e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

AUTOMATED EMERGING CYBER THREAT IDENTIFYING AND PROFILING BASED ON NATURAL LANGUAGE PROCESSING

N. Prasad[*1], Khandavalli Surya[*2], Kalepu Vinod[*3], Basina Ramya[*4], Devadasu Srujana[*5]

[*1]Assitant Professor, Department Of Information Technology, Sir C R Reddy College Of Engineering, Eluru, Inida.

[*2,3,4,5]Final Year Students Of Information Technology, Sir C R Reddy College Of Engineering, Eluru, Inida.

DOI: https://www.doi.org/10.56726/IRJMET60143

ABSTRACT

The duration of time that passes between a new cyber vulnerability and its use by cybercriminals has been getting smaller and smaller over time. Within hours after the exploit was released, attackers started scanning the internet looking for vulnerable hosts to deploy threats like cryptocurrency miners and ransomware on vulnerable systems. Thus, it becomes imperative for the cybersecurity strategy to detect threats and their capabilities as early as possible to maximize the success of prevention actions. The framework comprises three main parts: identification of cyber threats and their names; profiling the identified threat in terms of its intentions or goals by employing two machine learning layers to filter and classify tweets; and alarm generation based on the threat's risk. The main contribution of our work is the approach to characterizing or profiling the identified threats in terms of their intentions or goals, providing additional context on the threat and avenues for mitigation.

Keywords: Known Threats, NLP-Based System, Cyber Threat Identification, Texts, New Threats.

I. 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 by the Log4j vulnerability, vividly illustrate this trend. Within hours of its disclosure, malevolent entities swiftly initiated attacks, targeting vulnerable systems for deploying ransomware and crypto currency miners. This underscores the urgency for cyber security strategies to swiftly detect and comprehend emerging threats to maximize defence 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 MITRE ATT&CK for threat characterization." In our experimental endeavours, the profiling stage exhibited a commendable F1 score of 77%, demonstrating a robust capability in accurately profiling and understanding discovered threats. "This project stands at the forefront of proactive cybersecurity 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 defence actions against rapidly evolving cyber threats.

II. LITERATURE REVIEW

1. TITLE: Gathering Cyber Threat Intelligence from Twitter Using Novelty Classification

AUTHOR: B. D. Le, G. Wang, M. Nasim, and A. Babar

Preventing organizations from cyber exploits needs timely intelligence about cyber vulnerabilities and attacks, referred to threats. Cyber threat intelligence can be extracted from various sources including social media platforms where users publish threat information in real time. Gathering Cyber threat intelligence from social media sites is a time consuming task for security analysts that can delay a timely response to emerging cyber threats. We propose a framework for automatically gathering cyber threat intelligence from Twitter by using a novelty detection model. We evaluate our framework using a purpose-built data set of tweets from 50 influential cyber security related accounts over twelve months (in 2018). Our classifier achieves an F1-score of

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

0.643 for classifying cyber threat tweets and outperforms several baselines, including binary classification models. Our analysis of the classification results suggests that cyber threat relevant tweets on Twitter do not often include the CVE identifier of the related threats.

2. TITLE: Open-source intelligence: What is it? Why is it important to the military?

AUTHOR: R. D. Steele

OSINT is the collection and analysis of data from public sources, such as social media, websites, radio, and broadcast TV, to produce actionable intelligence. This data can be in the form of audio, image, video, or text. OSINT is used by businesses, organizations, and non-governmental organizations, and is primarily used in law enforcement, national security, and business intelligence functions.

3. TITLE: Multi-Task Deep Neural Networks for Natural Language Understanding

AUTHOR: Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao

In this paper, we present a Multi-Task Deep Neural Network (MT-DNN) for learning representations across multiple natural language understanding (NLU) tasks. MT-DNN not only leverages large amounts of cross-task data, but also benefits from a regularization effect that leads to more general representations in order to adapt to new tasks and domains. MT-DNN extends the model proposed in Liu et al. (2015) by incorporating a pre-trained bidirectional transformer language model, known as BERT (Devlin et al., 2018). MT-DNN obtains new state-of-the-art results on ten NLU tasks, including SNLI, SciTail, and eight out of nine GLUE tasks, pushing the GLUE benchmark to 82.7% (2.2% absolute improvement). We also show that MT-DNN representations enable effective domain adaptation with significantly fewer in-domain labels compared to pre-trained BERT representations, as demonstrated using the SNLI and SciTail datasets.

4. TITLE: Early Warnings of Cyber Threats in Online Discussions

AUTHOR: A. Sapienza, A. Bessi, S. Damodaran, P. Shakarian, K. Lerman, and E. Ferrara

We introduce a system for automatically generating warnings of imminent or current cyber threats. Our system leverages the communication of malicious actors on the dark web, as well as the activity of cyber security experts on social media platforms like Twitter. In a time period between September, 2016 and January, 2017, our method generated 661 alerts, of which about 84% were relevant to current or imminent cyber-threats. In the paper, we first illustrate the rationale and workflow of our system, then measure its performance. Our analysis is enriched by two case studies: the first shows how the method could predict DDoS attacks, and how it would have allowed organizations to prepare for the Mirai attacks that caused widespread disruption in October 2016. Second, we discuss the method's timely identification of various instances of data breaches.

Existing System:

As cyber threats constantly evolve, safeguarding organizations has become crucial. The Security Operations Center(SOC) acts as the central defense system, but its effectiveness hinges on timely and relevant threat intelligence Security analysts face the challenge of manually processing mountains of information, often with limited results due to irrelevant data. Thankfully, Open Source Intelligence (OSINT) emerges as a valuable tool for identifying emerging cyber threats.

Disadvantages

- An existing system never implemented Multi-Class machine learning (ML) algorithms - the next steps in the pipeline.
- An existing system didn't implement the following method to identify and classify threats.

III. PROPOSED SYSTEM

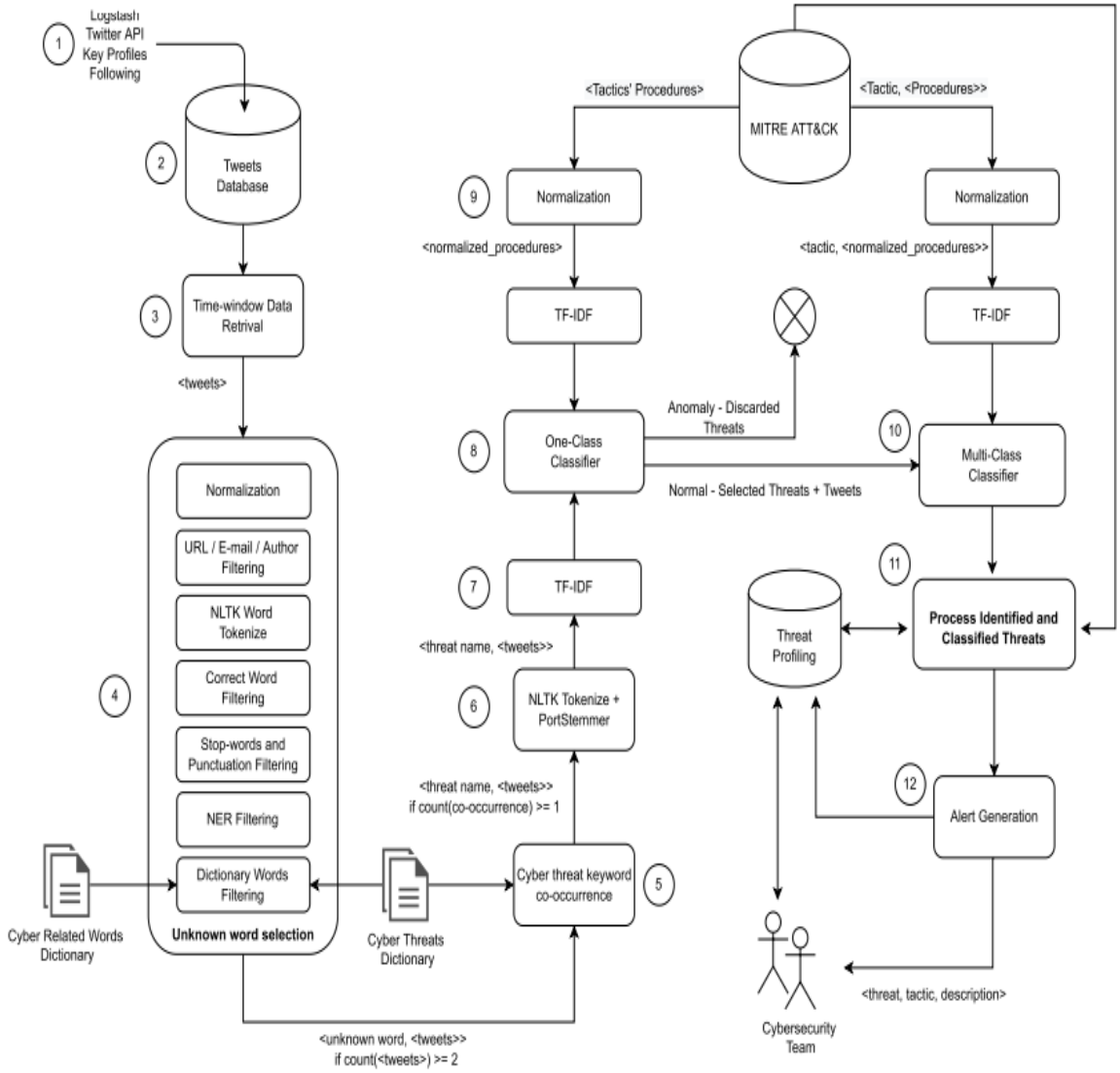
The overall goal of this work is to propose an approach to automatically identifying and profiling emerging cyber threats based on OSINT (Open Source Intelligence) in order to generate timely alerts to cyber security engineers. To achieve this goal, we propose a solution, continuously monitoring and collecting posts from prominent people and companies on Twitter to mine unknown terms related to cyber threats and malicious campaigns.

Advantages

To conduct a cyberattack, malicious actors typically have to

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

- Identify vulnerabilities,
- Acquire the necessary tools and tradecraft to successfully exploit them,
- Choose a target and recruit participants.



System Architecture

To employ word tokenization in the proposed solution, we use the Natural Language ToolKit (NLTK) Python module.11 The output of this step is, for each tweet, an array of its words or tokens. See in the example below how the tweet is split into tokens: Tweet: ‘‘The RobbinHood ransomware is using a vulnerable legacy Gigabyte driver in order to get around antivirus protections’’. Tokens: [’The’, ‘RobbinHood’, ‘ransomware’, ‘is’, ‘using’, ‘a’, ‘vulnerable’, ‘legacy’, ‘Gigabyte’, ‘driver’, ‘in’, ‘order’, ‘to’, ‘get’, ‘around’, ‘antivirus’, ’protections’].

IV. METHODOLOGY

- **Data Collection:**
  - Gather text data from various sources, such as news articles, social media posts, forums, blogs, and security reports.
- **Data Preprocessing:**
  - Clean and preprocess the collected text data by removing noise, irrelevant information, and standardizing the format. This includes tokenization, removing stopwords, stemming, and lemmatization.
- **Named Entity Recognition (NER):**

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

- Identify and classify named entities (like people, organizations, locations) in the text data to understand the context and relationships.

- **Sentiment Analysis:**
- Determine the sentiment or attitude expressed in the text (positive, negative or neutral) to gauge the public perception of cybersecurity issues.

- **Topic Modelling:**
- Use algorithms like Latent Dirichlet Allocation (LDA) to identify and categorize the main topics discussed in the text data.

- **Pattern Recognition and Anomaly Detection:**
- Develop algorithms to detect unusual patterns or anomalies in the text data that may indicate emerging cyber threats.

- **Machine Learning and Deep Learning:**
- Train models using historical data to recognize new and evolving threats based on identified patterns and trends.

- **Threat Profiling:**
- Create detailed profiles for different types of cyber threats, including their nature, potential impact, affected systems, and recommended mitigation strategies.

- **Alerting and Reporting:**
- Implement systems to generate alerts and detailed reports when significant threats are detected, notifying relevant stakeholders and providing actionable insights.

- **Continuous Monitoring and Updating:**
- Continuously monitor new data and update models and algorithms to adapt to evolving threats and improve the system's accuracy over time.

These methodologies work together to automatically detect, analyze, and profile emerging cyber threats, providing comprehensive and proactive cybersecurity protection.

V. RESULT

The final result of an automated system for cyber threat identification and profiling based on natural language processing (NLP) is a comprehensive and proactive cybersecurity solution. It provides real-time detection alerts, notifying cybersecurity professionals about potential threats as they emerge from analysing text data. The system generates detailed threat profiles, offering in-depth descriptions of each identified threat, including its nature, potential impact, affected systems, and recommended mitigation strategies. Additionally, it delivers sentiment insights, which help understand public reactions and concerns regarding various cyber threats. This allows organizations to gauge the seriousness and urgency of these threats. The system also produces comprehensive reports, summarizing the findings and offering actionable recommendations for enhancing cybersecurity measures.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

VI. CONCLUSION

Given the dynamism of the cyber security field, with new vulnerabilities and threats appearing at any time, keeping up to date on them is a challenging but important task for analysts. Even following the best practices and applying the best controls, a new threat may bring an unusual way to subvert the defences requiring 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 objective is exactly to cooperate with the hard work of following the rich source of information that is Twitter to extract valuable information about 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 tweets to the procedures conducted by real threats described in MITRE ATT&CK knowledge base. Our experiments showed that the profiling stage reached an F1 score of 77% in correctly profiling discovered threats among 14 different tactics and the percentage of false alerts of 15%. In future work, we consider it important to advance in tweets selection stages (Unknown Words and One-class), to improve the false positives rate and in the profiling stage, to reach higher accuracy in determining the technique associated with the identified threat.

VII. FUTURE SCOPE

We can enhance automated cyber threat identification and profiling using natural language processing by refining the algorithms to better understand nuanced language patterns indicative of threats. This involves training the system to recognize subtle cues, such as changes in tone or context, within text data. Additionally, integrating machine learning techniques can enable the system to adapt and improve its accuracy over time by learning from new data. Moreover, incorporating advanced linguistic analysis methods can help in identifying and extracting relevant information from unstructured text sources, enhancing the system's ability to detect emerging threats swiftly and accurately.

VIII. 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: Threat Intelligence, Gartner Research, 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] C. Sabottke, O. Suciu, and T. Dumitras, ‘‘Vulnerability disclosure in the age of social media: Exploiting Twitter for predicting real-world exploits,’’ in Proc. 24th USENIX Secur. Symp. (USENIX Secur.), 2015, pp. 1041–1056.

[5] 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.

[6] 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 cybersecurity threat intelligence,’’ in Proc. IEEE Conf. Intell. Secur. Informat. (ISI), Sep. 2016, pp. 7– 12.

[7] S. Mittal, P. K. Das, V. Mulwad, A. Joshi, and T. Finin, ‘‘CyberTwitter: Using Twitter to generate alerts for cybersecurity threats and vulnerabilities,’’ in Proc. IEEE/ACM Int. Conf. Adv. Social Newt. Anal. Mining (ASONAM), Aug. 2016, pp. 860–867.

[8] A. Attar Wala, S. Dimitrov, and A. Obeidi, ‘‘How efficient is Twitter: Predicting 2012 U.S. presidential elections using support vector machine via Twitter and comparing against Iowa electronic markets,’’ in Proc. Intell. Syst. Conf. (IntelliSys), Sep. 2017, pp. 646–652.

[9] Terrorism: Tracking the Mumbai terrorist attack through Twitter,’’ Inf. Syst. Frontiers, vol. 13, no. 1, pp. 33–43, Mar. 2011.

[10] T. Sakaki, M. Okazaki, and Y. Matsuo, ‘‘Earthquake shakes Twitter users: Real-time event detection by social sensors,’’ in Proc. 19th Int. Conf.World Wide Web, Apr. 2010, pp. 851–860.