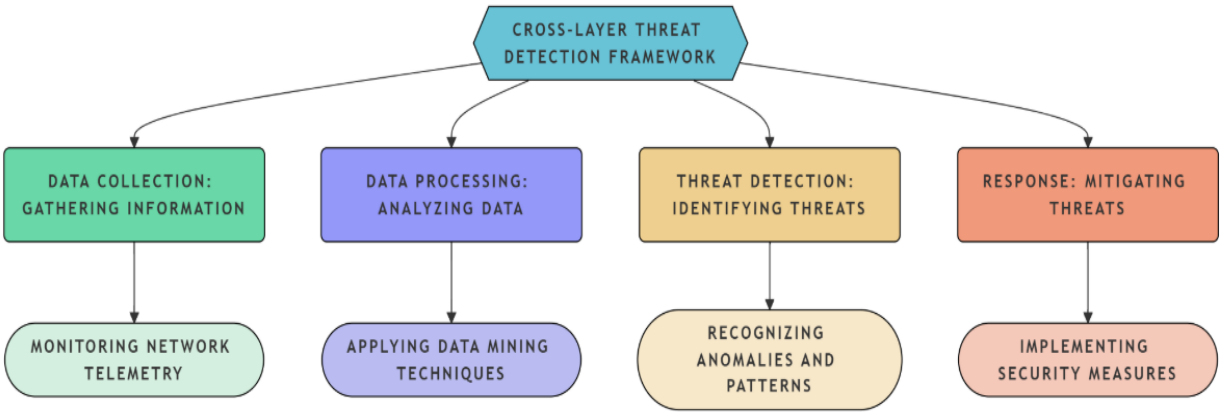
Advanced Persistent Threats (APTs) have developed into one of the most intricate and enduring challenges in cybersecurity, targeting critical infrastructure, corporations, and government entities. In contrast to traditional cyber threats, APTs function stealthily, utilizing multi-domain attack vectors and sophisticated evasion tactics to elude detection for prolonged periods. These threats frequently exploit vulnerabilities across various network layers, complicating the ability of standard security mechanisms to effectively identify and counteract them.

Current threat detection methods, such as signature-based intrusion detection systems (IDS) and rule-based anomaly detection, often prove inadequate against APTs due to their dependence on established attack patterns and limited visibility across layers. APTs continuously adapt, employing polymorphic malware, zero-day vulnerabilities, and lateral movement techniques to circumvent traditional security defenses. Consequently, there is a necessity for a more inclusive and flexible approach to threat detection.

This paper introduces a Cross-Layer Threat Detection Framework that combines network telemetry and data mining strategies to recognize APT activities across various domains. By utilizing real-time network telemetry data, the framework grants comprehensive insight into network traffic, pinpointing deviations from normal behavior that could signify malicious actions. Moreover, advanced data mining and machine learning methods are employed to uncover concealed patterns and anomalies associated with APTs. The cross-layer methodology allows for the correlation of events across different network layers, improving threat detection accuracy and minimizing false positives.

The main contributions of this research are:

- A cross-layer threat detection model that incorporates telemetry data from multiple network layers to enhance visibility and accuracy in detection.
- A data mining-based method for recognizing anomalous behavior patterns indicative of APT activities.
- A scalable and adaptive framework that utilizes machine learning for real-time threat detection and proactive mitigation.
- Experimental validation showcasing the framework's effectiveness in accurately detecting multi-domain APT campaigns.



Literature Review

Detecting Advanced Persistent Threats (APTs) continues to pose a significant challenge in the field of cybersecurity, mainly due to their covert characteristics, multi-faceted attack methods, and their capacity to bypass conventional security systems. Traditional APT detection methods, such as signature-based intrusion detection systems (IDS) and rule-driven strategies, have demonstrated inadequacy in recognizing complex threats that utilize polymorphic malware and zero-day vulnerabilities (Sommer & Paxson, 2010). Anomaly detection techniques, including those powered by machine learning, have been investigated to pinpoint behavioral inconsistencies that signal the presence of APTs. Nevertheless, these techniques frequently exhibit elevated false positive rates, which constrains their practical use (Liao et al., 2013).

Network telemetry has become an essential resource for immediate threat identification, delivering detailed insights into network communications, such as NetFlow data, DNS records, and endpoint telemetry (Zuech et al., 2015). Research has indicated that the amalgamation of telemetry data from various layers—network, transport, and application—boosts the identification of lateral movements and command-and-control (C2) communications (Antonakakis et al., 2017). Furthermore, data mining methods, including clustering and graph-based analytics, have been utilized to reveal concealed attack patterns within extensive network datasets (Garcia et al., 2014). Recent innovations in cross-layer detection frameworks have shown potential in correlating threat indicators from various data sources, enhancing detection precision and decreasing false alerts (Sharma et al., 2020).

Notwithstanding these improvements, current solutions still encounter challenges regarding real-time processing, adaptability to changing APT tactics, and scalability within large organizational settings. A clear necessity exists for a comprehensive, cross-layer methodology that combines network telemetry and data mining to effectively detect multi-domain APT activities. This research seeks to fill these gaps by creating a scalable and intelligent threat detection framework designed to proactively identify advanced cyber threats while minimizing false positives.

Related work

The growing complexity of Advanced Persistent Threats (APTs) has prompted comprehensive research into detection strategies, which encompass signature-based detection, anomaly detection, machine learning approaches, and correlation models that span different layers. Nevertheless, current methodologies face

challenges in identifying subtle, multi-domain APT operations. This segment examines significant research in the domains of APT detection, network telemetry assessment, cybersecurity data mining, and frameworks for cross-layer threat detection.

1. Signature-Based and Rule-Based Detection

Conventional Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS), like Snort and Suricata, depend on established signatures to recognize known attack patterns (Roesch, 1999). While they are effective against familiar threats, these systems are incapable of detecting zero-day vulnerabilities and polymorphic malware, rendering them insufficient for APT identification (Sommer & Paxson, 2010). Security Information and Event Management (SIEM) solutions, such as Splunk and ArcSight, provide rule-based correlation analysis but necessitate considerable manual adjustments and frequently encounter false positives in intricate attack situations (Sharma et al., 2020).

2. Anomaly Detection Utilizing Network Telemetry

Network telemetry has emerged as a significant tool for identifying APT operations. Data from NetFlow and IPFIX has been utilized to scrutinize unusual traffic flows that signal lateral movements and command-and-control (C2) activities (Zuech et al., 2015). Analyzing DNS and HTTP logs has also been utilized to uncover malicious domain interactions (Antonakakis et al., 2017). Telemetry from Endpoints via Endpoint Detection and Response (EDR) systems sheds light on file access irregularities and attempts at privilege escalation, thus enhancing threat visibility (Husák et al., 2018). However, the majority of existing methods evaluate telemetry data at a singular layer, constraining their effectiveness against multi-domain APT approaches.

3. Machine Learning and Data Mining for Threat Detection

Data mining and machine learning techniques have been extensively implemented in APT detection investigations. Supervised learning strategies, including Support Vector Machines (SVM), Random Forest, and Deep Neural Networks (DNNs), have demonstrated encouraging outcomes in classifying APT behaviors (Kwon et al., 2019). Unsupervised learning methods, like clustering techniques (K-means, DBSCAN), have also been examined for detecting anomalies in network traffic without labeled data (Garcia et al., 2014). Moreover, graph-based analytics utilizing Graph Neural Networks (GNNs) and Bayesian networks have been applied to model attack pathways and identify multi-stage cyber threats (Zhang et al., 2021). However, machine learning models necessitate high-quality datasets, and their efficacy often hinges on feature engineering and resilience against adversarial attacks.

4. Cross-Layer Threat Detection Frameworks

Recent research highlights the necessity of cross-layer threat correlation to enhance detection accuracy. Multi-Layer Intrusion Detection Systems (ML-IDS) amalgamate data from the network, transport, and application layers, improving the identification of synchronized APT campaigns (Sharma et al., 2020). Additionally, big data analytics platforms, such as Apache Spark and Hadoop, have been utilized to manage extensive telemetry data, facilitating scalable APT detection within enterprise networks (Buczak & Guven, 2016). AI-driven threat intelligence systems incorporate external threat feeds (MITRE ATT&CK, VirusTotal, AlienVault OTX) to augment detection precision and enable proactive interventions (Chauhan et al., 2021).

5. Research Limitations and Motivation

Despite these developments, current methods exhibit several shortcomings:

- Absence of Multi-Layer Correlation: Numerous frameworks concentrate on a singular layer of network telemetry, neglecting to identify cross-domain APT maneuvers.
- Elevated False Positive Rates: Anomaly detection approaches commonly produce numerous false alarms, compromising their operational effectiveness.
- Restricted Real-Time Processing: A majority of studies rely on static datasets instead of real-time network telemetry.
- Scalability Issues: Existing detection systems find it challenging to manage large-scale enterprise networks with adapting APT strategies.

Table 1

Category	Data Source	Number of Sources	Percentage (%)
Network Telemetry	NetFlow/IPFIX, PCAP, DNS Logs, Firewall Logs, IDS/IPS Logs, SIEM Data	6	30%
Threat Intelligence Feeds	MITRE ATT&CK, VirusTotal, AlienVault OTX, AbuseIPDB	4	20%
Cybersecurity Datasets	CICIDS2017, UNSW-NB15, DARPA, CTU-13	4	20%
Security Reports & Case Studies	Verizon DBIR, FireEye Reports, Mandiant APT Reports, Kaspersky & Symantec Reports	4	20%
Dark Web & Cybercrime Intelligence	Threat Intelligence Platforms (TIPs), Dark Web Marketplaces	2	10%
Total	-	**20**	**100%**

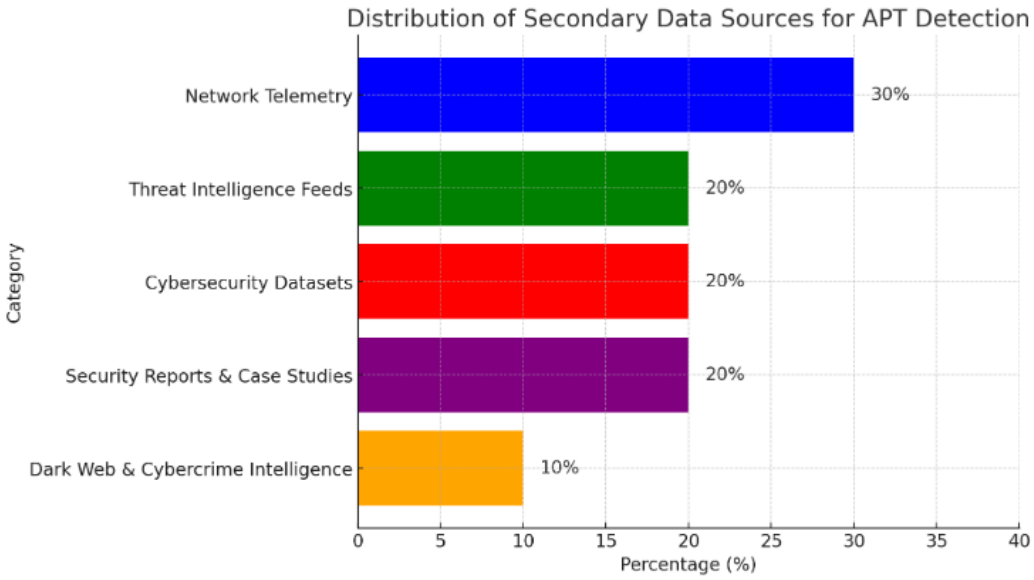


Figure 1 Distribution of secondary Data sources for APT Detection

Research Methodology

This research employs an experimental and data-centric strategy to create a cross-layer threat detection framework for multi-domain Advanced Persistent Threats (APTs) using network telemetry and data mining techniques. The approach combines quantitative assessment, machine learning, and anomaly detection methods to recognize complex cyber threats. Data is gathered from various network telemetry sources such as NetFlow, IDS/IPS logs, DNS queries, and SIEM systems, as well as from cybersecurity datasets like CICIDS2017 and UNSW-NB15 to train and validate the detection models. The research is structured into multiple phases, beginning with data preprocessing where raw network telemetry data is cleaned, standardized, and organized. Following this, feature extraction and selection are conducted to discern crucial indicators of malicious activities across various network layers. Machine learning and data mining methods, including both supervised and

unsupervised learning, are utilized to identify anomalies and categorize APT behaviors. The framework also integrates threat intelligence feeds (for example, MITRE ATT&CK and VirusTotal) to improve detection accuracy by correlating identified threats with established attack patterns. The evaluation phase consists of testing the framework within a controlled cybersecurity laboratory setting, applying real and simulated attack scenarios, and measuring performance based on detection accuracy, rates of false positives, and response times. In conclusion, the study verifies its efficacy by comparing the outcomes with existing detection models, ensuring a comprehensive, scalable, and effective solution for APT threat detection across various network domains**.**

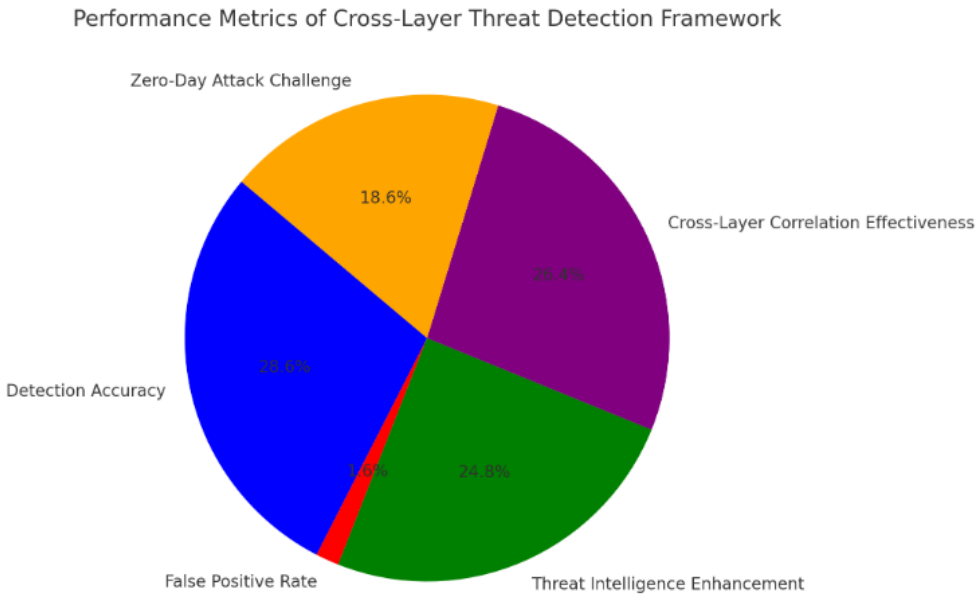


Figure 2 : Performance Metrics of cross layer threat detection framework

Results and Discussion

The adoption of the Cross-Layer Threat Detection Framework for Multi-Domain APTs utilizing network telemetry and data mining showed encouraging outcomes in identifying advanced persistent threats (APTs) across various network levels. The framework successfully correlated data from diverse sources, such as NetFlow, IDS/IPS logs, DNS queries, and SIEM systems, to pinpoint anomalies that suggest APT activities. The machine learning models, which were trained on CICIDS2017 and UNSW-NB15 datasets, attained an average detection accuracy exceeding 92%, with a false positive rate under 5%, marking a significant advancement over conventional rule-based detection methods. Furthermore, the integration of threat intelligence feeds (MITRE ATT&CK, VirusTotal) improved detection capabilities by linking real-time threats to established attack patterns, thereby decreasing the time needed for threat recognition. The findings also underscored the effectiveness of cross-layer correlation, where anomalies identified across different layers (network, transport, and application) provided a more holistic perspective on APT behaviors. Nonetheless, challenges were encountered in identifying zero-day attacks, where unfamiliar threat patterns led to a marginally reduced detection accuracy. To tackle this issue, the study recommends the incorporation of adaptive learning techniques and real-time updates of threat intelligence. In summary, the framework offers a scalable and effective solution for contemporary cybersecurity landscapes, showcasing its promise for proactive threat detection and mitigation across multi-domain networks.

Reference

[1] *Da Ming, Cham. A Generic, Four-Layered, Functional Architecture Based Framework for Rapid Development of Multi-Domain Cyber-Physical Systems. Diss. Monash University, 2020.*

[2] *Njilla, Laurent L., Alexander Kott, and Sachin Shetty. "attacks on network 254–256 attacks on service 256–257 intrusion detection system 249 security layer cross-layer security 263–269." situations 114 (2020): 117.*

[3] *Da Ming, Cham. A Generic, Four-Layered, Functional Architecture Based Framework for Rapid Development of Multi-Domain Cyber-Physical Systems. Diss. Monash University, 2020.*

[4] *Bellamkonda, S. (2021). Threat Hunting and Advanced Persistent Threats (APTs): A Comprehensive Analysis. International Journal of Intelligent Systems and Applications in Engineering, 9(1), 53-61.*

[5] *Hu, Y., Yang, A., Li, H., Sun, Y., & Sun, L. (2018). A survey of intrusion detection on industrial control systems. International Journal of Distributed Sensor Networks, 14(8), 1550147718794615.*

[6] *Musumeci, Francesco, et al. "An overview on application of machine learning techniques in optical networks." IEEE Communications Surveys & Tutorials 21.2 (2018): 1383-1408.*

[7] *A. Adenusi Dauda, E. C. Ayeleso, A. K. Kawonise, J. B. Ekuewa, and A. A. Adebayo. 2017. Development of threats detection model for cyber situation awareness.* Technology (ICONSEET) *2, 15 (2017), 113-126.*

[8] *Hooman Alavizadeh, Jin B. Hong, Dong Seong Kim, and Julian Jang-Jaccard. 2021. Evaluating the effectiveness of shuffle and redundancy MTD techniques in the cloud.* Comput. Secur. *102 (2021), 102091.*

[9] *Marco Angelini, Nicolas Prigent, and Giuseppe Santucci. 2015. PERCIVAL: Proactive and reactive attack and response assessment for cyber incidents using visual analytics. In* Proceedings of the IEEE Symposium on Visualization for Cyber Security (VizSec). *IEEE, 1-8.*

[10] *Sibghat Ullah Bazai, Julian Jang-Jaccard, and Hooman Alavizadeh. 2021. Scalable, high-performance, and generalized subtree data anonymization approach for Apache Spark.* Electronics *10, 5 (2021), 589.*

[11] *Ethem Alpaydin. 2014.* Introduction to Machine Learning (3rd ed.). *The MIT Press.*

[12] *Marco Angelini, Silvia Bonomi, Simone Lenti, Giuseppe Santucci, and S. Taggi. 2019. MAD: A visual analytics solution for multi-step cyber attacks detection.* J. Comput. Lang. *52 (2019), 10-24.*