AI System For MITRE ATT&CK Threat Classification

Prapatsorn Alongkornpradub, Vorameth Reantongcome, Ekkarat Techanawakarnkul
Department of Data Sciences and Artificial Intelligence, Asian Institute of Technology
{st124846, st124903, st124945}@ait.ac.th

Abstract

Cybersecurity threats are rapidly evolving, making traditional methods of threat analysis inefficient and time-consuming. This research proposes an AI-powered system that leverages Natural Language Processing (NLP) to automatically classify cybersecurity news articles into MITRE ATT&CK techniques. A fine-tuned BERT model automates the classification process, enabling organizations to analyze threat intelligence more effectively. The experiment involves training and evaluating the BERT model on cybersecurity news articles, assessing its performance using classification metrics. The expected result is an accurate system that automates the mapping of MITRE ATT&CK techniques, reducing manual effort and enhancing the accuracy of threat classification. Additionally, the system will help organizations make more informed security decisions by identifying which departments, systems, or assets are potentially affected by emerging threats. A web-based interface was also developed to demonstrate the system's classification capability.

1 Introduction

The increasing volume and complexity of cybersecurity threats pose significant challenges for organizations, governments, and individuals. As attacks become harder to detect and understand, it's more important than ever to quickly analyze threats and respond to them in time. As cyberattacks become more sophisticated, the ability to quickly analyze and respond to these threats is critical in protecting important information and systems. The MITRE ATT&CK framework helps security analysts understand how attackers work. However, identifying relevant MITRE ATT&CK from a high volume of cybersecurity news articles remains a challenging task due to the complexity of technical language, inconsistent writing styles, and the large volume of threat reports. Manual mapping of security incidents to the ATT&CK framework is inefficient and time-consuming, making it difficult for organizations to keep pace with emerging threats.

To address this challenge, this research proposes an AI-powered system that utilizes Natural Language Processing (NLP) to automatically classify cybersecurity news articles into MITRE ATT&CK techniques. By fine-tuning a BERT model on a labeled dataset of cybersecurity incidents and their corresponding ATT&CK techniques. This helps the system learn how to understand and classify threats based on the way information is written in news articles. Security teams can then use the classified techniques to identify the affected areas within an organization further.

Unlike traditional keyword-based or rule-based threat intelligence systems, which rely on predefined rules and patterns, this system will leverage deep learning-based text classification to significantly enhance both accuracy and scalability. Many existing systems analyze cybersecurity data, but few focus on automatically mapping cybersecurity news to the MITRE ATT&CK framework, which is crucial for a deeper understanding of attacker behaviors. Cybersecurity news articles often contain complex terminology, ambiguous descriptions of attacks, and diverse writing styles, making automated classification a particularly challenging task. By fine-tuning a BERT model specifically on cybersecurity threat intelligence data, this research aims to automate the process of threat classification, reducing the time and effort needed for manual analysis. This will not only enable security teams to respond to emerg-

ing threats more quickly but also minimize human error and inconsistencies that arise from manual mapping. Ultimately, the system will help organizations make more informed security decisions and improve their overall threat response efficiency.

The study addresses the following key research questions:

- How does the classification accuracy of the proposed system compare to existing chatbots?
- How effectively can the AI model understand and match complex attack descriptions to MITRE ATT&CK techniques, particularly when clear keywords are not present?

By addressing these questions, this research aims to develop an NLP-based system that automates the classification of cybersecurity threats, improving efficiency and reducing the manual effort required in the threat analysis process. The system, through the integration of a fine-tuned BERT model, will provide structured and actionable threat intelligence, helping organizations improve cybersecurity decision making and enhance incident response times. To support the practical application of the system, a web-based interface was developed to enable security analysts to interact with the model and automatically extract classified MITRE ATT&CK techniques from cyber threat news articles.

2 Related Works

2.1 Pre-trained Models and Classification Frameworks

Transformer-based models, especially BERT, have shown strong potential in cybersecurity tasks like threat detection and classification. Park and You (2023) developed CTI-BERT, a pre-trained model for cyber threat intelligence, outperforming general models in tasks like MITRE ATT&CK classification. However, limitations remain due to dataset size and language coverage. Similarly, Park and Lee (2022) proposed a full-stack NLP pipeline to extract threat intelligence from unstructured texts, though it lacked integration with structured frameworks like MITRE ATT&CK.

2.2 Handling Multi-Label Challenges in MITRE ATT&CK Classification

To handle the complexity of TTP classification, Nguyen et al. (2024) introduced a Noise Contrastive Estimation (NCE) framework to address multi-label challenges and missing annotations, achieving better performance than baseline models. Orbinato et al. (2022) also explored classifying unstructured CTI into MITRE ATT&CK categories, showing that deep learning outperformed traditional approaches, particularly at the sentence level.

2.3 Cybersecurity Datasets and Entity-Level Extraction

Lange et al. (2024) released the AnnoCTR dataset, which provides detailed annotations for threat reports linked to MITRE ATT&CK, useful for training classification models. Trong et al. (2020) introduced CySecED, expanding on earlier event detection datasets with broader event types and document-level context, though models still struggle with complex dependencies.

Additionally, Named Entity Recognition (NER) has become integral in identifying and classifying entities like malware, attack techniques, and affected systems. Studies have used NER to improve classification accuracy by linking identified entities to structured taxonomies such as MITRE ATT&CK. This approach is particularly important in automating the extraction of relevant information from unstructured cybersecurity texts, as discussed by Park and Lee (2022).

2.4 Aspect-Based and Context-Aware Modeling

Attention-based models have helped improve how well systems understand specific parts of a sentence, which is important in tasks like aspect-based sentiment analysis (ABSA). The context-guided BERT model by Wu and Ong (2020) adjusts the attention mechanism to better capture the relationship between targets and their context. A "quasi-attention" mechanism is introduced, allowing the model to give both positive and negative attention to words. This helps highlight important information while also reducing the influence of less relevant words. Such an approach is useful in cybersecurity, where understanding the link between parts of a threat report and MITRE ATT&CK techniques is critical. The approach can be mathematically represented as follows:

2.4.1 Context-Guided Quasi-Attention

The QACG-BERT model for enhancing the context guided in BERT model uses a quasi-attention

function for (T)ABSA where a new attention matrix combines the regular attention matrix with a quasi-attention matrix, enabling capturing more context of the combination of the sentence and aspects. The main equation is formulated below according to Wu and Ong (2020):

$$\hat{\mathbf{A}}^h = \mathbf{A}^h_{\text{Self-Attn}} + \lambda^h_A \, \mathbf{A}^h_{\text{Quasi-Attn}} \tag{1}$$

To briefly explain the calculation of $\mathbf{A}_{\text{Quasi-Attn}}$. The similarity measurement function will capture the similarity between two quasi-context and scaled with distance and α for a scaling factor.

$$\mathbf{A}^h_{\text{Quasi-Attn}} = \alpha \cdot \text{sigmoid}\left(\frac{f_\psi(\mathbf{C}^h_Q, \mathbf{C}^h_K)}{\sqrt{d_h}}\right) \tag{2}$$

For the quasi-attention matrix, the quasi-context query $\mathbf{C}^h_Q$ and the quasi-context key $\mathbf{C}^h_K$ must be defined where $\{\mathbf{Z}_Q, \mathbf{Z}_K\} \in \mathbb{R}^{d_e \times d_h}$ are weights of the linear layers that transform the embbeded raw context matrix. Next, the quasi-attention matrix is defined as:

$$\begin{bmatrix}\mathbf{C}^h_Q \\ \mathbf{C}^h_K\end{bmatrix} = \mathbf{C}^h \begin{bmatrix}\mathbf{Z}_Q \\ \mathbf{Z}_K\end{bmatrix} \tag{3}$$

Moving to the bidirectional gating factor or scalar controlling the effect of context-attention calculation λ^h_A is defined as:

$$\begin{bmatrix}\lambda^h_Q \\ \lambda^h_K\end{bmatrix} = \text{sigmoid}\left(\begin{bmatrix}\mathbf{Q}^h \\ \mathbf{K}^h\end{bmatrix}\begin{bmatrix}\mathbf{V}^h_Q \\ \mathbf{V}^h_K\end{bmatrix} + \begin{bmatrix}\mathbf{C}^h_Q \\ \mathbf{C}^h_K\end{bmatrix}\begin{bmatrix}\mathbf{V}^{\mathbf{C}}_Q \\ \mathbf{V}^{\mathbf{C}}_K\end{bmatrix}\right) \tag{4}$$

$$\lambda^h_A = 1 - (\beta \cdot \lambda^h_Q + \gamma \cdot \lambda^h_K), \tag{5}$$

Both $\{\beta, \gamma\}$ are scalars that control the composition weightings.

From the aforementioned model, it can allow the model to capture more attention in different complex compositional operations.

3 Methodology

With the proposed solution, the deep detail will be explained according to the tasks which will be part of the whole system.

3.1 Workflow

As a cybersecurity engineer, identifying the threat from news or articles could cost a lot of time for manually processing. From the time consuming problem, the cybersecurity engineers must spend much of their man-day in reading and understanding the news or articles, identifying the technique and tactics using and mapping into the potential organization resources. After exploring the Natural Language Understanding course, the old workflow could be improved by applying applications of natural language processing, hoping to reduce the time consuming problem and help to quickly identify the threats before hackers could attack the organization.

3.1.1 Old Workflow

As shown in Figure 1, the old workflow started with searching through the websites and filtering any threat or cybersecurity incident news, articles, or forums which related to the organization. Then, the information will be checked against each team in order to identify potential effects of the systems. After verification, the information will be extracted for techniques and tactics which will help to identify the pattern that attacks in organization. If the patterns are matched, the relevant parties and SOC team will be informed for initializing a mitigation plan.

From the aforementioned workflow, the inefficiencies process is in the extraction of techniques and tactics which require security expertise and knowledge to understand and map the attack with MITRE ATT&CK Framework due to the complexity of the framework. This could take a day for a security expert to identify and classify.

3.1.2 New Workflow

As shown in Figure 2, in the very first step, the security datasets will be pre-processed according to the techniques for the selected pre-trained model. The common pre-processing methods are cleaning text, formatting, or adding marked characters. Additionally, the datasets will be explored for the compatibility and some context on the meaning in order to understand the learnability of the model. After that, the processed datasets will be trained through the pre-trained models available on the Internet. In this preliminary proposal, the BERT model will be chosen as the trained model for specific classification on the MITRE ATT&CK frameworks with cybersecurity context. Moreover, in order to provide the best context capturing for understanding true meaning of sentences related to the MITRE ATT&CK Framework, the Context-Guided Quasi-Attention BERT (Wu and Ong, 2020) will be used for a base model com-

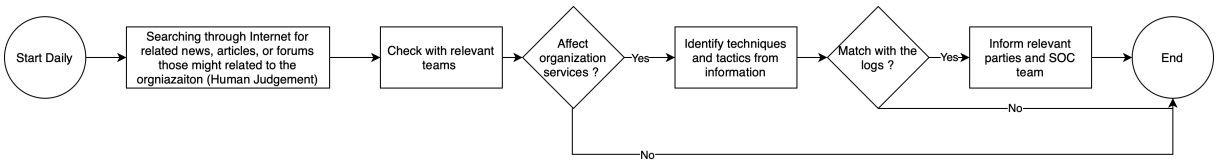

Figure 1: Old Workflow

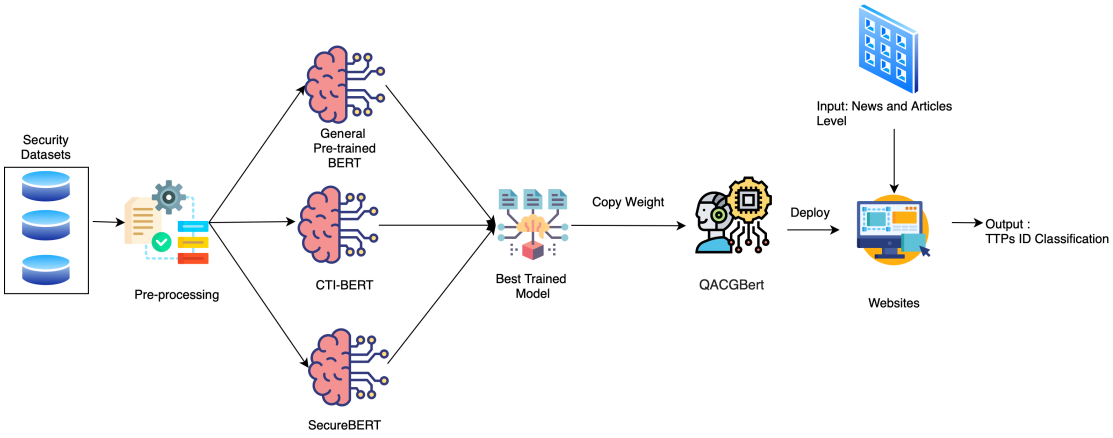

Figure 2: New Workflow

bined with weight from a pre-trained model. In the deployment model, a fine-tuned model will be used to deploy the model.

3.2 Datasets

There is only one dataset that will be used in this methodology for training the model for complying with the classification task for news or articles. The potential dataset is tumeteor/Security-TTP-Mapping (Tumeteor, n.d.) where the authors show the mapping of the security text with the Tactics, Techniques, and Procedures which could imply that hackers might use this kind of process according to the news to convey the attack. This dataset consists of around 14.9K rows for train, around 2.63K rows for validation, and around 3.17K for test. In addition, the dataset is part of paper from "Noise Contrastive Estimation-based Matching" (Nguyen et al., 2024) where the text describes cyber attacks that will be mapped with one or multiple MITRE ATT&CK IDs.

3.3 Data preprocessing

In the data preprocessing stage, the sentences in the dataset are standardized to prepare them for model training. This process includes converting all text to lowercase and removing any unnecessary characters. To convert the raw text into a format suitable for the model, the tokenizer from the pre-trained model is used. The tokenizer transforms words into tokens and ensures the sentences follow the model's input requirements, including padding, start-of-sentence tags, and end-of-sentence tags. Additionally, since the data come with only multiple labels within one sentence. Each multiple will be unpacked and added into another row which allows the model to learn.

As the dataset is already structured with sentence-level inputs, each sentence is mapped to its corresponding label, which will guide the model during training. During preprocessing for the QACGBertForSequenceClassification model, an additional step is performed: named entity recognition (NER) is applied to extract key entities from each sentence. These entities are then incorporated into the dataset as a new column. This added entity information will help the model capture contextual relationships between the entities and improve the accuracy of classification for MITRE ATT&CK techniques and tactics.

3.4 Models

With classification problems, BERT will be the main model that will be used in this proposed solution. The model that will be used in this project is the CTI-BERT (Park and You, 2023) which has been trained to specialize in the cybersecurity domain extended from pre-trained SecBert (Jackad-

uma, 2022) model. Secure-BERT (Aghaei et al., 2022) is another model that will be compared with other BERT models which extend from the pre-trained Ro-Bert model. The base model will be bert-base-uncased from Hugginface. These three models will be compared in order to find the best model that will be used in the system.

Furthermore, QACGBertForSequenceClassification was trained to enhance classification accuracy by focusing on identifying specific words and entities crucial for detecting MITRE ATT&CK techniques. This model incorporates contextual embeddings to capture more nuanced relationships in threat narratives, improving disambiguation between similar attack techniques.

3.5 Training

In the first stage of training, the pre-trained models (CTI-BERT, Secure-BERT, and bert-base-uncased) were trained on a dataset consisting solely of sentences and their corresponding labels. The primary objective at this stage was to determine the best-performing model based on classification accuracy, training loss, and validation loss. Default hyperparameters were used to ensure a fair comparison between the models.

After identifying the best model, the next step involves further fine-tuning using the QACGBertForSequenceClassification model. This model builds on the strengths of the selected pre-trained model but incorporates additional contextual information to improve classification accuracy.

During this phase, the training dataset is enhanced by adding an entity column, which includes entities extracted from each sentence using Named Entity Recognition (NER). By introducing these additional entity features, the model can better focus on key cybersecurity-related terms and concepts, improving its ability to classify MITRE ATT&CK techniques. This fine-tuning process with QACGBertForSequenceClassification allows the model to capture more specific and contextually relevant information, resulting in better overall performance in classifying complex attack patterns.

3.6 Experiment

The experiment consists of two main components to evaluate the effectiveness of different models and input processing methods.

3.6.1 Model Experiment

The objective of the model experiment is to determine the best-performing model for classifying MITRE ATT&CK techniques. Three models will be compared: bert-base-uncased, CTI-BERT, and Secure-BERT. These models will be trained on a dataset consisting of sentences mapped to their corresponding labels (MITRE ATT&CK techniques). The comparison will focus on training loss, validation loss, and classification accuracy. The best-performing model from this experiment will be selected for further fine-tuning.

After identifying the best model, the QACGBertForSequenceClassification model will be used for further training. This model is designed to capture more specific entities within each sentence, improving the accuracy of MITRE ATT&CK techniques classification. The model configuration will be inherited from the selected pre-trained model, with additional contextual parameters to handle cybersecurity-specific entities.

3.6.2 Input Pre-processing Experiment

In the input pre-processing experiment, the dataset (already consisting of sentences and their corresponding labels) will be processed in two ways to determine which results in the most accurate classification:

1. Sentence-Level Processing: Each sentence in the dataset will be treated as a separate input for classification.
2. Entity-Augmented Sentence-Level Processing: The sentences will be enhanced by adding an additional column containing entity information extracted through Named Entity Recognition (NER). This entity information helps the model focus on relevant cybersecurity-related terms.

Both methods will be evaluated based on their impact on classification accuracy and relevance. The performance of each method will be compared to identify the most effective input pre-processing approach.

3.7 Evaluation / Metrics

To evaluate the models, two mechanisms will be used. The first method involves calculating performance metrics such as accuracy, precision, recall, and F1-score, which will evaluate how well the model classifies data during the training process with validation datasets. The second method

is to compute a Normalized Discounted Cumulative Gain (NDCG) score for calculating the different top k rank for recommendations accuracy.

4 Results

This section presents the results from the experiments conducted to evaluate the performance of various BERT-based models in classifying cybersecurity news into MITRE ATT&CK techniques. The evaluation involved three models: BERT-base-uncased, CTI-BERT, and Secure-BERT. These models were trained on a cybersecurity-specific dataset and evaluated using standard classification metrics, including training loss, validation loss, accuracy, precision, recall, and F1-score after applying the Context-Guided Quasi-Attention BERT.

4.1 Standard Classification Metrics

The performance of the four models was evaluated using standard classification metrics, as summarized in Table 1 as follows:

Metrics	BERT-base -uncased	CTI -BERT	Secure BERT	QACG BERT
Training loss	3.575500	2.774000	3.579800	1.706076
Validation loss	3.438241	2.707260	3.307534	4.723936
Accuracy	0.437643	0.546388	0.444487	0.168077
Precision	0.292122	0.431662	0.299807	0.163138
Recall	0.437643	0.546388	0.444487	0.168077
F1-Score	0.333476	0.466138	0.337116	0.144451

Table 1: Standard Classification Metrics Results of BERT-based Models.

From Table 1, QACG-BERT achieved the lowest training loss (1.71) but had the highest validation loss (4.72). Additionally, its accuracy (16.81%) and other metrics were significantly lower compared to the other models. In contrast, CTI-BERT outperformed the other models in terms of accuracy (54.64%), precision (43.17%), recall (54.64%), and F1-score (46.61%), making it the most balanced and effective model overall.

4.2 Normalized Discounted Cumulative Gain (NDCG) Metric

In order to evaluate the model, the ground truth from ChatGPT will be used. A total of 10 news articles from security websites will be collected, and the model will classify Techniques, Tactics, and Procedures (TTPs) IDs. Then, the NDCG will be computed using the ground truth and the probability output of the model, as shown in the following Table 2:

Metrics	CTI-BERT	QACG BERT
NDCG Score with k = 20	0.0000	0.0163

Table 2: NDCG Scores for CTI-BERT and QACG-BERT models.

From the evaluation, it can be seen that the ground truth couldn't be predicting on the CTI-BERT at all. However, when applied the Context-Guided Quasi Attention ideas into the classification. The result is getting better. However, the accuracy is still very low. This could be a limitation of the ground truth with only 10 samples or the datasets couldn't provide coverage of aspect or context in news.

5 Discussion

This study explores how an AI-based system can improve the process of mapping cybersecurity news to the MITRE ATT&CK framework. The proposed model aims to reduce manual workload, improve accuracy, and enable faster incident response. To better understand the strengths and contributions of the system, the discussion is structured around the two main research questions.

5.1 RQ1: How does the classification accuracy of the proposed system compare to existing chatbots?

The initial results indicate that the system's classification accuracy is currently lower than that of existing chatbots. While the fine-tuned BERT model demonstrates significant potential in automating the classification process, further improvements are necessary to achieve the high accuracy levels. One possible reason for the low accuracy could be insufficient training data or issues in the model's generalization across various types of attack descriptions. Future improvements will focus on refining the training process and expanding the dataset to ensure better performance.

5.2 RQ2: How effectively can the AI model understand and match complex attack descriptions to MITRE ATT&CK techniques, particularly when clear keywords are not present?

The model shows potential in understanding complex and unclear attack descriptions, but it cur-

rently struggles to achieve high accuracy in matching these descriptions with the correct MITRE ATT&CK techniques. This is likely due to the model's challenges in generalizing from the training data, especially when dealing with new or more complex attack scenarios. Fine-tuning the BERT model with a more diverse and comprehensive dataset could help improve its ability to match attack descriptions with the correct techniques.

5.3 Impact of the Work

This research can still make a meaningful difference in how cybersecurity teams handle large amounts of threat data, even though the current model's low accuracy is a limitation. The AI-powered system can process data faster than manual methods, but its lack of consistent accuracy reduces how useful it is right now. Fixing these performance issues will be key to making sure the system can help security teams respond to threats quickly and effectively. By improving the system's understanding of complex attack descriptions, it will be able to better assist in decision-making by clearly linking attacks to MITRE ATT&CK techniques.

To demonstrate the usability of the proposed system, a web-based application was also developed. The web app allows security analysts to input unstructured cybersecurity news text and receive predictions of relevant MITRE ATT&CK techniques.

Figure 3 shows the web interface without any input, displaying how the system prepares for user interaction.

Figure 4 shows the interface with an example input and output, where the user submits a cybersecurity news article and the system returns the mapped MITRE ATT&CK techniques.

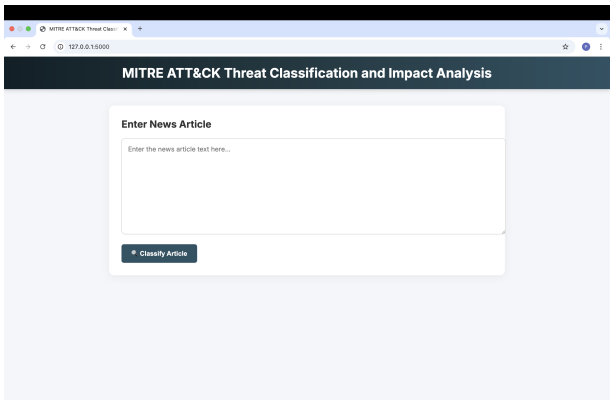

Figure 3: Web interface without any input.

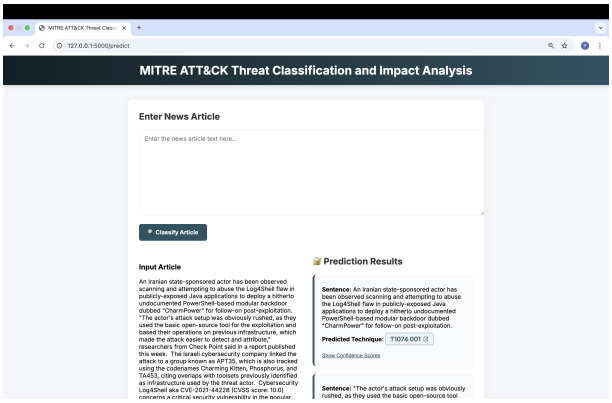

Figure 4: Web interface with sample input and output.

6 Conclusion and Future Work

6.1 Conclusion

The primary goal of this research is to reduce the effort involved in manual analysis while enhancing the consistency and accuracy of threat classification. However, the current low accuracy, especially during the evaluation phase, highlights the need for further improvement. The model's ability to better understand the context and language of cybersecurity news articles is crucial to improving the reliability of threat detection. Therefore, improving the model's accuracy remains a crucial area for future development.

6.2 Limitations and Challenges

Despite its potential, the system faces several limitations that impact its overall performance:

A primary challenge is the low accuracy observed during both the training and evaluation phases. This issue may be due to factors such as an insufficient or imbalanced dataset, overfitting, or the model's difficulty in generalizing to new and unseen attack types.

The variability in attack descriptions across cybersecurity news articles presents another challenge. Different writing styles and terminologies complicate the model's ability to generalize effectively, especially when the descriptions are ambiguous or complex.

Moreover, the quality and size of the training data play a crucial role in the model's performance. The current dataset may not cover a wide enough range of attack techniques or may be too small, leading to poor learning of certain attack behaviors and limited robustness.

Lastly, the system's current language limitation, which supports only English, restricts its applicability in non-English-speaking regions. Expanding the model's language capabilities is crucial for broader adoption and enhanced usability.

6.3 Future Work

Future improvements should focus on several key areas. The dataset should be expanded to include more types of attacks and descriptions, along with using techniques like creating new data or testing with tough examples to improve the model robustness. The model should also be fine-tuned with more specific cybersecurity data to better understand attack reports and the MITRE ATT&CK framework.

Additionally, multilingual support should be added so the system can handle reports in different languages. Finally, integration with live threat feeds should allow the system to quickly adapt to new attack methods and classify threats in real time.

Acknowledgments

We would like to express our sincere gratitude to **Professor Chaklam Silpasuwanchai** for his valuable guidance and support throughout this project. We also thank **Mr. Todsavad Tangtortan** for his helpful advice and feedback during the development of our system.

References

Ehsan Aghaei, Xi Niu, Waseem Shadid, and Ehab Al-Shaer. 2022. SecureBERT: A Domain-Specific Language Model for Cybersecurity. arXiv preprint arXiv:2204.02685. https://arxiv.org/abs/2204.02685.

Hieu Man Duc Trong, Duc Trong Le, Amir Pouran Ben Veyseh, Thuat Nguyen, and Thien Huu Nguyen. 2020. Introducing a New Dataset for Event Detection in Cybersecurity Texts. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5381–5390, Online. Association for Computational Linguistics.

Jackaduma. 2022 SecBERT. https://github.com/jackaduma/SecBERT.

Lukas Lange, Marc Müller, Ghazaleh Haratinezhad Torbati, Dragan Milchevski, Patrick Grau, Subhash Chandra Pujari, and Annemarie Friedrich. 2024. AnnoCTR: A Dataset for Detecting and Linking Entities, Tactics, and Techniques in Cyber Threat Reports. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 1147–1160, Torino, Italia. ELRA and ICCL.

Vittorio Orbinato, Mariarosaria Barbaraci, Roberto Natella, and Domenico Cotroneo. 2022. Automatic Mapping of Unstructured Cyber Threat Intelligence: An Experimental Study. arXiv preprint arXiv:2208.12144. https://arxiv.org/abs/2208.12144.

Tu Nguyen, Nedim Šrndić, and Alexander Neth. 2024. Noise Contrastive Estimation-based Matching Framework for Low-Resource Security Attack Pattern Recognition. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 355–373, St. Julian's, Malta. Association for Computational Linguistics.

Tumeteor. n.d. Security-TTP-Mapping. https://huggingface.co/datasets/tumeteor/Security-TTP-Mapping.

Youngja Park and Taesung Lee. 2022. Full-Stack Information Extraction System for Cybersecurity Intelligence. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 531–539, Abu Dhabi, UAE. Association for Computational Linguistics.

Youngja Park and Weiqiu You. 2023. A Pretrained Language Model for Cyber Threat Intelligence. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 113–122, Singapore. Association for Computational Linguistics.

Zhengxuan Wu and Desmond C. Ong. 2020. Context-Guided BERT for Targeted Aspect-Based Sentiment Analysis. arXiv preprint arXiv:2010.07523. https://arxiv.org/pdf/2010.07523.