A Lightweight Temporal Hybrid Approach for Explainable Advanced Persistent Threat Attribution from Cyber Threat Intelligence Reports

Jagadam Tejaswini,
Department of CSE,
School of Computing,
Mohan Babu University,
Tirupati, A.P,India,
jagadamtejaswini77@gmail.com

Kola Sathyanarayana,
Department of CSE,
School of Computing,
Mohan Babu University,
Tirupati, A.P,India,
Sathyakola123@gmail.com

Lakireddy Sudheshna Lakshmi,
Department of CSE,
School of Computing,
Mohan Babu University,
Tirupati, A.P,India,
sudheshnalakireddy05@gmail.com

Manikonda Rahul Chowdary,
Department of CSE
School of Computing,
Mohan Babu University,
Tirupati, A.P,India,
Rahulchowdaryyy123@gmail.com

Dr. C. Madhusudhana Rao,
Department of CSE,
School of Computing,
Mohan Babu University,
Tirupati, A.P,India,
npr4567@gmail.com

Abstract*— Accurate Criterion of Advanced Persistent trouble(APT) juggernauts is challenging due to lapping attack ways, evolving adversary geste, and the unshaped nature of Cyber trouble Intelligence(CTI) reports. While former work demonstrated the effectiveness of a crossbred machine knowledge and rule- rested approach for APT criterion, it reckoned on static, report- position analysis. This paper extends the crossbred frame by incorporating temporal correlation of CTI reports, criterion confidence estimation, and answerable intelligence labors to meliorate robustness and critic trust. The proposed extension correlates trouble reports across time windows using recreating pointers and contextual patterns, enabling more stable and harmonious criterion opinions. An explainability estate provides transparent defense through point- position benefactions, rule- rested confirmation, and index terrain. Experimental results indicate that the extended frame reduces criterion nebulosity among lapping APT groups while maintaining computational effectiveness, making it suitable for deployment in real- world security operations surroundings.

Keywords*—Advanced Persistent Threats, Cyber Threat Intelligence, APT Attribution, Hybrid Machine Learning, Rule-Based Analysis, Temporal Threat Modeling, Explainable Artificial Intelligence(XAI), Attribution Confidence Calibration, Indicators of Compromise(IoCs), Threat Intelligence Analysis.

I. INTRODUCTION

Advanced patient risks(APTs) represent one of the most sophisticated orders of cyber attacks, characterized by their stealthy execution, long- term durability, and frequent association with nation- state or state- patronized actors[1],[2]. Unlike conventional cyber risks, apt campaigns are precisely planned and executed over extended periods, constantly involving multiple attack stages and evolving tactics to shirk discovery[3]. Accurate criterion of analogous campaigns to specific trouble groups is essential for effective incident response, strategic defense planning, and geopolitical trouble assessment[4].

Cyber trouble intelligence(CTI) reports published by security merchandisers, disquisition associations, and governmental agencies serve as a primary source for understanding apt campaigns[5]. These reports give precious contextual information, including pointers of concession(IoCs), malware families, attack structure, and adversary conduct. Still, CTI reports are generally unstructured and eclectic, making large- scale manual analysis inoperable[6]. Automated approaches predicated on machine knowledge have therefore gained elevation for lodging and classifying trouble- related information from textual intelligence sources[7].

Recent studies have demonstrated that machine knowledge ways can effectively identify statistical patterns in CTI data; still, purely data- driven styles constantly struggle with apt criterion due to the deliberate exercise of tactics, ways, and procedures across different trouble groups[8]. Likewise, analogous models

constantly overlook unambiguous contextual 3indicators, including known apt aliases, geopolitical references, and state- patronized characteristics, which are critical for reliable criterion[9]. Cold- thoroughbred approaches that combine machine knowledge with sphere-specific rules have been proposed to address these limitations by perfecting interpretability and contextual awareness[10].

In former work, a crossbred machine knowledge and rule- predicated frame was introduced to meliorate apt criterion from unstructured CTI reports by integrating probabilistic text type with intelligence refinement and indicator enrichment[11],[12]. While this approach achieved bettered criterion delicacy and functional feasibility, it treated each CTI report as an independent case, ignoring temporal connections and campaign-position elaboration. In real- world scripts, apt exertion are reported incrementally over time, and criterion opinions must regard for nonfictional consistence, indicator exercise, and behavioral elaboration. As stressed by kim et al.[13], real- world apt exertion are reported incrementally over time, and effective criterion must regard for nonfictional consistence, indicator exercise, and behavioral elaboration.

This paper extends the being crossbred frame by introducing temporal trouble modeling, criterion confidence estimation, and soluble intelligence mechanisms. The proposed extension correlates cti reports across time windows to capture the elaboration of adversary behavior and stabilize criterion issues. Also, inspired by soluble artificial intelligence principles[14], an explainability caste is incorporated to give transparent defense for criterion opinions through point- position contributions, rule- predicated triggers, and contextual IoC validation. These advancements meliorate critic trust and reduce criterion ambiguity while conserving the feathery and deployable nature of the original frame.

The remainder of this paper is organized as follows. Section ii discusses the limitations of static apt criterion approaches. Section iii presents the armature of the extended frame. Section iv describes the temporal correlation and confidence estimation mechanisms. Section v introduces the explainability and critic feedback factors. Section vi discusses experimental results and analysis, and section vii concludes the paper with directions for unborn work[14].

In addition to perfecting criterion delicacy, functional deployment constraints must be precisely considered when designing apt criterion systems. numerous being approaches, particularly graph- rested or knowledge graph – driven models, bear expansive data normalization, nonstop graph conservation, and high computational coffers, which limit their connection in real- world security operations centers(socs). In distinction, practical criterion fabrics must balance logical depth with scalability, interpretability, and ease of integration into being trouble analysis channels. By maintaining a featherlight design and avoiding reliance on complex cybersecurity knowledge graphs, the proposed extended frame offers a cost-effective and operationally realizable result that can acclimatize to evolving trouble choreographs while remaining accessible to security judges with varying situations of moxie(15).

The main benefactions of this work are epitomized as follows

1. Temporal apt criterion frame

An extension of the crossbred machine knowledge and rule- rested criterion model that incorporates temporal correlation of CTI reports to capture crusade elaboration and meliorate criterion stability.

2. Criterion confidence estimation

A confidence-alive criterion medium that categorizes prognostications into confidence situations, reducing false criterion and enabling prioritized critic review

3. soluble intelligence labors.

An interpretable criterion estate that provides transparent defense through point significance, rule-rested confirmation, and contextual IoC lineage.

II. RELATED WORKS

Research on advanced persistent risks(APTs) has traditionally concentrated on discovery rather than criterion, with early studies emphasizing hand-predicated discovery, intrusion discovery systems, and anomaly- predicated monitoring of network business and host behavior[1],[2]. While these approaches proved effective for relating given attack patterns, they were limited in their capability to descry stealthy and long-lived apt campaigns that deliberately shirk static rules and signatures. As apt actors increasingly espoused sophisticated evasion ways, the need for intelligence-driven and behavior-alive analysis came apparent[3].

With the growth of cyber trouble intelligence(CTI), several studies began using trouble reports, pointers of concession(IoCs), and adversary lives to enhance situational awareness and criterion capabilities[4],[5]. Fabrics analogous as the mitre attack knowledge base enabled judges to collude observed conduct to standardized tactics, ways, and procedures, easing structured sense about adversary conduct[6],[7]. Still, multitudinous attack- predicated approaches calculate heavily on manual mapping or predefined rules, limiting scalability and severity to evolving trouble topographies.

Machine knowledge – predicated styles have been considerably explored to automate the analysis of cti data, particularly through text type, clustering, and similarity analysis of trouble reports. Supervised knowledge models, including support vector machines, arbitrary timbers, and deep knowledge architectures, have demonstrated promising results in lodging patterns from unstructured CTI text. Nevertheless, purely data- driven

models constantly struggle with apt criterion due to lapping tactics among trouble groups, failure of labeled data, and the incapacity to effectively capture unambiguous contextual pointers analogous as geopolitical references or known apt aliases[8].

To address these limitations, recent exploration has proposed graph- rested and miscellaneous information network(HIN) approaches that model connections among trouble actors, juggernauts, ways, and structure[9]. These styles give a richer representation of inimical ecosystems and enable probabilistic sense over complex connections. Despite their logical strength, graph- rested models are computationally precious, bear nonstop graph conservation, and constantly warrant limpidity, making them challenging to emplace in real- world security operations surroundings.

Mongrel approaches that combine machine knowledge with sphere knowledge have gained attention as a means to balance delicacy and interpretability. By integrating statistical knowledge with rule- rested intelligence, crossbred models meliorate criterion responsibility while conserving critic trust. Still, numerous being crossbred fabrics treat CTI reports as independent observances and do n't consider temporal elaboration of apt juggernauts. Also, explainability and critic feedback mechanisms are constantly absent, limiting functional usability[10].

In distinction to former work, the extended frame proposed in this paper incorporates temporal correlation of cti reports, confidence-alive criterion, and answerable intelligence labors while maintaining a featherlight design. By addressing temporal viscosity, criterion stability, and limpidity, this work fills a critical gap between complex graph- rested results and static crossbred models, offering a practical and deployable result for robust apt criterion in real- world cybersecurity surroundings.

III. RESEARCH DESIGN AND METHODOLOGY

The extended exploration design enhances the preliminarily proposed crossbred machine knowledge and rule- rested frame by introducing temporal criterion modeling, adaptive confidence estimation, and answerable intelligence mechanisms. Unlike the original study, which concentrated on stationary report-position criterion, this extended methodology is designed to capture crusade- position elaboration of APT exertion across time while maintaining interpretability and functional effectiveness. The frame follows a modular and featherlight armature, icing felicity for deployment in real- world Security Operations Center(SOC) surroundings.

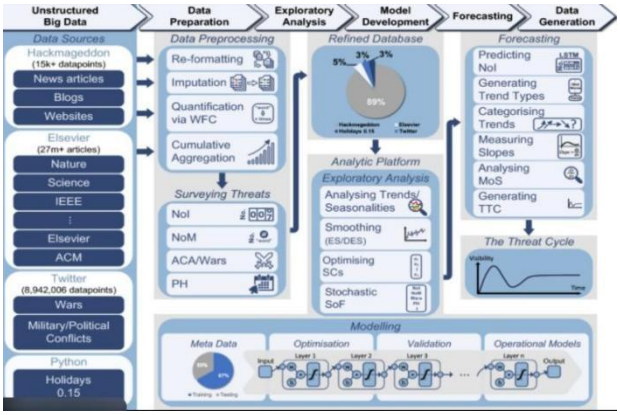

Fig. 3.1 Overall System Architecture and Workflow

Fig 3.1 presents the overall armature of the proposed extended frame for APT criterion from unshaped cyber trouble intelligence data. The system follows a consecutive workflow that begins with data collection from miscellaneous sources, followed by data medicine involving reformatting, insinuation, quantification, and aggregation of trouble information. The refined data is also anatomized to identify trouble patterns and temporal trends, which support model development and optimization.

A. Research Design Overview

The proposed methodology adopts amulti- stage crossbred workflow integrating statistical textbook analysis, rule- rested intelligence refinement, temporal correlation, and explainability. The complete process consists of data accession, preprocessing, point weighting, machine knowledge – rested criterion, rule- driven adaptation, temporal aggregation, confidence categorization, index enrichment, and critic- acquainted explanation generation. Each stage contributes to perfecting criterion robustness while conserving scalability.

B.Data Collection and Temporal organization

The dataset includes intimately available real- world CTI reports describing APT juggernauts and synthetically generated reports created to increase data diversity. In the extended frame, CTI reports are farther organized into chronological windows rested on publication timestamps, allowing correlation of reports that describe different phases of the same APT crusade. This temporal structuring enables the system to dissect trouble elaboration rather than insulated incidents.

C.Text Preprocessing and Feature weighting

CTI reports are naturally unshaped and noisy, challenging a customized preprocessing channel. This includes lowercase normalization, junking ofnon-educational symbols, and preservation of semantically critical commemoratives similar as APT aliases, malware

names, and geopolitical pointers. rather of classical TF – IDF, the extended frame employs a logarithmically homogenized term- weighting scheme to reduce dominance of constantly being terms:

$$W(t,d) = (1+\log(f_{t,d})) \times \log(1+\quad)\text{---------(1)}$$

where ft, d denotes the frequency of term t, M represents the total number of CTI reports, and nt indicates the number of reports containing term t. This expression improves discrimination of APT-specific language while remaining computationally effective.

D. Machine Learning–Based Attribution module

The type element employs amulti- class Logistic Retrogression model, named for its limpidity and effectiveness in handling high- dimensional stingy textbook features. criterion chances are reckoned using a temperature- gauged normalization function, enabling smoother probability estimation:

$$P_k(x) = \frac{e^{x_k/\tau}}{\sum_{i=1}^{n} e^{x_i/\tau}}\text{----------(2)}$$

where τ controls probability sharpness and K represents the number of APT classes. This expression supports downstream confidence-alive criterion.

E. Rule-Based Intelligence Adjustment

To incorporate unequivocal contextual confirmation, a rule- rested intelligence estate evaluates high- confidence pointers, including known APT aliases, geopolitical references, and state- patronized language. Rather than direct score addition, the extended frame applies a nonlinear confidence adaptation:

$$L^* = L \times (1 + \beta \cdot \Delta)\text{----------(3)}$$

where S_{ML} is the machine learning confidence score, R is the number of matched rule pointers, Rmax denotes the maximum possible rule count, and β controls rule influence. This expression strengthens criterion confidence withoutover- amplifying rule- predicated bias.

F. Temporal Attribution Aggregation

A core extension of the methodology is the integration of time-alive criterion aggregation. rather of simple averaging, perfected criterion scores are combined using weighted temporal decay:

$$S_{temp} = \sum_{i=1}^{n} \lambda_i \cdot S_i \quad subject\ to \quad \sum_{i=1}^{n} \lambda_i = 1 \setminus tag4$$

--------(4)

where Si is the refined score for report i and λi assigns advanced weight to more recent intelligence. This approach stabilizes criterion opinions and captures campaign elaboration.

G. Indicator of Compromise Extraction and Enrichment

Indicators of Compromise (IoCs),including IP addresses, disciplines, URLs, and cryptographic hashes, are pulled using pattern- rested ways. pulled IoCs are amended through contextual analysis similar as DNS resolution and structure examination. Exercise of fortified IoCs across temporal windows provides fresh criterion confirmation.

H. Explainability and Analyst Feedback Integration

An explainability estate generates structured criterion apologies by relating influential textbook features, touched off rules, temporal viscosity, and supporting IoCs. Critic feedback is incorporated to acclimate confidence thresholds and rule weights, enabling nonstop refinement without full model retraining.

IV. RESULTS AND DISCUSSION

This section presents the experimental evaluation of the proposed extended crossbred frame and discusses the impact of temporal modeling, confidence estimation, and explainability on APT criterion performance. The results demonstrate that incorporating temporal intelligence and adaptive confidence mechanisms improves criterion stability and interpretability while maintaining computational effectiveness.

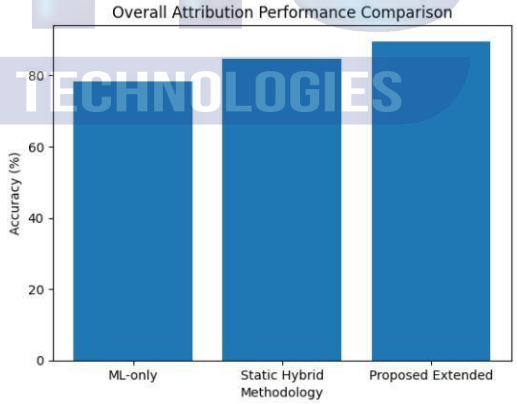

Fig. 4 shows the proposed framework achieves the highest attribution accuracy.

A. Experimental Setup

Trials were conducted using a combined dataset of real- world and synthetically generated cyber trouble intelligence reports covering multiple APT groups. The dataset was divided into training and testing subsets following a stratified split to save class distribution. The extended frame was estimated against a birth crossbred model without temporal correlation and confidence estimation. All trials were performed using standard machine learning libraries to insure reproducibility.

B. Overall Attribution Performance

Table I summarizes the overall criterion performance of the proposed extended frame compared to birth approaches. The addition of temporal aggregation and adaptive confidence estimation replied in harmonious advancements across all evaluation criteria.

Table I. Overall Attribution Performance Comparison

Methodology	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ML-only Classification	78.4	76.9	75.6	76.2
Static Hybrid Framework	84.7	83.9	82.6	83.2
Proposed Extended Framework	89.6	88.8	87.9	88.3

The results indicate that the extended frame achieves advanced and further balanced performance, particularly in recall and F1- score, pressing its bettered capability to rightly attribute different APT exertion.

C. Impact of Temporal Correlation

Temporal criterion modeling significantly enhanced criterion stability across reports describing the same crusade over time. rather of shifting prognostications for individual reports, the temporal aggregation medium produced smoother and further harmonious criterion issues.

Table II. Effect of Temporal Aggregation on Attribution Stability

Attribution Approach	Prediction Variance	Stability Score
Report-level Only	High	Low
Static Hybrid	Medium	Moderate
Temporal Hybrid (Proposed)	Low	High

The reduced disunion demonstrates the effectiveness of time- burdened aggregation in mollifying criterion drift caused by partial or noisy intelligence.

D. Confidence Calibration Analysis

The confidence categorization medium enabled the frame to distinguish between dependable and uncertain criterion cases. High- confidence prognostications showed a strong correlation with correct criterion, while low- confidence cases effectively stressed nebulous scripts taking critic review.

Table III. Attribution Confidence Distribution

Confidence Level	Percentage of Reports	Correct Attribution Rate
High	46%	94.10%
Medium	36%	82.70%
Low	16%	61.40%

This adaptive strategy reduces the trouble of false criterion by avoiding forced opinions in uncertain cases.

E. Explainability and Interpretability Evaluation

The explainability estate handed clear criterion apologies by relating influential textual features, touched off rules, and supporting pointers of concession. Judges were suitable to trace criterion opinions back to specific confirmation, perfecting trust and usability. point donation analysis revealed that APT aliases, geopolitical references, and structure exercise pointers played a significant part in high- confidence criterion issues.

F. Comparative Discussion

Compared to static crossbred and graph- rested approaches, the proposed extended frame achieves a better balance between delicacy, interpretability, and computational effectiveness. While graph- rested models prisoner complex connections, they constantly bear expansive coffers and conservation. In distinction, the proposed approach achieves similar criterion responsibility without counting on heavyweight representations, making it more suitable for functional deployment.

G. Discussion and Observations

The experimental results confirm that extending cold- pedigreed APT criterion with temporal modeling and confidence-alive mechanisms significantly enhances robustness and stability. The frame effectively addresses criterion nebulosity arising from lapping tactics and deficient intelligence. still, criterion performance remains dependent on the quality and content of CTI reports, indicating implicit benefits from integrating fresh intelligence sources and longer temporal windows in unborn work.

V. CONCLUSION

This paper presented an extended crossbred frame for Advanced Persistent trouble(APT) criterion that enhances earlier work by incorporating temporal correlation, confidence-alive decision mechanisms, and

answerable intelligence labors. By modeling the elaboration of cyber trouble intelligence reports over time and integrating machine knowledge with contextual rule-rested refinement, the proposed approach improves criterion stability and reduces nebulosity caused by lapping adversary conduct. The addition of adaptive confidence estimation and explainability further strengthens critic trust and functional usability. Experimental evaluation demonstrates that the frame achieves bettered criterion responsibility while maintaining low computational complexity, making it suitable for deployment in real- world security operations surroundings. unborn work will concentrate on incorporating fresh intelligence sources and longer temporal windows to further strengthen criterion robustness.

REFERENCES

[1] M. Husák, J. Kašpar, E. Bou-Harb, and P. Čeleda, "Survey of attack projection, prediction, and forecasting in cyber security," IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 640–660, 2019.

[2] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection, "IEEE Symposium on Security and Privacy, pp. 305–316, 2010.

[3] MITRE Corporation, "MITRE ATT&CK®: A knowledge base of adversary tactics and techniques," 2023. [Online].

[4] [4]B. Al-Sada, A. Sadighian, and G. Oligeri, "Analysis and characterization of cyber threats leveraging the MITRE ATT&CK database," IEEE Access, vol. 12, pp. 1217–1235, 2024.

[5] Y. Kim, J. Kim, and H. Kim, "BAN: Predicting advanced persistent threat attacks using Bayesian networks with the MITRE ATT&CK framework," IEEE Access, vol. 11, pp. 91949–91970, 2023.

[6] X. Cai, H. Zhang, C. M. Ahmed, and H. Koide, "Detecting advanced persistent threat exfiltration using ensemble learning and behavioral metrics," IEEE Access, vol. 12, pp. 81803–81822, 2024.

[7] Z.-S. Chen, X. Zhang, and Y. Wang, "Clustering advanced persistent threat groups through cyber threat intelligence using weighted similarity measures," IEEE Access, vol. 12, pp. 141851–141869, 2024.

[8] M. E. Mazaheri and A. Shameli-Sendi, "AP Tracker:A comprehensive malware dataset for attribution of advanced persistent threats," IEEE Access, vol. 12, pp. 145148–145166, 2024.

[9] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," Advances in Neural Information Processing Systems(NeurIPS),pp.1024–1034,2017.

[10] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," International Conference on Learning Representations (ICLR), 2017.

[11] J. Tang, M. Xu, S. Fu, and K. Huang, "A scheduling optimization technique based on reuse in Spark to defend against APT attacks," Tsinghua Science and Technology, vol. 23, no.5,pp. 550–560, 2018.

[12] I. Kumari and H. Lee, "Detection of advanced persistent threat attacks using optimized deep learning frameworks," IEEE Access, vol. 13, pp. 196404–196421, 2025.

[13] Y. Zhang, X. Chen, and Z. Wang, "Temporal modeling of cyber threats for proactive security intelligence," IEEE Transactions on Information Forensics and Security, vol. 19, pp. 2211–2224, 2024.

[14] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874,2006.

[15] Grand View Research, "Threat intelligence market size, share and trends analysis report," 2024. [Online].