A Survey on Detection of cybersecurity threats on Twitter using deep learning

Omar Alsodi
The School of Business
University of Southern Queensland
Brisbane, Australia
u1089882@umail.usq.edu.au

Xujuan Zhou
The School of Business
University of Southern Queensland
Brisbane, Australia
xujuan.zhou@usq.edu.au

Raj Gururajan
The School of Business
University of Southern Queensland
Brisbane, Australia
Raj.Gururajan@usq.edu.au

Anup Shrestha
The School of Business
University of Southern Queensland
Brisbane, Australia
Anup.Shrestha@usq.edu.au

Abstract— **In these times of increasing cybersecurity threats, monitoring and analysing cybersecurity events in a timely and effective way is the key to promote social media security. Twitter is one of the world's widely used social media platforms where users can share their preferences, images, opinions, and events. The Twitter platform can promptly aggregate cyber-related events and provide a source of information about cyber threats. Likewise, Deep Learning can play a critical role to help social media providers achieve a more accurate assessment of cybersecurity threats. In this paper, we have reviewed various threats and discussed deep learning techniques to detect cybersecurity threats on Twitter.**

Keywords: Cyber security, Deep Learning, Twitter, Cybersecurity threats, social media

I. Introduction

Social media comprises online services that enable users to establish a public profile and connect with a list of other users to view and share their profiles and content. The application? of social media links is different from one service to another [1]. Twitter is one of the social media platforms that helps users to discover and extend their networks with new followers. An important characteristic of any social media is 'data sharing' wherein users can share their preferences, videos, images and events. Twitter is currently a dominant means of communication for billions of Internet users [2]. However, Twitter continues to offer a rich environment for illegal activities that relate to cybersecurity incidents [3]. Cyber threats can occur at any point, and severely affect peoples' lives and trigger social and economic disruptions [4]. Lately, the cybersecurity threats have expanded impacting users of all social media platforms including Twitter. Many studies have attempted to propose solutions for identifying and detecting the anticipated threats in cybersecurity. These efforts have resulted in the extraction of helpful data that identifies expected cybersecurity threats. With the advent of Artificial Intelligence (AI), Deep Learning techniques have emerged that play a critical role to help social media platforms achieve a more accurate detection and assessment of cybersecurity threats.

This paper presents the relevant cybersecurity threats in Twitter and Deep Learning techniques to detect cyber threats in Twitter. The rest of this paper is organized as follows: Section II described the cyber security threats in Twitter. Section III reviewed deep learning as a technique to detect these threats and incidents in Twitter. Ssection IV discussed the challenges and opportunities for cyber security threats using deep learning. Finally, section V concluded the review and offered future research directions.

II. CYBER SECURITY THREATS IN TWITTER

A. Cybersecurity threats

Cybersecurity threats occur more frequently with the popularity of today's use of Twitter. Consequently, these threats may seriously impact the lives of individuals and cause social and financial unrest. Researchers have been using Twitter at least since 2010 as an extensive, publicly available database for analyzing and extracting cybersecurity threats.

B. Types of cybersecurity threats

Various types of social media platforms report several attacks against them. Many of the attacks are targeted at stealing the identity of users or undermining the privacy and trust of the network, such as hijacking, identity theft, spamming, social phishing, malware attacks, facial image retrieval and analysis, impersonation, fake requests and Sybil attacks. Cybersecurity Threats are carried out as attacks on Twitter for a wide range of motivating factors, that include political, emotional, financial, entertainment, ideological, personal, cyber warfare, or commercial purposes [7].

- Spamming

Attackers in a spam attack send unsolicited messages (spam) in bulk to internet users. In Twitter, this type of attack appears to be more successful compared to traditional spamming attacks where email is used to spread spam. This is due to the ease of exploiting social connections among Twitter users where the targeted user can be easily persuaded to read and trust the spam information to be safe [5].

- Identity theft

The concept of "Identity theft" can be understood in a variety of ways, but all of them boil down to one basic definition which states that it is illegal or unauthorized use of personal information belonging to someone else for one's own gain [6].

- Malware

Among social networking platforms, circulation of malware is getting more widespread. The attackers send the scripts for the inserted malware to the legitimate user. A virus may be installed on the attacker's computers by clicking the malicious URL, or it may redirect to a fake website that tries to steal personal information from the target user [7].

- Sybil threats

With this type of attack, Sybil attack and false account, criminals create many false identities that assist them in the distributed system and peer-to-peer system to obtain unauthorized benefits. A Sybil attack is a major concern for Twitter security because it involves a large number of users connecting as peers to a peer-to-peer network, which means that one online person can manage and maintain multiple fake identities in the Twitter environment [5]. For example, by proclaiming itself-using the keyword "best", an attacker may falsely increase the reputation and popularity of a Twitter account. A Sybil attack can decrease credibility values, corrupt data, and outvote genuine Twitter users [8].

- Hijacking

This refers to the benefit of control over the social media profile of someone else. If an attacker can crack the authentication details of an account, the attacker succeeds in hijacking a legitimate profile. Using typical passwords are also a bad choice, as they increase the risk of dictionary attacks that can hijack such passwords. Strong passwords that are regularly updated provide a good protection against Hijacking [7].

- Face image retrieval and analysis

Twitter users also appear to add photos to their individual profiles that can be used to recognize the corresponding holders of the profile. Thus, an outsider may use this data source to compare profiles across services using facial recognition, which is part of the wider challenge raised by so-called mashups. Facial recognition can be used to link image instances (and accompanying information) across services and websites that, in turn, allow connectivity, for example, a pseudo-anonymous dating profile with a corporate website profile defined. As a result, the adversary will collect considerably more knowledge about the user than is expected [9].

- Fake requests

Using their own account, the attacker sends a fake request to extend their network. If users accept false requests, the suspect will have more privileges and even more data from the victim reports. It is not possible to completely avoid such fraudulent requests, therefore social network platforms should be more accountable to users [7].

- Social phishing

Phishing is a type of deception in which the perpetrator tries to extract confidential information from the victim by impersonating a trustworthy individual. Phishing attacks usually employ generic lures. For example, a phisher that is misinterpreted as a large banking company or a successful online auction site would have a fair return, despite little or no knowledge of the recipient [10].

- Impersonation

In this case, the attacker's aim is to create a fake account to impersonate a real-world individual successfully. This threat depends heavily on the methods of authentication faced by users when registering to build a new account. These threats may cause the entity that is being impersonated to suffer severe damage [7].

TABLE 1: COMPARISON OF MOST POPULAR TYPES OF CYBER SECURITY THREATS ON TWITTER [7]

Threats	Impact on user	Effectiveness of user side protection mechanism	Effectiveness of server-side protection mechanism	Threat to data integrity	Threat to data privacy
Spamming	Small	Poor	Strong	No	No
Identity theft	Average to high	Poor	Poor	Yes	No
Malware	High	Medium	Medium	Yes	Yes
Sybil threats	Average	Poor	Strong	Yes	No
Hijacking	High	Poor	Poor	Yes	Yes
Image and Face retrieval	Average to high	Medium	Medium	No	Yes
Fake requests	Low	Strong	Poor	No	Yes
Social phishing	High	Strong	Poor	Yes	Yes
Impersonation	High	Poor	Poor	Yes	Yes

III. DETECTION OF CYBER SECURITY THREATS IN TWITTER USING DEEP LEARNING

A. Deep learning

Deep learning is a type of machine learning method that allows machines to learn from their mistakes and comprehend the world as a hierarchy of concepts [11]. Deep learning enables computational models consisting of several layers of processing to learn data representation at multiple abstraction levels. These methods have greatly improved state-of-the-art speech recognition, visual object recognition, object detection, and many other domains such as drug discovery and genomics [12]. The use of deep learning technology for cybersecurity research and intrusion detection is extremely important since most attacks use invasive software families that can be detected and classified [13]. Deep learning is commonly used in pattern recognition. Furthermore, the issue of classification, such as text classification and image classification has also shown efficiency using Deep learning [4].

Convolutional Neural Networks (CNNs).

ConvNets is designed to process data that come in the form of multiple arrays, such as a colour image consisting of three 2D pixel-intensity arrays in three colour channels. There are many data modalities in the form of multiple arrays: 1D for signals. Sequences, like language; 2D images or audio spectrograms; and 3D images, either video or volumetric. The four key ideas behind ConvNets that take advantage of the characteristics of natural signals are: local connections, shared weights, pooling, and the use of multiple layers [14]. Convolutional networks integrate three architectural ideas to ensure a certain degree of transition, size, and distortion invariance: 1) local receptive fields; 2) shared weights (or duplication of weights), and 3) spatial or temporal subsampling [15].

The authors of [16] also suggested CNN to strive for image recognition. The basic idea of CNN is to capture a data function by transferring the kernel, a convolution matrix, to a region in the image. Generally, while neural networks cannot retain spatial information in the image, CNN can maintain it by adding the kernel to each area of the image. In the case of natural language processing (NLP), we can also add the convolutional layer of CNN to the vector space translated from the text corpus. Since each kernel can learn how to insert in a region i.e., one sentence in the NLP and capture the semantic and structural features of the sentence, CNN performs well in the text classification.

The researchers in [4] proposed a multi-task learning approach based on the natural language processing technology and machine learning algorithm of the Iterated Dilated Convolutional Neural Network (IDCNN) and Bidirectional Long Short-Term Memory (BiLSTM) to establish a highly accurate network model. Their results show that the proposed model operates well to predict cyber hazard incidents from tweets and greatly outperforms a variety of baselines.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory Networks (LSTM)

A recurrent neural network (RNN) is a recurrent structure where a directed graph along a chain is generated by node associations. This helps RNN to view time dynamic behaviour for a time series applied to natural language processing (NLP). RNNs can use their internal state to process input agreements and may do so only for a limited period, i.e., they cannot remember long-term information [17]. In other words, RNN is a neural network that simulates a complex system of discrete time that has an input x_t, an output y_t, and a hidden state h_t. The subscript $_t$ represents time in our notation. RNNs have a very elegant way of dealing with sequential (time) data that embodies connections between data points like the sequence [18].

The researchers in [19] have proposed recurrent neural network (RNN) for sequential data processing such as voice and text processing. The defining characteristic of RNN, which is distinct from that of RNN, the general neural networks are the introduction of the hidden state vector. The secret state represents the description of the previous input data which is modified once the new input is reached. Finally, after processing all input results, the secret state is the summarization of all sequences, which is somewhat like the processing of a sequence done by a human being. Of course, RNN has the benefit of reading sentences that are read by a human. However, as the layer deepens, gradient explosions and vanishing problems occur, which can degrade performance [39]. To avoid this problem, the long-term memory (LSTM) technique has been proposed in [20]. In order to prevent the gradient from bursting and disappearing problems, LSTM adds the cell state to change the previous knowledge. LSTM has been commonly used for text classification because it can learn high-level representation using a deeper layer due to the cell status while maintaining the sequence of representations given by RNN. The LSTM has been applied to the emotion classification of short texts on social media [21]. In [22], the authors suggested a densely connected Bi-LSTM composed of several Bi-LSTM layers, which shows improved efficiency than Bi-LSTM.

Deep Neural Networks (DNN)

A neural network can be a deep neural network (DNN) with many layers that make it very mind-boggling. DNN contains one layer of data, at least one hidden layer, and one layer of output. A rectilinear unit (ReLU) is contained in a hidden sheet. ReLU is a mechanism for activation which has specified the positive part of its argument. It has fewer gradient problems and is efficient in terms of computation. As each neuron in a single layer relates to each neuron in the next layer, the secret layer is also called a fully linked layer [17]. A typical neural network (NN) consists of several neuron-called simple, interconnected network processors, each producing a series of activations of real value. Sensors that sense the environment activate input neurons, and weighted connections from previously active neurons trigger other neurons [23]. Dionísio, Alves, Ferreira, and Bessani [24] presented a new tool for analysing information obtained from Twitter using deep neural networks to process cybersecurity threats.

B. Cyber security threat detection on Twitter

Twitter is one of the world's widely used online services with over 340 million users that includes 186 million monetizable Daily Active users. Over 500 million tweets are sent per day by Twitter users. A variety of public events are reported on the web every day as a social network and microblogging tool. People also post online risks in their tweets, including zero-day attacks, malware, data breaches, security violations, DDoS, bugs, and more [25]. Researchers have been using Twitter as a robust data source platform to analyse and extract incidents. Phuvipadawat and Murata extract the attributes of the hashtags, fans, and timestamps from the posts to detect the important news activity [26]. The authors of [27] investigated on detecting these events by analysing the text stream in Twitter and attempting to solve these problems with EDCoW (Event Detection with Clustering of Wavelet-based Signals) that are mainly focused on Tweet trends and the frequency of hashtags [28]. Verma and Sofat provided an overview of the techniques, functionality, detection rate, and limitations used to detect spam profiles, primarily on Twitter [29]. Likewise, the authors of [30] studied the detection of incidents dependent on timestamps, geolocations, and cluster sizes. Rao, Kamhoua, Njilla, and Kwiat [31] reviewed the prevalent methods for identifying spam, spammers, cyber criminals, malicious material, and suspicious users on Twitter, and provided a unified framework for modelling and reasoning on the veracity of Twitter posts using advances in statistical relational learning [31]. Similarly, the researchers [32] studied the detection of incidents by analysing the dataset tags. Likewise, in [33], the authors focused on a supervised learning approach by proposing a novel multitask learning-based model to detect cyber threat event on tweets datasets. Nazir et al. [34] combined the moving threshold average algorithm with the Gaussian tweet sentiment signal detection algorithm and the top hashtag. The authors of [35] have also used deep learning techniques to detect cybersecurity threats.

Fang et al. [4] proposed a multi-task learning approach based on the Iterated Dilated Convolutional Neural Network (IDCNN) and Bidirectional Long Short-Term Memory (BiLSTM) natural language processing technology. Machine learning algorithm was presented to set up a highly accurate network model [4]. In [36], the authors used locality-sensitive hashing to roughly find related items and incremental clustering to implement a realistic, real-time event detection algorithm. Researchers are trying to define the features of tweets and use suitable algorithms to solve the problems they are researching. Rodriguez and Okamura [37] show a framework for classifying OSINT data into cyber-security-related to be introduced and analysed, and the accuracy of those data was subsequently improved using an unsupervised method.

Table 2 summarizes the comparison of previous studies on the detection of cybersecurity threats in Twitter. All previous studies indicate the focus on the usage of the effective classifiers to improve the detection accuracy.

TABLE 2: COMPARISON OF PREVIOUS STUDIES ON THE DETECTION OF CYBERSECURITY THREATS IN TWITTER

Study	Focus	Methodology	Data Sets
[38]	- Events from Twitter that requires only minimal supervision - DoS attacks, data breaches, and account hijacking.	Weakly supervised learning	Tweets containing “DDoS”
[39]	An automatic, self-learned framework that can detect, geolocate, and categorize cybersecurity events in near-real time over the Twitter stream	First story detection	Streaming Tweets
[31]	Machine learning techniques by considering user behaviour, content of tweets, social relationships, etc., to detect different types of cyberthreats.	Social KB	- Tweets containing “URLs” - Streaming Tweets
[40]	A novel application of NLP models to detect denial of service attacks using only social media as evidence.	Basic neural network	Tweets written on attack day
[32]	Treat the event detection problem in a multimodal twitter hashtag network.	Expectation-maximization (EM) algorithm	- Tweets containing hashtag
[24]	a novel tool that uses deep neural networks to process cybersecurity information received from Twitter.	SVM, MLP, CNN, BiLSTM	Tweets filtered by keywords
[41]	Collection method of Cyber threat tweets	Centroid, One-class SVM, CNN, LSTM	Streaming Tweets
[42]	a multi-task learning approach combining two Natural Language Processing tasks for cyberthreat intelligence.	Multi-Task Learning (MTL)	Streaming Tweets
[43]	a novel word embedding model, called contrastive word embedding, that enables to maximize the difference between base embedding models.	CNN, RNN and LSTM	Curated data, OSINT data, and background knowledge.
[4]	- Detection of cyber threat events on tweets. - Named Entity Recognition (NER) for tweets	Multitask learning NLP, IDCNN, BiLSTM	Streaming Tweets

IV. CHALLENGES AND FUTURE WORKS FOR DETECTON OF CYBER SECURITY THREATS

As cyber threats increase security risks, numerous researchers and security firms have been developing several solutions. Watermarking [44], Steganalysis [45], and digital oblivion [46] are some of the solutions for protecting social network users against threats from compromised multimedia data. While traditional solutions such as spam detection [47] and phishing detection [48] mitigate the conventional risks, there are also some established security solutions such as mechanisms for authentication [49] and privacy settings [50] as well as commercial solutions such as minor monitor and social protection applications that offer safeguards against cyberthreats in social media platforms. However, it is very challenging to detect cybersecurity attacks in Twitter. There are two primary reasons for this problem. Firstly, since the tweets are limited to 140 characters and the writing patterns of people are flexible, the meaning and context of words are diverse and different phrases and term shifts are varied [52]. Secondly, there are too diverse and confounding contents of advertisement tweets and people misuse hashtags in their posts to get likes and attention. Due to these reasons, it is extremely difficult to detect cybersecurity threats and researchers find it very difficult to automatically detect cybersecurity threats from tweets [4]. Therefore, to succeed in cybersecurity threats detection in Twitter, researchers should explore the option of improving the deep learning model to different types of data in order to demonstrate high accuracy and less time to detect cybersecurity threats.

V. CONCLUSIONS

Twitter offers new engagement and networking possibilities, but also pose new questions about cybersecurity and privacy. Other Users of social networks are responsible for the bulk of 'online risks' resulting from social networking their use of twitter. Connection to knowledge-rich information about individual users is created and this can be accumulated from their Twitter usage. In unforeseen or unconsidered circumstances, this information can provide the ingredients for criminal activities, such as identity theft and related fraudulent activities. In this respect, Twitter users of social media sites may be typically at a higher risk of attacks from social engineering. By its essence, the social networking platform is rich in social interaction and association. Such virtual social environment provides a fertile ground for scams and manipulation since Twitter users are more likely to believe fellow users (or messages from people who seem to be acquaintances). Consequently, social networks inadvertently act as a facilitator of more powerful vulnerabilities, and these situations can be exploited by cybersecurity attackers. Twitter users should be aware of the threats to their personal and financial details and should be able to behave responsibly. Deep learning technologies, which have been applied in various domains such as medicine and healthcare [53], [54], [55], can play a critical role to help cybersecurity experts achieve a more accurate detection and assessment of cybersecurity threats based on twitter data.

REFERENCES

[1] Boyd, D.M. and Ellison, N.B., 2007. Social network sites: Definition, history, and scholarship. Journal of computer-mediated Communication, 13(1), pp.210-230.

[2] Rathore, S., Sharma, P.K., Loia, V., Jeong, Y.S. and Park, J.H., 2017. Social network security: Issues, challenges, threats, and solutions. Information sciences, 421, pp.43-69.

[3] Weir, G.R., Toolan, F. and Smeed, D., 2011. The threats of social networking: Old wine in new bottles?. Information security technical report, 16(2), pp.38-43.

[4] Fang, Y., Gao, J., Liu, Z. and Huang, C., 2020. Detecting cyber threat event from twitter using IDCNN and BILSTM. Applied Sciences, 10(17), p.5922.

[5] Rathore, S., Sharma, P.K., Loia, V., Jeong, Y.S. and Park, J.H., 2017. Social network security: Issues, challenges, threats, and solutions. Information sciences, 421, pp.43-69.

[6] Irshad, S. and Soomro, T.R., 2018. Identity theft and social media. International Journal of Computer Science and Network Security, 18(1), pp.43-55.

[7] Zhang, Z. and Gupta, B.B., 2018. Social media security and trustworthiness: overview and new direction. Future Generation Computer Systems, 86, pp.914-925.

[8] Noh, G., Oh, H., Kang, Y.M. and Kim, C.K., 2014. PSD: Practical Sybil detection schemes using stickiness and persistence in online recommender systems. Information Sciences, 281, pp.66-84.

[9] Al Hasib, A., 2009. Threats of online social networks. IJCSNS International Journal of Computer Science and Network Security, 9(11), pp.288-93.

[10] Jagatic, T.N., Johnson, N.A., Jakobsson, M. and Menczer, F., 2007. Social phishing. Communications of the ACM, 50(10), pp.94-100.

[11] Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep learning. MIT press.

[12] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.

[13] [13] Hatcher, W.G. and Yu, W., 2018. A survey of deep learning: Platforms, applications and emerging research trends. IEEE Access, 6, pp.24411-24432.

[14] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.

[15] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), pp.2278-2324.

[16] Xu, H., Dong, M., Zhu, D., Kotov, A., Carcone, A.I. and Naar-King, S., 2016, October. Text classification with topic-based word embedding and convolutional neural networks. In Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (pp. 88-97).

[17] Simran, K., Balakrishna, P., Vinayakumar, R. and Soman, K., 2020, April. Deep learning approach for enhanced cyber threat indicators in Twitter stream. In Proc. Int. Symp. Secur. Comput. Commun. (Vol. 2019, pp. 135-145).

[18] Schuster, M. and Paliwal, K.K., 1997. Bidirectional recurrent neural networks. IEEE transactions on Signal Processing, 45(11), pp.2673-2681.

[19] Elman, J.L., 1990. Finding structure in time. Cognitive science, 14(2), pp.179-211.

[20] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.

[21] Wang, J.H., Liu, T.W., Luo, X. and Wang, L., 2018, October. An LSTM approach to short text sentiment classification with word embeddings. In Proceedings of the 30th conference on computational linguistics and speech processing (ROCLING 2018) (pp. 214-223).

[22] Ding, Z., Xia, R., Yu, J., Li, X. and Yang, J., 2018, August. Densely connected bidirectional lstm with applications to sentence classification. In CCF International Conference on Natural Language Processing and Chinese Computing (pp. 278-287). Springer, Cham.

[23] Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural networks, 61, pp.85-117.

[24] Dionísio, N., Alves, F., Ferreira, P.M. and Bessani, A., 2019, July. Cyberthreat detection from twitter using deep neural networks. In 2019 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

[25] Yagcioglu, S., Seyfioglu, M.S., Citamak, B., Bardak, B., Guldamlasioglu, S., Yuksel, A. and Tatli, E.I., 2019. Detecting cybersecurity events from noisy short text. arXiv preprint arXiv:1904.05054.

[26] Phuvipadawat, S. and Murata, T., 2010, August. Breaking news detection and tracking in Twitter. In 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (Vol. 3, pp. 120-123). IEEE.

[27] Weng, J. and Lee, B.S., 2011, July. Event detection in twitter. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 5, No. 1).

[28] Cordeiro, M., 2012, January. Twitter event detection: combining wavelet analysis and topic inference summarization. In Doctoral symposium on informatics engineering (Vol. 1, pp. 11-16).

[29] Verma, M. and Sofat, S., 2014. Techniques to detect spammers in twitter-a survey. International Journal of Computer Applications, 85(10).

[30] Kaleel, S.B. and Abhari, A., 2015. Cluster-discovery of Twitter messages for event detection and trending. Journal of computational science, 6, pp.47-57.

[31] Rao, P., Kamhoua, C., Njilla, L. and Kwiat, K., 2018. Methods to detect cyberthreats on twitter. Surveillance in Action, pp.333-350.

[32] Yılmaz, Y. and Hero, A.O., 2018. Multimodal event detection in Twitter hashtag networks. Journal of Signal Processing Systems, 90(2), pp.185-200.

[33] Ji, T., Zhang, X., Self, N., Fu, K., Lu, C.T. and Ramakrishnan, N., 2019, August. Feature driven learning framework for cybersecurity event detection. In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 196-203).

[34] Nazir, F., Ghazanfar, M.A., Maqsood, M., Aadil, F., Rho, S. and Mehmood, I., 2019. Social media signal detection using tweets volume, hashtag, and sentiment analysis. Multimedia Tools and Applications, 78(3), pp.3553-3586.

[35] Dabiri, S. and Heaslip, K., 2019. Developing a Twitter-based traffic event detection model using deep learning architectures. Expert systems with applications, 118, pp.425-439.

[36] Sani, A.M. and Moeini, A., 2020, April. Real-time event detection in twitter: A case study. In 2020 6th International Conference on Web Research (ICWR) (pp. 48-51). IEEE.

[37] Rodriguez, A. and Okamura, K., 2020. Enhancing data quality in real-time threat intelligence systems using machine learning. Social Network Analysis and Mining, 10(1), pp.1-22.

[38] Ritter, A., Wright, E., Casey, W. and Mitchell, T., 2015, May. Weakly supervised extraction of computer security events from twitter. In Proceedings of the 24th international conference on world wide web (pp. 896-905).

[39] Le Sceller, Q., Karbab, E.B., Debbabi, M. and Iqbal, F., 2017, August. Sonar: Automatic detection of cyber security events over the twitter stream. In Proceedings of the 12th International Conference on Availability, Reliability and Security (pp. 1-11).

[40] Chambers, N., Fry, B. and McMasters, J., 2018, June. Detecting denial-of-service attacks from social media text: Applying nlp to computer security. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers) (pp. 1626-1635).

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

[42] Dionísio, N., Alves, F., Ferreira, P.M. and Bessani, A., 2020, July. Towards end-to-end cyberthreat detection from Twitter using multi-

task learning. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

[43] Shin, H.S., Kwon, H.Y. and Ryu, S.J., 2020. A new text classification model based on contrastive word embedding for detecting cybersecurity intelligence in twitter. Electronics, 9(9), p.1527.

[44] Zigomitros, A., Papageorgiou, A. and Patsakis, C., 2012, June. Social network content management through watermarking. In 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (pp. 1381-1386). IEEE.

[45] Li, F., Wu, K., Lei, J., Wen, M., Bi, Z. and Gu, C., 2015. Steganalysis over large-scale social networks with high-order joint features and clustering ensembles. IEEE Transactions on Information Forensics and Security, 11(2), pp.344-357.

[46] Stokes, K. and Carlsson, N., 2013, July. A peer-to-peer agent community for digital oblivion in online social networks. In 2013 Eleventh Annual Conference on Privacy, Security and Trust (pp. 103-110). IEEE.

[47] Miller, Z., Dickinson, B., Deitrick, W., Hu, W. and Wang, A.H., 2014. Twitter spammer detection using data stream clustering. Information Sciences, 260, pp.64-73.

[48] Lee, S. and Kim, J., 2013. Warningbird: A near real-time detection system for suspicious urls in twitter stream. IEEE transactions on dependable and secure computing, 10(3), pp.183-195.

[49] Joe, M.M. and Ramakrishnan, B., 2017. Novel authentication procedures for preventing unauthorized access in social networks. Peer-to-Peer Networking and Applications, 10(4), pp.833-843.

[50] Ghazinour, K., Matwin, S. and Sokolova, M., 2016. YOURPRIVACYPROTECTOR, A recommender system for privacy settings in social networks. arXiv preprint arXiv:1602.01937.

[51] De Souza, G.A. and Da Costa-Abreu, M., 2020, July. Automatic offensive language detection from twitter data using machine learning and feature selection of metadata. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-6). IEEE.

[52] Sapienza, A., Bessi, A., Damodaran, S., Shakarian, P., Lerman, K. and Ferrara, E., 2017, November. Early warnings of cyber threats in online discussions. In 2017 IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 667-674). IEEE.

[53] Bargshady, G., Soar, J., Zhou, X., Deo, R. C., Whittaker, F. and Wang, H., 2019. A joint deep neural network model for pain recognition from face. In: 4th IEEE International Conference on Computer and Communication Systems (ICCCS 2019), 23-25 Feb 2019, Singapore.

[54] Kok, C., Jahmunah, V., Oh, S. L., Zhou, X., Gururajan, R., Tao, X., Cheong, K. H., Gururajan, R., Molinari, F. and Acharya, U. R., 2020. Automated prediction of sepsis using temporal convolutional network. Computers in Biology and Medicine, 127:103957. pp. 1-10. ISSN 0010-4825

[55] Bargshady, G., Soar, J., Zhou, X., Deo, R. C., Whittaker, F. and Wang, H., 2020. Ensemble neural network approach detecting pain intensity from facial expressions. Artificial Intelligence in Medicine 109, 101954