Extracting Multiple Relations between Entities from Unstructured Threat Intelligence Reports

Nguyen Dai Tho, Nguyen Trung Hieu, Nguyen Ngoc Bao Tran, Nguyen Phuong Anh

***Abstract*— Cyber threats are becoming an ever-increasing concern for organizations and even countries. Attackers are constantly in search of new and sophisticated attack vectors. On the other side, security defenders need to gather as much as possible information about the threats on the Internet and analyze them to understand current and emerging attack trends for effectively detecting and mitigating potential threats with fast response. This paper addresses the problem of automatically extracting threat intelligence from unstructured text sources. We focus specifically on the possibility of multiple relations between two entities and propose a two-stage process that allows any binary classifier to be used for multi-class classification without interfering with the binary algorithm used. The experimental results illustrate the efficiency of our proposed approach.**

***Tóm tắt*— Các mối đe dọa an ninh mạng đang trở thành mối lo ngại ngày càng tăng đối với các tổ chức và thậm chí là cả các quốc gia. Những kẻ tấn công luôn tìm kiếm các phương pháp tấn công mới và tinh vi. Mặt khác, các chuyên gia bảo mật cần thu thập càng nhiều thông tin về các mối đe dọa trên Internet nhằm phân tích chúng để hiểu rõ các xu hướng tấn công hiện tại và mới nổi để có thể phát hiện, ngăn chặn các mối đe dọa tiềm ẩn một cách hiệu quả và nhanh chóng. Bài báo này đề cập đến vấn đề trích xuất thông tin đe dọa tự động từ các nguồn văn bản không cấu trúc. Nhóm tác giả tập trung đặc biệt vào khả năng tồn tại của nhiều mối quan hệ giữa hai thực thể, đồng thời đề xuất một quy trình hai giai đoạn cho phép bất kỳ bộ phân loại nhị phân nào được sử dụng cho phân loại đa lớp mà không làm ảnh hưởng đến thuật toán nhị phân đã sử dụng. Kết quả thực nghiệm minh họa hiệu suất của phương pháp mà nhóm tác giả đề xuất.**

Keywords — Relation extraction; named entity recognition; threat intelligence; deep learning.

Từ khóa— Trích xuất quan hệ; nhận dạng thực thể được đặt tên; tình báo mối đe dọa; học sâu.

I. INTRODUCTION

In today's interconnected world, cyber threats have become more and more sophisticated and complex. Attackers have been continually changing their tactics and refining their techniques, making attacks increasingly more difficult to defend against. Defenders need to stay alert and be prepared to avoid financial losses, infrastructure disruptions, and damage to reputation.

Threat intelligence is an important aspect of cybersecurity focused on the collection and analysis of information about potential threats that can affect the confidentiality, integrity, and availability of information and systems. Threat intelligence enables security teams to be more proactive and agile in responding to emerging security challenges. Organizations can prioritize focusing resources on the most critical risk areas, rather than wasting time and effort on threats that pose little or no risk. It also helps to improve incident management and reduce response time. Refer to 0 for a more comprehensive definition of threat intelligence and an overview of the related issues.

Threat intelligence comes at a cost. Security teams are required to stay attentive and always keep an eye out for potential threats. They need to turn large volumes of raw data into valuable and useful knowledge fast and accurately for early identifying the

This manuscript is received on August 24, 2023. It is commented on September 14, 2023 and is accepted on October 26, 2023 by the first reviewer. It is commented on October 23, 2023 and is accepted on November 03, 2023 by the second reviewer.

most critical threats and taking steps to address them before an incident occurs. The most effective way to acquire threat intelligence feeds is through automated processes. Structured information is extracted from unstructured textual data by a process called information extraction. Information extraction consists of two consecutive tasks: named entity recognition and relation extraction. Named entity recognition seeks to identify entity mentions and classify them into predefined entity names. Then, with the extracted entities, the relation extraction task aims to detect semantic relations between pairs of entities and categorize them into a predefined set of relation types.

There have been many studies on automated cyber threat intelligence collection and analysis. However, most of them only consider the named entity recognition problem, e.g. [2-8]. Only a few attempts to address the issue of relation extraction [9-11].

Authors in [3-5] proposed CTI (Cyber threat intelligence) systems that collect CTI data from unstructured text. They performed threat analysis based on the collected CTI data. [1] presented an automated technique to extract and validate IOCs for web applications. [2] introduced a method to generate interpretable and accurate IOCs based on the results of running malware samples in a sandbox environment. There are some studies to design a system that extracts threat actions from unstructured CTI reports [6, 7]. To do this, [6] analyzed the grammatical structure of a sentence using dependency parsing, while [7] analyzed the part of speech (POS) containing three elements (subject, verb, object) that make up a threat action.

Our work differs from previous works in that we focus on extracting descriptive or static CTI data other than IOCs from unstructured text. Our work also considers the behavioral relationships between CTI data.

Although the data source is rich and updated promptly, the data on cyber security threats have not been fully exploited, mainly in the form of raw, unstructured documents, and has not been processed. Meanwhile, information about cybersecurity threats is often processed and stored through data standards to store entities and relationships between entities. A widely used data standard in the field is the STIX (Structured Threat Information Expression) data standard. Information, according to the STIX standard, is represented as objects and the relationships between them. This information can also be visualized graphically, allowing STIX-compliant data to be clearly represented. The relation extraction is an important task to show the relationship between named entities in the text. Deep learning has made recent breakthroughs to solve the relational extraction problem.

In fact, according to the STIX data standard with information about network security threats, there can be more than one relationship between two entities. This means that when we examine entities in network threat data, they can interact and connect in various ways. For example, one entity may have a “related to” relationship with multiple other entities, or there may exist multiple types of different relationships, such as “performed”, “related to” or “occurred at”. The complexity in these relationships between entities can vary depending on the context and the type of information under consideration. Understanding and extracting these relationships can help create a more intricate network of connections between entities in network threat information. This can assist researchers and security professionals in gaining a better understanding of how various elements come together in threat situations and enhance the ability to detect and respond to network threats more effectively. Due to these limitations, current studies have not provided sufficient information to help restructure the text and support risk warnings. Based on the foundation of simpler models and using a data set about the relationship between two certain entities, we propose a model that is consistent with data standards in the field of Information on Cybersecurity Threats.

Problem Definition: The task of Extracting Multiple Relations between Entities from unstructured text in the domain of CTI involves

identifying and categorizing various relationships between named entities, which may include but are not limited to threats, actors, and targets. These relationships can be diverse, encompassing multiple types such as “related to”, “performed”, “occurred at” and others. The challenge is to extract these relationships comprehensively, considering the context, and represent them in a structured manner that conforms to data standards, like STIX. The ultimate goal is to create a more detailed network of connections between entities in CTI to enhance threat detection and response capabilities.

Our contribution is multiple. First, we develop a dataset of relationships between pairs of labeled STIX standard entities for testing and evaluation purposes. Then, we propose a deep learning model to extract information about cyber security threats, with the task of extracting two possible relationships on a pair of entities based on the entities collected from a sub-problem - entity recognition.

The paper will consist of 5 sections. In Part II, we will discuss the System Model of Named Entity Recognition. Part III will detail the proposed architectural model, while Part IV will cover the experiments. The final section will provide the conclusion.

II. System Model

Named Entity Recognition (NER) is responsible for assigning entities with their respective labels. Through word classification in the text, NER is the necessary first step for information extraction.

A. Bi-LSTM

In our entity recognition model, we employ a Bi-LSTM architecture. This architecture comprises several essential components, including an input layer, a character embedding layer, a word embedding layer, a word-level Bi-LSTM network, a text-level Bi-LSTM network, and an output layer. This entire system is visually represented in Figure 1**.**

When we feed a piece of text into our model, it undergoes a multi-step transformation. First, the text is divided into both word and character tokens. These tokens then pass through two layers: the word embedding layer and the character embedding layer. These layers convert the tokens into numeric vectors that the model can work with.

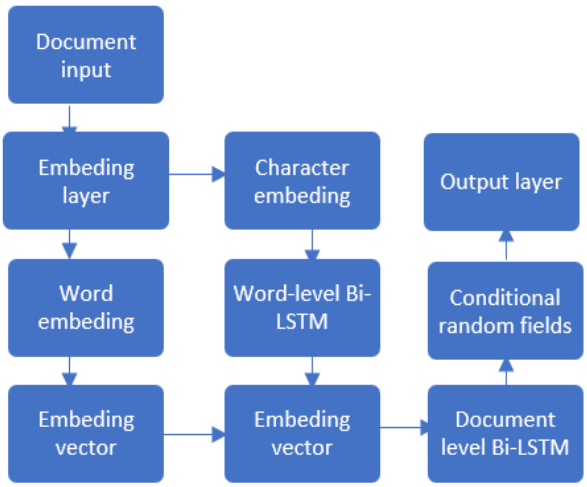

Figure 1. Entity recognition model using Bi-LSTM

The character embedding layer plays a critical role. It takes the character sequence that forms a word and processes it in both forward and backward directions using the word-level Bi-LSTM network. The output of this network is a unique representation vector for the word, which combines information from the characters and the word embedding vector generated at the previous word embedding layer.

The matrices formed by these word representation vectors are then handed over to the text-level Bi-LSTM network. This network's job is to predict the labels for each word. However, the model goes a step further by utilizing Conditional Random Fields (CRF) to assess the likelihood that a label corresponds to a particular word.

Finally, the model makes its prediction by selecting the label with the highest score, effectively identifying the entity or label of the word in question. This intricate process allows our system to accurately recognize entities within text.

B. Bi-GRU

In addition, we've implemented an entity recognition model that utilizes Bi-GRU. This

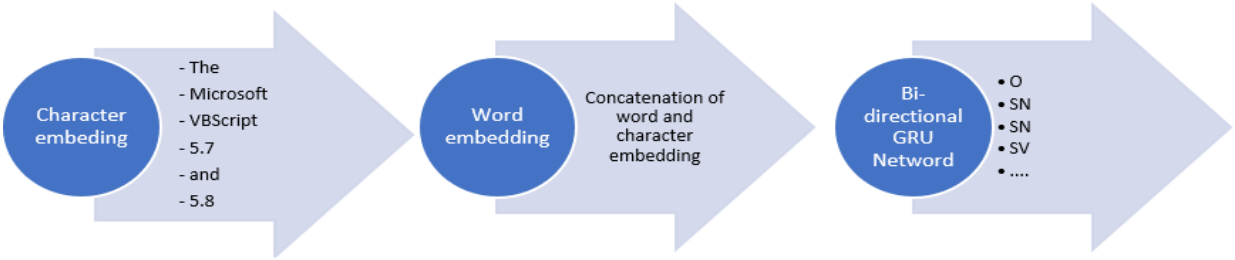

Figure 2. Entity recognition model using Bi-GRU

model follows a structure similar to the one we discussed with Bi-LSTM. It begins by embedding characters and words at both character and word levels. These embeddings are then merged into a vector that serves as a representation for each word. This vector is subsequently fed into a Bidirectional Gated Recurrent Unit (Bi-GRU) network to make determinations about word labels. You can find a visual representation of this system in Figure 2, which illustrates our Entity recognition model using Bi-GRU.

III. SOLUTION APPROACH

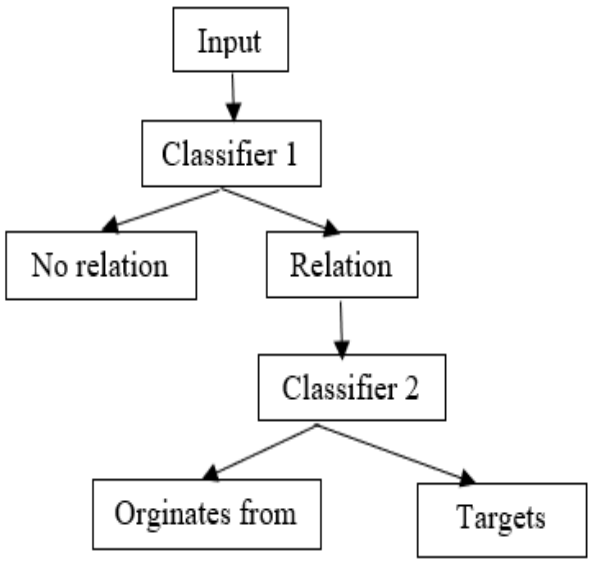

Figure 3. Two-stage relation extraction model

There are two possible ways to solve the problem of multiple relation extraction: (1) The first one, called two-stage extraction, is running the single relation extraction algorithm to extract each relationship one by one; (2) The second, called simultaneous extraction, adjusting the output from 2-label prediction to 3-label prediction. With the first method, a pair of entities can be both assigned to this relationship and assigned to the other, leading to misclassification (reducing Precision, leading to a decrease in F1 score). With the second method, we need to use a multi-class classification algorithm, but not all deep learning algorithms allow the use of multi-class classification, only binary classification. The multi-class classifier can also be less sensitive (low Recall, leading to a low F1 score), or less accurate because the difference in probability between the prediction labels is not as large as the 2-class classifier, so the classification can be more falsified.

From there, the idea arises to use the classification method in 2 stages (see Figure 3): the stage of classification with relation and without relation; and the stage of classifying pairs of related entities into specific relations. The advantages of this idea are:

- It is possible to apply a binary classifier that is not capable of converting to a multiclass classifier.
- It is possible to take advantage of binary classification with large probability differences, making classification decisions more certain and reducing the possibility of misclassification.

Specifically, we create a labeled dataset of relationships between specific entity pairs suitable for the field of Information on Cybersecurity Threats, ideal for future studies on relation extraction in this area. The labeled dataset consists of 2129 sentences, with three labels: “No relation”, “Originates from” and “Targets” (2). Propose a two-stage framework model that allows the extraction of any relationship, be it many possible relationships

between different pairs of entities or double relationships within the same entity pair.

The model has the task of recognizing the names of entities as input for the problem of extracting two possible relations of the HackOrg and Area pair of entities (the entity "Organization that caused the attack" and the entity "Region") is the relation Originates from (relationship "Derived from") and relation Targets (relation "Attack on"). The input to the model is the unstructured text about the cybersecurity threat, and the model's output is the relationship of a pair of entities (HackOrg, Area).

IV. EXPERIMENTS

A. Dataset

We use the dataset in [16] made for studies on entity recognition. We proceed to preprocess and label the data to have a separate dataset for the relational extraction problem. The obtained dataset has many entities that match the STIX data standard, suitable for training a model specific to the Cybersecurity Threat Information field. At the same time in this data set, there are two entities, HackOrg and Area, which are equivalent to two entities Threat Actor and Location in STIX standard, which is a pair of entities defined with two possible relations that the goal of this research project are aiming for a solution.

The data before being fed into the learning and prediction models are all converted to lowercase, removing associations and special characters, leaving only characters commonly used in entities like ".", ":", "-", and "_".

The original labeled dataset for entity recognition contains 175220 words. The labels cover 13 entities: HackOrg, OffAct, SamFile, SecTeam, Tool, Time, Purp, Area, Idus, Org, Way, Exp, and Features (see Table 1). We use 85% of the data for training and 15% for evaluation.

The entities HackOrg and Area can be considered equivalent to the objects Threat Actor and Location respectively in the STIX standard. Between these entities, there can be two possible relationships: Originates from and Targets. We select the sentences containing the HackOrg and Area entity pairs and then label them as having either Non or Originates from or Targets relationship, numbered 0, 1, and 2 respectively. The dataset for the relation extraction task has 2129 sentences (See Table 2), of which 85% is used for training and 15% for evaluation.

TABLE 1. DATASET FOR NAMED ENTITY RECOGNITION IN THREAT INTELLIGENCE

Entity	**Characteristic**	
	Amount	**Ratio**
HackOrg	5465	15,01%
OffAct	2669	7,33%
SamFile	2400	6,59%
SecTeam	1921	5,28%
Tool	4784	13,13%
Time	2659	7,30%
Purp	2424	6,66%
Area	3447	9,47%
Idus	2136	5,87%
Org	2489	6,84%
Way	2018	5,54%
Exp	1559	4,28%
Features	2441	6,70%

TABLE 2. RELATION DATASET FOR HACKORG AND AREA ENTITIES

Relation	**Characteristic**		**Note**
	Amount	**Ratio**	
Originates from	706	33,16%	
Targets	1294	60,77%	
None	129	6,07%	No relationship

B. Experimental results

The models are run experimentally on Google Colab with 13GB RAM configuration, Nvidia K80 or T4 GPU. The word embedding model used in the two models is GloVe.

Tables from 3 to 6 show the parameters used for the training phase and the results obtained in the evaluation phase for entity recognition using the Bi-LSTM and Bi-GRU models.

Tables 7 and 8 describe the parameters used in the training phase and present the evaluation

results of the different options for relation extraction. We compare the performances of the proposed two-stage relation extraction method with the simultaneous extraction alternative and also with the trivial solution of applying the two-class classifier twice, once for recognizing each relation independently. We call this trivial solution dual extraction.

TABLE 3. PARAMETERS FOR THE TRAINING PHASE OF ENTITY RECOGNITION USING BI-LSTM

Parameter	**Value**	**Note**
optimization	Adam	The optimization algorithm used to find a minimum of the loss function and increase the accuracy of the model
learning rate	0,001	A tuning parameter to determine the step size at each iteration while finding a minimum of the loss function
n_epochs	100	Number of iterations to be made over all the training data
batch_size	128	Number of samples used in each iteration
cell_char	100	Number of characters in each cell of the neural network
cell_word	100	Number of words in each cell

TABLE 4. RESULT OF ENTITY RECOGNITION MODEL USING BI-LSTM

Embedding dimension (dim_char)	**F1 score**
100	0,7748
300	0,7629

TABLE 5. PARAMETERS FOR TRAINING THE ENTITY RECOGNITION MODEL USING BI-GRU

Parameter	**Value**
optimization	Adam
learning rate	0,001
n_epochs	100
cell_char	100
cell_word	100

In Table 9 we give the overall results when the ratio of Originates from and Targets relations measured by the number of sentences in the dataset is 0.353:0.647. The precision and recall are both calculated in a weighted fashion, according to the proportion of sentences of each relation in the dataset. The precision and recall obtained when predicting three classes in the two-stage extraction method as well as when predicting two classes in the single extraction model are calculated from the individual results of each of two two-class classifiers.

TABLE 6. RESULTS OBTAINED FROM THE ENTITY RECOGNITION MODEL USING BI-GRU

embedding_size	**batch_size**	**F1 score**
100	128	0,7558
100	256	0,7148
200	128	0,6228

TABLE 7. PARAMETERS FOR THE TRAINING PHASE OF RELATION EXTRACTION

Parameter	**Value**	**Note**
optimization	Adam	
learning rate	0,001	
epochs	30	
gru_size	230	Size of gated recurrent unit (GRU) cell in the recurrent neural network
keep_prob	0,5	Used to control the dropout rate [10] when training the neural network
char_embedding_dim	50	Number of dimensions to use for character embeddings

With the problem of entity recognition, the experimental results of this work in both models show better results than the results of other works, such as [16], with F1-score of 0.7129. However, compared with the results of the two

studies with the model we use, the common F1-score of the labels is 0.92 in the Bi-LSTM model and 0.9872 in the Bi-GRU model, then the results are not good. It is because the number of entity names in the dataset used in this work is extensive, with 13 labels, and the number of entities is not uniform, it can significantly affect the measurement results. In addition, the model can predict well with entities with a clear structure, such as Time or Area. Still, poorly with complex entities, without particular structures and with a small number of samples like Purp or Features, for example, the phrase "military aviation capabilities" is the Purp entity, or "contain information" is the Features entity.

TABLE 8. RESULTS OBTAINED FROM THE RELATION EXTRACTION METHODS

Method	**Relation**	**Efficiency**	
Two-stage (2 layers)	Have relation	F1 score Precision Recall	0,4906 0,4906 0,4906
Two stage extraction (2 layers)	Originate from or Targets	F1 score Precision Recall	0,9033 0,9010 0,9066
Simultaneous extraction (3 layers)	Originates from	F1 score Precision Recall	0,8780 0,9230 0,8372
Simultaneous extraction (3 layers)	Targets	F1 score Precision Recall	0,9147 0,9076 0,9218
Dual extraction (2 layers)	Originates	F1 score Precision Recall	0,9139 0,9124 0,9156
Dual extraction (2 layers)	Targets	F1 score Precision Recall	0,8781 0,8756 0,8819

With the relation extraction problem, the experimental results of this work are relatively lower than the results in the study containing related models, with F1-score of 0.989 with 01 relations. Because of the complexity of the sentence structure in the Security Threat Intelligence domain, the lack of contextual clarity between the sentences with the Originates relation and the Targets relation makes it difficult for the model to distinguish these two labels from each other. Some other studies using HAN model give better results when applied on a simple data set with a single relationship between two entities. In addition, the results show that the prediction by the two-stage method gives better overall results than the simultaneous method and the dual extraction method, with the F1-score of these three methods being 0,9068, 0,9023, and 0,8926, respectively. Thus, the two-stage extraction gives the best results of the three methods on the data set with two relations Originate from and Targets. Besides, this extraction method can also take advantage of the available binary extraction algorithm without requiring modification or conversion into a multi-class extraction algorithm.

TABLE 9. OVERALL RESULTS OF THE WEIGHTED RELATION EXTRACTION METHODS

Method	**F1 score**	**Precision**	**Recall**
extraction	0,9068	0,9078	0,9065
Simultaneous Extraction	0.9023	0.913	0.8919
Dual Extraction	0.8926	0.8917	0.8937

V. CONCLUSION

In this work, we labeled data and tested relational extraction methods to apply in the information about security threats field with the subproblem of entity identification as input for the relation extraction. With the goal of proposing and comparing possible double-relationship extraction methods between two certain entities, this work has concluded that the result from the proposed two-class extraction method is the best among the two classes. The method were tested against the dataset for the dual relationship Originates from – Targets with the HackOrg - Area entity pair. The F1-scores obtained from the two-stage extraction method, the simultenous method, and the dual extraction are 0.9068, 0.9023, and 0.8926, respectively.

These results show that our approach provides a realistic manner to assess the threats and vulnerabilities early.

REFERENCES

[1] Wiem Tounsi and Helmi Rais. A Survey on Technical Threat Intelligence in the Age of Sophisticated Cyber Attacks. *Computers & Security*, 2017.

[2] Robert A. Bridges, Corinne L. Jones, Michael D. Iannacone, Kelly M. Testa, and John R. Goodall. Automatic Labeling for Entity Extraction in Cyber Security. In *Proceedings of the 2014 ASE International Conference on Cyber Security*, 2014.

[3] Nuno Dionísio, Fernando Alves, Pedro M. Ferreira, and Alysson Bessani. Towards End-to-End Cyberthreat Detection from Twitter Using Multi-Task Learning. In *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*, 2020.

[4] Nuno Dionísio, Fernando Alves, Pedro M. Ferreira, and Alysson Bessani. Cyberthreat Detection from Twitter Using Deep Neural Networks. In *Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN)*, 2019.

[5] Ghaith Husari, Ehab Al-Shaer, Mohiuddin Ahmed, Bill Chu, and Xi Niu. TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources. In *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC '17)*, 103–115, 2017.

[6] Gyeongmin Kim, Chanhee Lee, Jaechoon Jo, and Heuiseok Lim. Automatic Extraction of Named Entities of Cyber Threats Using a Deep Bi-LSTM-CRF Network. *International Journal of Machine Learning and Cybernetics*, 11, 2341-2355, 2020.

[7] Xuren Wang, Xinpei Liu, Shengqin Ao, Ning Li, Zhengwei Jiang, Zongyi Xu, Zihan Xiong, Mengbo Xiong, and Xiaoqing Zhang. DNRTI: A Large-Scale Dataset for Named Entity Recognition in Threat Intelligence. In *Proceedings of the 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2020.

[8] Jun Zhao, Qiben Yan, Jianxin Li, Minglai Shao, Zuti He, and Bo Li. TIMiner: Automatically Extracting and Analyzing Categorized Cyber Threat Intelligence from Social Data. *Computers & Security*, Volume 95, 2020.

[9] Hyeonseong Jo, Yongjae Lee, and Seungwon Shin. Vulcan: Automatic Extraction and Analysis of Cyber Threat Intelligence from Unstructured Text. *Computers & Security*, Volume 120, Issue C, 2022.

[10] Aditya Pingle, Aritran Piplai, Sudip Mittal, Anupam Joshi, James Holt, and Richard Zak. RelExt: Relation Extraction Using Deep Learning Approaches for Cybersecurity Knowledge Graph Improvement. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '19)*, 879-886, 2019.

[11] Xiaojing Liao, Kan Yuan, XiaoFeng Wang, Zhou Li, Luyi Xing, and Raheem Beyah. Acing the IOC Game: Toward Automatic Discovery and Analysis of Open-Source Cyber Threat Intelligence. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*, 755–766, 2016.

[12] MITRE Corporation, Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX™), 2012.

[13] Nadeesha Perera et al., "Named Entity Recognition and Relation Detection for Biomedical Information Extraction", Front. Cell Dev. Biol., 2020.

[14] Pennington et al., "GloVe: Global Vectors for Word Representation," inProc. of the Empirical Methods in NaturalLanguage Processing, 2014.

[15] Awais Ahmed Shujrah et al., "Measurement of E-Learners' Level of Interest in Online Course Using Support Vector Machine", Indian Journal of Science and Technology, 2019.

[16] N. Srivastava et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research, 2014.

[17] Ashish Vaswani et al., "Attention Is All You Need", 31st Conference on Neural Information Processing Systems (NIPS 2017), 2017.

[18] Haoyu Wang et al., "Extracting Multiple-Relations in One-Pass with Pre-Trained Transformers", Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019.

[19] Kyle Wilhoit, Joseph Opacki, Operationalizing Threat Intelligence, Packt Publishing, 2022, pp. 3-5, 11, 16, 36, 317.

[20] Zichao Yang et al., "Hierarchical attention networks for document classification", InProc. of NAACL-HLT 2016, 2016.

[21] Ningyu Zhang et al., "Attention-Based Capsule Networks with Dynamic Routing for Relation Extraction", 2018.

ABOUT THE AUTHORS

Nguyen Dai Tho

Workplace: University of Engineering and Technology, Vietnam National University, Hanoi.

Email: nguyendaitho@vnu.edu.vn

Education: He received his engineer's degree from the Hanoi University of Science and Technology in 1995; his master's degree from the Francophone Institute of Computer Science (IFI) in 1997, and his PhD degree from the University of Technology of Compiègne, France in 2000.

Recent research interests: Information security; Computer networks; Distributed computing.

Cơ quan làm việc: Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.

Email: nguyendaitho@vnu.edu.vn

Quá trình đào tạo: Nhận bằng Kỹ sư tại Đại học Bách Khoa Hà Nội vào năm 1995; Thạc sĩ tại Viện Khoa học máy tính Pháp ngữ (IFI) vào năm 1997; Tiến sĩ tại Đại học Công nghệ Compiègne, Pháp vào năm 2000.

Hướng nghiên cứu hiệ nay: Bảo mật thông tin; Mạng máy tính; Phân phối máy tính.

Nguyen Trung Hieu

Workplace: People's Security University.

Email: hieutn2709@gmail.com

Education: He completed his MSc degree in Information communication technology from LaTrobe University, Australia.

Recent research interests: Machine learning; Deep learning; Network security.

Cơ quan làm việc: Học viện An ninh nhân dân.

Email: hieutn2709@gmail.com

Quá trình đào tạo: Thạc sĩ Công nghệ thông tin truyền thông tại Đại học LaTrobe - Úc.

Hướng nghiên cứu hiện nay: Học máy; Học sâu; An toàn mạng.

Tran Nguyen Ngoc Bao

Workplace: Money Forward Vietnam.

Email: trannguyen61st@gmail.com

Education: She received her bachelor's degree from the University of Engineering and Technology, Vietnam National University, Hanoi in 2022.

Recent research interests: Deep learning; Information security; Software engineering.

Cơ quan làm việc: Tập đoàn Money Forward Việt Nam.

Email: trannguyen61st@gmail.com

Quá trình đào tạo: Cử nhân tại Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội vào năm 2022.

Hướng nghiên cứu hiện nay: Học sâu; An toàn thông tin; Kỹ thuật phần mềm.

Nguyen Phuong Anh

Workplace: FPT University.

Email: anhnp75@fe.edu.vn

Education: She received her PhD degree from the University of Lorraine, France.

Recent research interests: Information security; Artificial intelligence for cybersecurity.

Cơ quan làm việc: Đại học FPT.

Email: anhnp75@fe.edu.vn

Quá trình đào tạo: Tiến sĩ tại Đại học Lorraine, Pháp.

Hướng nghiên cứu hiện nay: Bảo mật thông tin; Trí tuệ nhân tạo cho an ninh mạng.