International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

OPEN ACCESS

Available Online at :www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET2512317

Automated Cyber Threat Identification and Natural Language Processing

Kuruvamanikindi Venkatesh, M Sai Kumar, Shaik Mohammed Maaz, Surekari Yashwanth Teja, Dr. K. Pavan Kumar

Department of Artificial Intelligence and Machine Learning, Dr K V Subba Reddy Institute of Technology, Kurnool, Andhra Pradesh, India

ARTICLEINFO

Article History:

Accepted : 05 May 2025
Published: 09 May 2025

Publication Issue :
Volume 12, Issue 3
May-June-2025

Page Number :
89-97

ABSTRACT

The time window between the disclosure of a new cyber vulnerability and its use by Cyber criminals has been getting smaller and smaller over time. Recent episodes, such asLog4j vulnerability, exemplifies this well. Within hours after the exploit being released, attackers started scanning the internet looking for vulnerable hosts to deploy threats like crypto currency miners and ransom ware on vulnerable systems.Thus,it becomes imperative for the cyber security defense strategy to detect threats and their capabilities as early as possible to maximize the success of prevention actions.Althoughcrucial, discovering new threatsisa challenging activity for security analysts due to the immense volume of data and information sources to be analyzed for signs that a threat is emerging. In this sense,wepresenta framework for automatic identification and profiling of emerging threats using Twitter messages as a source of events and MITRE ATT&CK as a source of knowledge for threat characterization.Theframeworkcomprises three main parts: identification of cyber threats and their names; profilingtheidentifiedthreatintermsofitsintentionsorgoalsbyemploying two machine learning layers to filter and classify tweets; and alarm generation based on the threat'srisk.Themaincontributionofourworkistheapproachtocharacterizeor profile the identified threats in terms of their intentions or goals, providing additional context on the threat and avenues for mitigation. In our experiments, theprofilingstagereachedanF1 score of 77% in correctly profiling discovered threats.

Keywords: Cybersecurity, Emerging Threats, Machine Learning, Threat Profiling, MITRE ATTACK

INTRODUCTION

As the cyber landscape continues to evolve, the shrinking timeframe between the disclosure of vulnerabilities and their exploitation by threat actors presents a pressing challenge for cyber security. Recent incidents, exemplified bytheLog4jvulnerability, vividly illustrate this trend. Within hours of its disclosure, malevolent entities swiftly initiated attacks, targeting vulnerable systems for deploying ransom ware and crypto currency miners. This underscores the urgency for cyber security strategies to swiftly detect and comprehend emerging threats to maximize pre-emptive defense actions. Yet, amidst vast volumes of data, identifying nascent threats remains a formidable task for security analysts. To address this challenge, our project introduces a novel framework designed for the automatic identification and profiling of emergent cyber threats, utilizing Twitter as an event source and leveraging MITRE ATT&CK for threat characterization."

"The framework orchestrates three core components: first, the identification of cyber threats and their nomenclature; second, the profiling of these identified threats, discerning their intentions and goals through a sophisticated machine learning architecture; and finally, the generation of alerts based on the risk posed by the identified threats. A significant stride in our work lies in our approach to characterizing these emergent threats, providing contextual insights in to their intentions. This added layer of understanding not only facilitates. Threat identification but also offers avenues for effective mitigation strategies. In our experiment alendeavors, the profiling stage exhibited a commendableF1 score of 77%, demonstrating a robust capability in accurately profiling and understanding discovered threats."

"This project stands at the forefront of proactive cyber security measures, aiming to equip defenders with a sophisticated system capable of early threat detection and nuanced threat characterization. By leveraging Twitter as a valuable source of event data and employing cutting-edge machine learning techniques, the framework not only identifies threats but also delves deeper into their intentions, providing invaluable insights for proactive defense actions against rapidly evolving cyber threats.

LITERATURESURVEY

Cyber Threat Intelligence and Identification:

Severalstudieshighlighttheneedforautomatedthreatdetectionsystemsduetotheever- increasing volume of cyber-attack information. Traditional signature-based methods suchasthosediscussedbyKimetal.(2017)intheirstudyonintrusiondetectionsystems (IDS) are effective in detecting known threats but struggle with new and emerging threats. To address this, the adoption of cyber threat intelligence (CTI) has gained prominence. CTI frameworks, such as the one described by Husák et al. (2019), focus on aggregating and analyzing threat data from various sources, but often require manual intervention, making them less scalable for rapid identification of emerging threats.

Our study builds on the so insights by automating the identification a profiling process, leveragingsocialmediaplatformssuchasTwitterasareal-timesourcefordetectingnew threats.

Natural Language Processing in Cyber security:

NLP has been widely adopted in cyber security for analyzing textual data, such as security reports, threat feeds, and social media posts, to detect anomalies or suspicious behaviour. In Sarker et al. (2020), NLP is applied to security texts and logs to extract indicators of compromise (IoCs). However, the challenge of analyzing unstructured data in natural language remains due to the nuances of human language, suchasslang, jargon, and abbreviations commonly used in cyber security discussions.

In Yao et al. (2018), NLP is used to mine Twitter data for early detection of vulnerabilities, but their

method lacks robust framework for classifying and profiling threats. Our framework builds on this by incorporating MITRE ATT&CK for structured threat profiling, enhancing the quality of threat intelligence derived from NLP.

Social Media as a Source for Threat Detection:

Social media platforms, particularly Twitter, have proven to be valuable sources for detecting emerging threats due to their real-time nature and the speed at which informationspreads.Sabottkeetal.(2015)demonstratedthatTwittercouldbeminedfor early vulnerability disclosures, which were often reported before official data bases such as CVE were updated. However, the volume of data and the challenge of filtering relevant information pose a significant obstacle to effective real-time analysis.

Alruilyetal.(2020)expandedonthisbydevelopingasystemthatusesNLPandmachine learningtoclassifysecurity-relatedtweets.Despitetheirsuccessinidentifying relevant tweets, the system does not provide deeper profiling of threats, leaving a gap in understanding their potential impact andmitigationstrategies.Our proposedframework addresses this gap by not only identifying relevant cyber threats from Twitter but also profilingthemusingatwo-layer machinelearningapproach.Thisaddsanewdimension of understanding regarding the threat's objectives and avenues for exploitation.

MITREATT&CKforThreatProfiling:

The MITRE ATT&CK framework is an industry-standard knowledge base that categorizes adversarial tactics, techniques, and procedures (TTPs). It has become a criticaltoolforthreatprofiling,asseeninstudiessuchasSharmaetal.(2019),wherethe frameworkwasusedtoanalyzeattackvectorsandimprove incidentresponse.ATT&CK provides a structured methodology for characterizing threats, but manually correlating threat data with the ATT&CK framework remains time-consuming and complex.

In our work, MITRE ATT&CK is leveraged as a knowledge base for threat characterization, but wego further by automating theprofiling process. Using machine learning classifiers, our framework automatically aligns detected threats with relevant ATT&CKtacticsandtechniques,makingiteasierforsecurityanalyststounderstandthe nature and potential impact of an emerging threat.

AlarmGenerationandRiskAnalysisinCybersecurity:

Ghafir et al. (2018) introduced an alert generation systemthat triggers alarms based on network anomalies. However, most alarm systems focus solely on known threats and often generate a high number of false positives. Alarm fatigue becomes an issue when security analysts must sift through irrelevant alerts to identify genuine threats.

Our framework improves upon traditional alarm generation systems by combining risk assessment withNLP-basedthreat identification. Byprofilingeachemergingthreat and analyzing its risk level based on its intentions and potential impact, we generate more relevant and actionable alarms, reducing the burden on security analysts.

MachineLearningforThreatDetectionandProfiling:

The application of machine learning in cybersecurity has seen significant growth, particularly for threat detection and profiling. Sommer and Paxson (2010) highlighted theadvantagesofmachinelearningindetectingnewthreatsthroughpatternrecognition, but theyalsoemphasizedthedifficultyof handlingnoisydata. Morerecently, Kumar et al. (2021) used deep learning to classify network threats based on their behavior.

EXISTING SYSTEM

Cybersecurity remains a critical concern for modern organizations due to the increasing number and sophistication of cyber threats. A central component in managing these threats is the **Security Operations Center (SOC),** which is responsible for monitoring and safeguarding an organization's digital

infrastructure. The effectiveness of SOCs depends heavily on receiving **timely, accurate, and actionable threat intelligence.**

Currently, most SOCs rely on **manual processes** for gathering information from various sources—technical blogs, social media, security advisories, and forums—to stay informed about new vulnerabilities and attack patterns. This method is time-intensive and often inefficient due to the **overwhelming amount of irrelevant or redundant data,** increasing the likelihood of **missed warnings or delayed responses.**

1) Role of OSINT in Cybersecurity

One of the most valuable resources in cyber threat intelligence is **Open Source Intelligence (OSINT),** which comprises data collected from publicly available platforms. Among OSINT sources, **Twitter stands out** due to its real-time nature and the active participation of cybersecurity professionals, researchers, and even malicious actors. These users frequently share information about zero-day vulnerabilities, malware campaigns, and ongoing attacks.

2) Existing Research and Limitations

Several studies have attempted to automate the extraction of threat intelligence from OSINT sources, particularly social media. While progress has been made, significant **limitations remain in current methodologies,** including:

1. **Keyword-Based Filtering Approaches**

One early method involves scanning OSINT sources using predefined keywords to flag tweets or posts that mention threats. While fast, this technique suffers from **high false-positive rates**, as many unrelated posts may contain similar keywords. Moreover, such systems **often fail to contextualize or classify** the nature of the threat, rendering them less useful for preventive action.

2. **CNN-Based Threat Classification**

Some systems use **Convolutional Neural Networks (CNNs)** to classify text data (such as tweets) into known categories of threats like **ransomware, phishing, or DDoS attacks.** Although this improves classification accuracy, it **does not name the specific threat instance** (e.g., "WannaCry" ransomware) or offer deeper insight into the **threat's behavior, propagation method, or risk level.**

3. **Novelty Detection Techniques**

Other research utilizes **novelty detection models** to identify tweets or content that differ from known vulnerabilities listed in databases like **CVE (Common Vulnerabilities and Exposures).** These models flag unfamiliar content as potentially new threats. However, they only **label tweets as relevant or irrelevant,** without performing **threat profiling or risk assessment**, limiting their usefulness in decision-making for SOC teams.

4. **NER and BiLSTM-Based Automation**

More advanced tools use **Named Entity Recognition (NER)**and**Bidirectional Long Short-Term Memory (BiLSTM)** neural networks to automate entity extraction and classification. These systems use Twitter's streaming API to collect data and automatically parse relevant security information. While they reduce manual effort, they still **lack multi-class classification**capabilities and do not effectively **link threats to specific tactics or attack goals.**

Disadvantages of the Existing Systems

- **No Multi-Class ML Implementation**: Existing systems typically implement **binary classification** (e.g., threat vs. non-threat) without **multi-class machine learning algorithms** to differentiate between various types of cyber threats in a detailed manner (e.g., ransomware vs. botnet vs. spyware).
- **Lack of Threat Profiling**: Most systems do not **profile or characterize threats** beyond surface-level classification. They do not map threats to frameworks such as **MITRE ATT&CK**, which could help analysts understand

the attacker's tactics, techniques, and procedures (TTPs).

- **No Full Pipeline Execution**: The current systems fail to implement a **comprehensive detection-to-profiling pipeline,** where identified threats are processed, classified, and contextualized with actionable intelligence.
- **Over-Reliance on Keywords**: Excessive dependence on static keyword filters results in **low precision**and**high maintenance costs,** as the keywords must be continuously updated to stay effective.
- **Minimal Risk Assessment**: Most tools do not assess or categorize the **potential risk or severity** of the identified threats, which is vital for prioritizing responses and allocating resources.

PROPOSEDSYSTEM

Objective

The **primary goal** of this research is to develop a comprehensive and automated framework for**identifying and profiling emerging cyber threats**by leveraging**Open Source Intelligence (OSINT)** sources, specifically Twitter. The system aims to generate **timely alerts** for cybersecurity engineers, enabling them to detect, assess, and respond to threats as early as possible in the attack lifecycle.

To fulfill this goal, the proposed solution involves several macro steps, forming a continuous pipeline of threat discovery, classification, and alerting.

Methodology: Macro Steps

1. Continuous OSINT Monitoring

The system will **continuously collectdata** from reputable individuals, researchers, cybersecurity experts, and organizations posting on Twitter. By monitoring these sources in real-time, the system can detect **unfamiliar terms or anomalies** that might indicate the emergence of new cyber threats, malicious campaigns, or novel tactics.

2. Threat Term Extraction

The collected posts are subjected to **Natural Language Processing (NLP)**and**Machine Learning (ML)** techniques to identify which terms are most likely to represent **threat names.** Irrelevant terms or common language noise are filtered out. This step ensures a focused dataset containing only potentially significant cybersecurity terms.

3. Tactic Identification via MITRE ATT&CK

Each detected term is matched against the **MITRE ATT&CK framework**, a globally accessible knowledge base of adversary tactics and techniques. This allows the system to infer the **most likely tactic and technique** associated with the threat. For instance, if a tweet suggests credential harvesting activity, the system will tag it with the corresponding MITRE tactic like "Credential Access" and technique such as "Phishing."

4. Threat Profiling and Risk Assessment

After identifying and mapping the threat, the system builds a **profile** containing key attributes: name, tactic, likely intention, potential target sectors, and inferred methods of execution. Using these attributes, a **risk rate** is calculated based on the **speed of threat propagation**, volume of mentions, and community concern levels. This risk score helps prioritize which alerts are critical.

5. Timely Alert Generation

As soon as a threat is confirmed and profiled, the system generates a **timely alert** containing detailed characterization of the threat. This alert is disseminated to cybersecurity teams and system administrators, giving them an early warning window to investigate, strengthen defenses, and deploy countermeasures.

Advantages of the Proposed System

1. End-to-End Automation

Unlike current systems that stop at detection, this framework completes the entire cycle: **data collection → term identification → tactic mapping → profiling**

→ **alert generation,** all with**minimal human intervention**.

2. Early Threat Detection

By monitoring real-time OSINT sources like Twitter, the system can **detect threats within hours** of public disclosure—long before they're documented in traditional databases like CVE or exploited widely.

3. Context-Rich Threat Profiling

With the integration of MITRE ATT&CK techniques, the system not only identifies threats but also **describes their behavior, goals, and methods**, giving security professionals deeper insight into potential attack strategies.

4. Proactive Defense Enablement

The framework empowers organizations to take **proactive security measures** by providing **early and actionable threat intelligence**, helping to mitigate risks before attacks escalate.

5. Risk-Based Prioritization

By assigning a **risk score** based on the velocity and impact of threat dissemination, the system aids security teams in **prioritizing responses** and focusing resources on the most pressing risks.

Attack Lifecycle Insight

To successfully conduct a cyberattack, adversaries must typically:

1. **Identify vulnerabilities** in systems or software;
2. **Acquire tools and techniques**—either by developing or purchasing from dark web markets;
3. **Select targets** and possibly recruit participants or insiders;
4. **Establish infrastructure** such as command-and-control servers, fake domains, or proxy networks;
5. **Plan and execute** the attack by exploiting the target.

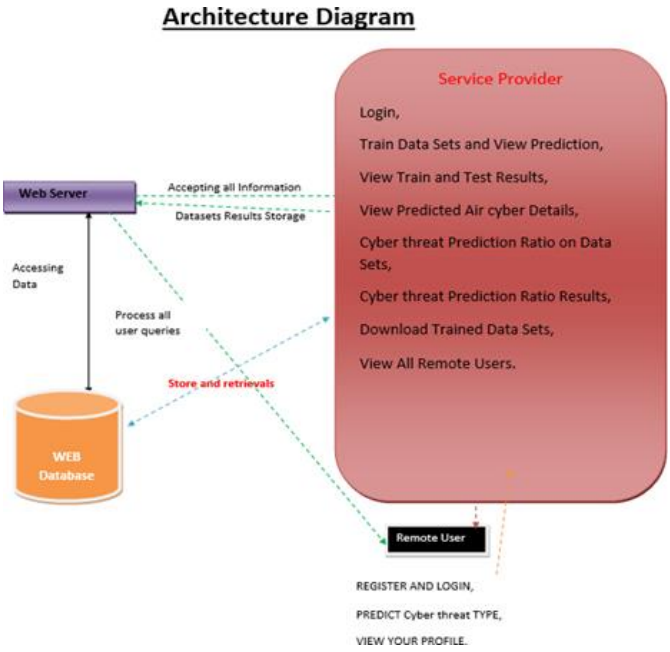
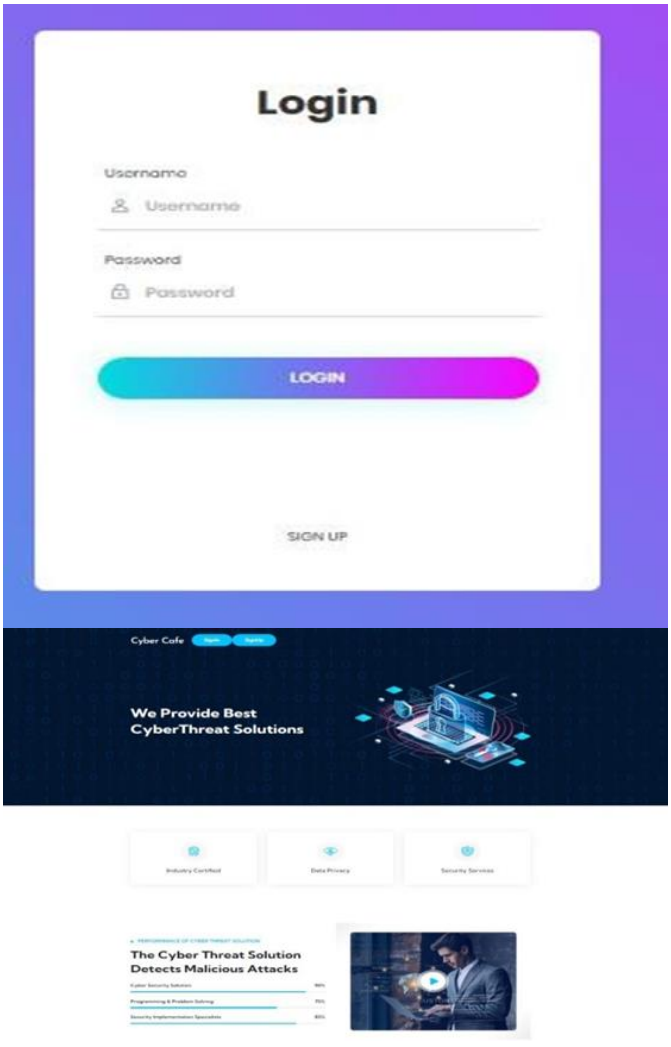
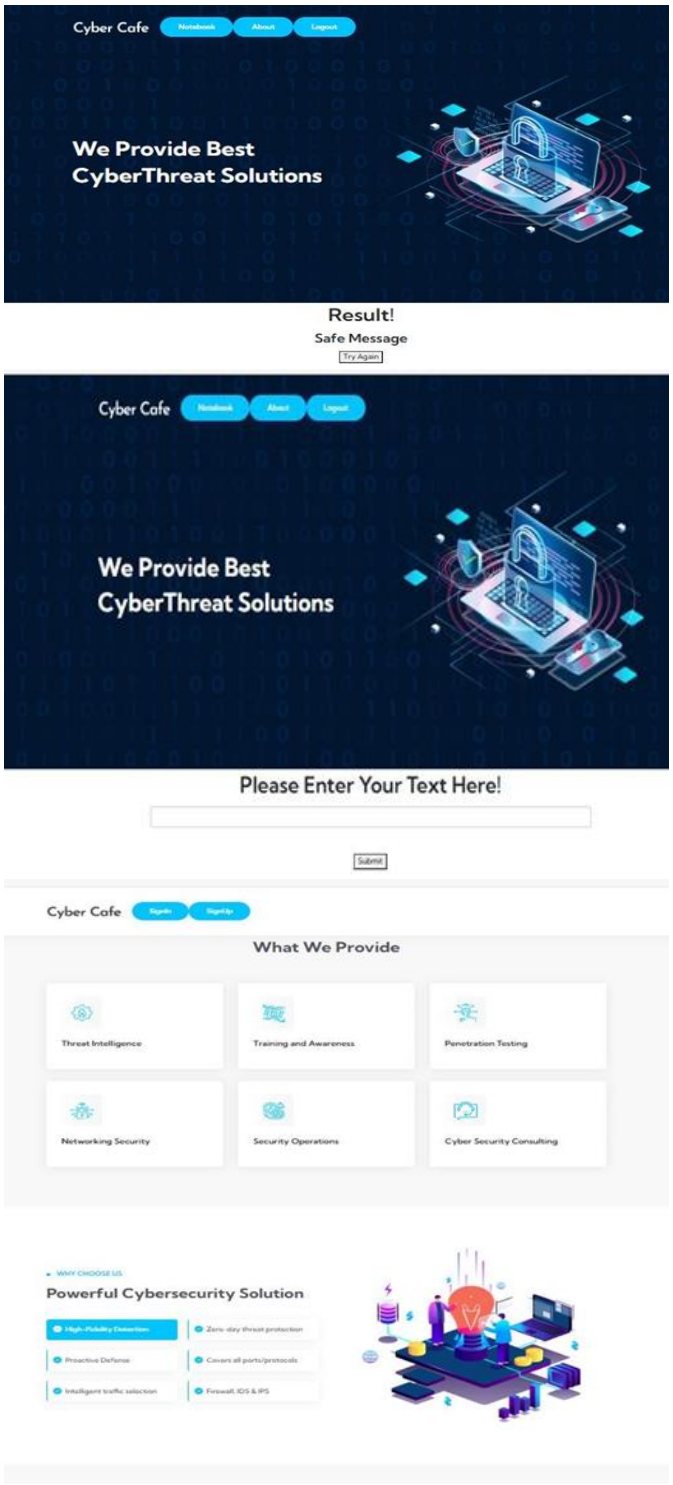


Fig:Architecturediagram

V. RESULTS





CONCLUSION

Given the dynamism of the cyber security field, with new vulnerabilities and threatsappearing at anytime, keeping up to dateonthemisachallenging but importanttaskforanalysts.Evenfollowingthebestpractic esandapplyingthebest controls, anewthreat maybringanunusualwayto subvert thedefensesrequiring a quick response. This way, timely information about emerging cyber threats becomes paramount to a complete cyber security system.

This research proposes an automated cyber threat identification and profiling based on the natural language processing of Twitter messages. The objectiveisexactlytocooperatewiththehardworkoffollo wingtherichsourceof informationthat isTwitterto extract valuable informationabout emerging threats in a timely manner.

This work differentiates itself from others by going a step beyond identifying the threat. It seeks to identify the goals of the threat by mapping the text from tweetsto the procedures conducted by real threats described in MITRE ATT&CK knowledge base. Taking advantage of this evolving and collaborative knowledge base to train machine learning algorithms is a way to leverage the efforts of cyber security community to automatically profile identified cyber threats in terms of their intents.

To put in test our approach, in addition to the research experiment, we implementedthe proposed pipeline and run it for70 days generating online alerts fortheThreatIntelligenceTeamofabigfinancialinstitutio ninBrazil.Duringthis period, at least three threats made the team take preventive actions, such as the PetitPotamcase,describedinsectionV.Oursystemalerte dtheteammakingthem awareofPetit-Potam17daysbeforetheofficialpatchwaspublishedbyMi crosoft. Within this period, the defense teamwas able to implement mitigations avoiding potential exploits and, consequently, incidents.

Ourexperimentsshowedthat theprofiling stagereached an F1 scoreof77%inworking onthiswaybyexperimenting witha different NLP approach using the part of speech (POS) algorithm implementation from Spacy29 Python library. The object is to identify the root verb, the subject, and the objectofthephrasestoselecttweetswheretheactiondescr ibed(therootverb) is referencing the unknown word (the subject)

Future Enhancements :

1. Integration withAdditionalOSINT Sources:
   - Expandingthesystemtocollectandanalyzedata fromabroaderrangeofopen-source intelligence (OSINT) platforms, such as dark web forums, security blogs, and technical vulnerability databases, will improve the ability to detect emerging threats in a wider context.
2. Real-TimeThreat Intelligence:
   - Enhancingthesystemtoprocesstweetsandother OSINTdatainrealtimecan helpprovideinstantthreatalerts.Byreducingthetime it takestodetectand profile threats, organizations can respond more rapidly to emerging vulnerabilities.
3. MultilingualThreatDetection:
   - Implementing multilingual natural language processing (NLP) models would allow the system to identify and profile cyber threats across different languages, improving its global applicability and coverage.
4. AdvancedThreatProfilingand Contextualization:
   - Enhancements in profiling could provide a deeper understanding of threat actors' motivations,tactics,andpatterns.Thiscanbeachievedbyintegratingthesystemwith knowledge bases like MITRE ATT&CK to map threats across the cyber kill chain stages.
5. SentimentAnalysisforThreatRiskAssessment:
   - Adding sentiment analysis can help assess the urgencyor risk levelof detected threatsbasedonthetoneandlanguageusedintheetweets,allowingforamore nuanced understanding of threat severity.

REFERENCES

[1]. B. D. Le, G. Wang, M. Nasim, and A. Babar, "Gathering cyber threat intelligence from Twitter using novelty classification," 2019, arXiv:1907.01755.

[2]. Definition:ThreatIntelligence,GartnerResearch, Stamford,CO,USA,2013.

[3]. R. D. Steele, "Open source intelligence: What is it? why is it important to the military," Journal, vol. 17, no. 1, pp. 35–41, 1996.

[4]. Saqib, M., Malhotra, S., Mehta, D., Jangid, J., Yashu, F., & Dixit, S. (2025). Optimizing spot instance reliability and security using cloud-native data and tools. Journal of Information Systems Engineering and Management, 10(14s), 720–731. https://doi.org/10.52783/jisem.v10i14s.2387

[5]. C.Sabottke,O.Suciu,andT.Dumitras,"Vulnerabilitydisclosureintheageofsocial media:ExploitingTwitterforpredictingreal-worldexploits,"inProc.24thUSENIX Secur. Symp. (USENIX Secur.), 2015, pp. 1041– 1056.

[6]. Jangid, J. (2025). Secure microservice communication in optical networks. Journal of Information Systems Engineering and Management, 10(21s), 911–926. https://doi.org/10.52783/jisem.v10i21s.3455

[7]. A. Sapienza, A. Bessi, S. Damodaran, P. Shakarian, K. Lerman, and E. Ferrara, "Early warnings of cyber threats in online discussions," in Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW), Nov. 2017, pp. 667–674.

[8]. Dixit, S., & Jangid, J. (2025). Exploring smart contracts and artificial intelligence in FinTech. Journal of Information Systems Engineering and Management, 10(14s), 282–295. https://doi.org/10.52783/jisem.v10i14s.2208

[9]. E. Nunes, A. Diab, A. Gunn, E. Marin, V. Mishra, V. Paliath, J. Robertson, J. Shakarian, A. Thart, and P. Shakarian, "Darknet and deepnet mining for proactive cybersecuritythreat intelligence,"inProc.

IEEEConf.Intell.Secur.Informat.(ISI), Sep. 2016, pp. 7– 12.

[10]. S.Mittal,P.K.Das,V.Mulwad,A.Joshi,andT.Finin, "CyberTwitter:UsingTwitter togeneratealerts for cybersecuritythreatsand vulnerabilities," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2016, pp. 860–867.

[11]. A.Attarwala,S.Dimitrov,andA.Obeidi,"Howeffic ientisTwitter:Predicting2012 U.S. presidential elections using support vector machine via Twitter and comparing against Iowaelectronicmarkets,"inProc.Intell.Syst.Conf. (IntelliSys),Sep.2017, pp.646– 652.

[12]. N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, "Towards end-to-end cyberthreat detection from Twitter using multi-task learning," in Proc. Int. Joint Conf.Neural Netw. (IJCNN), Jul. 2020, pp. 1–8. [10]O. Oh, M. Agrawal, andH. R. Rao, "Information control and terrorism: Tracking the Mumbai terrorist attack through Twitter," Inf. Syst. Frontiers, vol. 13, no. 1, pp. 33–43, Mar. 2011.