Chen GAO, Xuan ZHANG, Mengting HAN, Hui LIU, 2021. A review on cyber security named entity recognition. *Frontiers of Information Technology & Electronic Engineering,* 22(9):1153-1168. https://doi.org/10.1631/FITEE.2000286

A review on cyber security named entity recognition

Key words: Named entity recognition (NER); Information extraction; Cyber security; Machine learning; Deep learning

Corresponding author: Xuan ZHANG
E-mail: zhxuan@ynu.edu.cn
ORCID: Xuan ZHANG, https://orcid.org/0000-0003-2929-2126

NER approaches

- Different approaches are used to identify named entities from unstructured cyber security data sources. These methods are rule-based NER, dictionary-based NER, and machine learning based NER. Machine learning methods can be divided into statistical machine learning and deep learning methods.

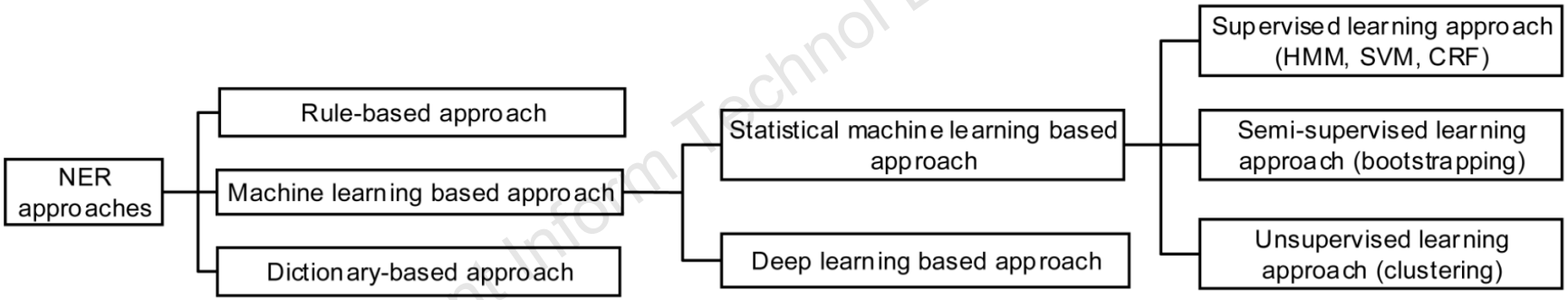

Fig. 1 Classification of named entity recognition (NER) approaches
HMM: hidden Markov model; SVM: support vector machine; CRF: conditional random fields

Problems with cyber security NER

- First, the cyber security domain has many technical terms and complex naming conventions. Some examples are as follows:
  - Conjunction and disjunction
  - Non-standardized naming convention
  - Abbreviation
  - Massive nesting
- Second, in the cyber security domain, there are few available datasets. There is also a lack of unified and standardized classification standards.
- Finally, the cyber security entity categories are unevenly distributed.

Cyber security NER systems

- Cyber security NER systems use three different learning approaches: supervised learning, semi-supervised learning, and unsupervised learning. Table 1 lists the works related to cyber security NER in the past 10 years. As can be seen from Table 1, most of the literature was published in the last three years.

Cyber security NER systems

Table 1 Literature on cyber security NER systems*

No.	Reference	Method description	Country	Number of citations
1	Joshi et al. (2013)	Extracting cyber security related linked data from text	USA	70
2	Mulwad et al. (2011)	Extracting information about security vulnerabilities from web text	USA	63
3	Bridges et al. (2013)	Automatic labeling for entity extraction in cyber security	USA	25
4	McNeil et al. (2013)	PACE: pattern accurate computationally efficient bootstrapping for timely discovery of cyber security concepts	USA	22
5	Lal (2013)	Information extraction of security related entities and concepts from unstructured text	USA	15
6	Weerawardhana et al. (2014)	Automated extraction of vulnerability information for home computer security	USA	9
7	Dionísio et al. (2019)	Cyberthreat detection from Twitter using deep neural networks	Portugal	6
8	Gasmi et al. (2018)	LSTM recurrent neural networks for cyber security NER	USA	4
9	Zhou et al. (2018)	Automatic identification of indicators of compromise using neural network based sequence labeling	China	3
10	Xiao (2018)	Towards a two-phase unsupervised system for cyber security concept extraction	USA	1
11	Shang et al. (2017)	A framework to construct a knowledge base for cyber security	China	1
12	Mazharov and Dobrov (2018)	NER for an information security domain	Russia	1
13	Tikhomirov et al. (2020)	Using BERT and augmentation in named-entity recognition for a cyber security domain	Russia	1

To be continued

Cyber security NER systems

Table 1

14	Qin et al. (2019)	A network security entity recognition method based on a feature template and CNN-BiLSTM-CRF	China	0
15	Ma et al. (2021)	Cyber security NER using bidirectional long short-term memory with conditional random fields	China	0
16	Zhang et al. (2019)	Multifeature NER in information security based on adversarial learning	China	0
17	Long et al. (2019)	Collecting indicators of compromise from unstructured text of cyber security articles using neural network based sequence labeling	China	0
18	Gu et al. (2020)	Tweet malware name recognition based on an enhanced BiLSTM-CRF model	China	0
19	Georgescu et al. (2019)	Named-entity-recognition-based automated system for diagnosing cyber security situations in IoT networks	Romania	0
20	Li et al. (2019)	A self-attention-based approach for NER in cyber security	China	0
21	Liu (2020)	Network security entity recognition methods based on the deep neural network	China	0
22	Wu et al. (2020)	An effective approach of NER for cyber threat intelligence	China	0
23	Simran et al. (2020)	A deep learning approach for intelligent NER of cyber security	USA	0
24	Kim et al. (2020)	Automatic extraction of named entities of cyber threats using a deep Bi-LSTM-CRF network	Korea	0
25	Wang et al. (2020)	NER in a threat intelligence domain with triple loss function based on metric learning (TSFL)	China	0

* Ordered by the number of citations in Google Scholar (till Nov. 20, 2020)

Cyber security NER systems

- Fig. 3 shows how the F1 score of the cyber security methods changes over time. As can be seen from the figure, the research methods gradually change from the statistical machine learning method to the deep learning method, and the F1 score continuously increases.

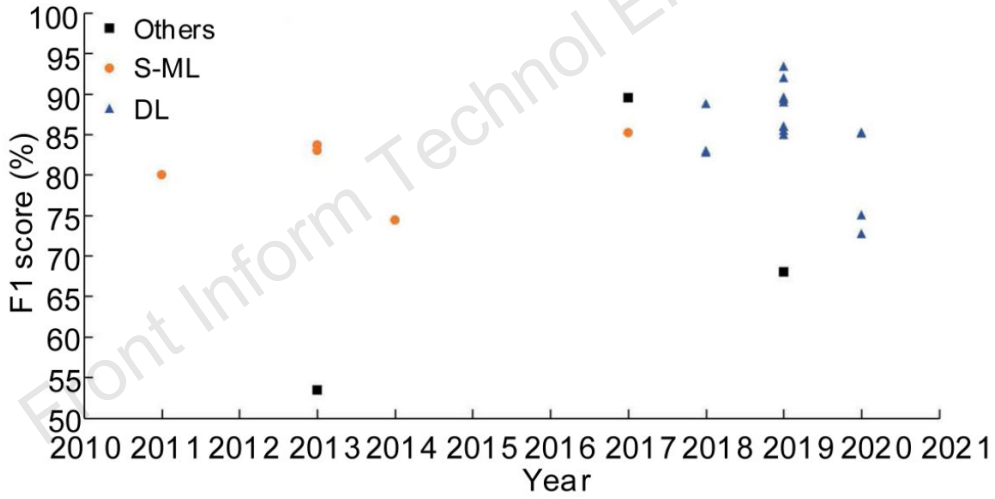

Fig. 3 Distribution of the cyber security NER F1 score over the years

DL: deep learning; S-ML: statistical machine learning; Others: semi-supervised and unsupervised learning

Cyber security NER systems

- The supervised learning method is the main learning method in machine learning systems. Early supervised learning models based on statistical machine learning include mainly HMM, SVM, and CRF. Table 2 lists the works related to machine learning network security NER.

Table 2 Summary of related work on statistical machine learning cyber security NER

Reference	Method	Dataset	F1 score	Pros	Cons
Mulwad et al. (2011)	SVM, Wikitology	Internet cyber-security text	80.0%	Effective way to link entities to a knowledge base	Fully dependent on Wikitology, feature engineering
Weerawardhana et al. (2014)	CRF, rule template	Bridges et al. (2013)'s	74.4%, 67.8%	Comparison of different methods	Feature engineering, poor mobility
Joshi et al. (2013)	CRF, DBpedia	Internet cyber-security text	83.0%	Effective way to link entities to a knowledge base	Fully dependent on DBpedia, difficulty in reproducing the method
Lal (2013)	CRF	Internet cyber-security text	83.7%	Comparison of different datasets	Small feature engineering size of the dataset
Shang et al. (2017)	CRF, rule template	Lal (2013)'s	85.2%	Combining the rule and machine learning	Relying on experts to define rules for feature engineering

Cyber security NER systems

- Table 3 summarizes the cyber security NER research based on deep learning.

Table 3 Summary of deep learning cyber security NER research

Reference	Method	Dataset	F1 score	Pros	Cons
Gasmi et al. (2018)	LSTM, CRF	Bridges et al. (2013)'s	82.80%	No feature engineering	Difficulty in identifying complex entities
Mazharov and Dobrov (2018)	LSTM, CRF	Internet cyber security text	83.00%	Improved performance	Small examples providing small improvement
Wu et al. (2020)	LSTM, CRF, dictionary	Internet cyber security text	85.27%	Improved performance	Building a dictionary manually
Ma et al. (2021)	LSTM, CRF	Lal (2013)'s	89.38%	Improved network structure	More features
Simran et al. (2020)	GRU, RNN, CRF	Bridges et al. (2013)'s	93.40%	Improved performance and efficiency	A lot of label data
Qin et al. (2019)	CNN, LSTM, CRF, feature template	Internet cyber security text	86.00%	Identification of mixed security entities in Chinese and English	Manually defining feature templates
Liu (2020)	CNN, LSTM, CRF, unified language model pre-training	Internet cyber security text	86.00%	Improved NER effect on small datasets	Manually defining feature templates
Zhou et al. (2018)	LSTM, CRF, attention, token spelling features	Internet cyber security text	88.80%	Improved low-frequency entity recognition	Difficulty in distinguishing tokens similar to IOCs but not malicious
Long et al. (2019)	LSTM, CRF, multi-attention, token spelling features	Internet cyber security text	89.60%	Better contextual expression learning	No pre-embedded language model

To be continued

Cyber security NER systems

- Table 3

Li et al. (2019)	LSTM, CRF, multi-attention	Internet cyber security text	84.98%	Improved performance	More features required
Dionísio et al. (2019)	CNN, LSTM, CRF	Internet cyber security text	92.00%	End-to-end processing of short text data	Fewer entity categories
Gu et al. (2020)	BERT, LSTM, CRF, multi-attention	Internet cyber security text	85.50%	Improved recognition effect of a single complex entity category	Low efficiency
Wang et al. (2020)	BERT, Word2Vec, new loss function, attention	Internet cyber security text	85.16%	Solving the problem of unbalanced data label distribution	High time complexity
Tikhomirov et al. (2020)	BERT, dataset augmentation	Internet cyber security text	72.74%	Proposing a new method of dataset augmentation for NER tasks	Poor recognition caused by unbalanced data categories
Zhang et al. (2019)	GAN, LSTM, CRF	Internet cyber security text	89.00%	Improved quality of crowdsourced annotations in information security	High time complexity
Kim et al. (2020)	LSTM, CRF, bag-of-character (BOC)	Internet cyber security text	75.05%	More effective BOC character embedding than RNN and CNN	Low performance

Cyber security NER systems

- Table 4 summarizes cyber security NER research based on semi-supervised learning. Table 5 summarizes the NER cyber security research based on unsupervised learning.

Table 4 Summary of related work on semi-supervised learning cyber security NER

Reference	Method	Dataset	F1 score	Pros	Cons
McNeil et al. (2013)	Bootstrapping pattern	Internet cyber security text	53.4%	Selection of informative samples from the unlabeled data	Small dataset and very low performance
Georgescu et al. (2019)	Ontology, Watson Knowledge Studio tool	Internet cyber security text	68.0%	Combining manually annotated data with context-based knowledge extraction	Low performance

Table 5 Summary of related work on unsupervised learning cyber security NER

Reference	Method	Dataset	F1 score	Pros	Cons
Xiao (2018)	Word2Vec ontology	Internet cyber security text	89.5%	Exploiting a large amount of unlabeled data	Entity ambiguity and difficulty in identifying new entities

Cyber security NER systems

- Fig. 4 shows how methods have changed over the years. As can be seen, before 2017, a feature template based approach was adopted for cyber security NER. Starting in 2017, word embedding paradigm methods took the lead. The pre-training paradigm method has been gradually applied to cyber security NER since 2019.

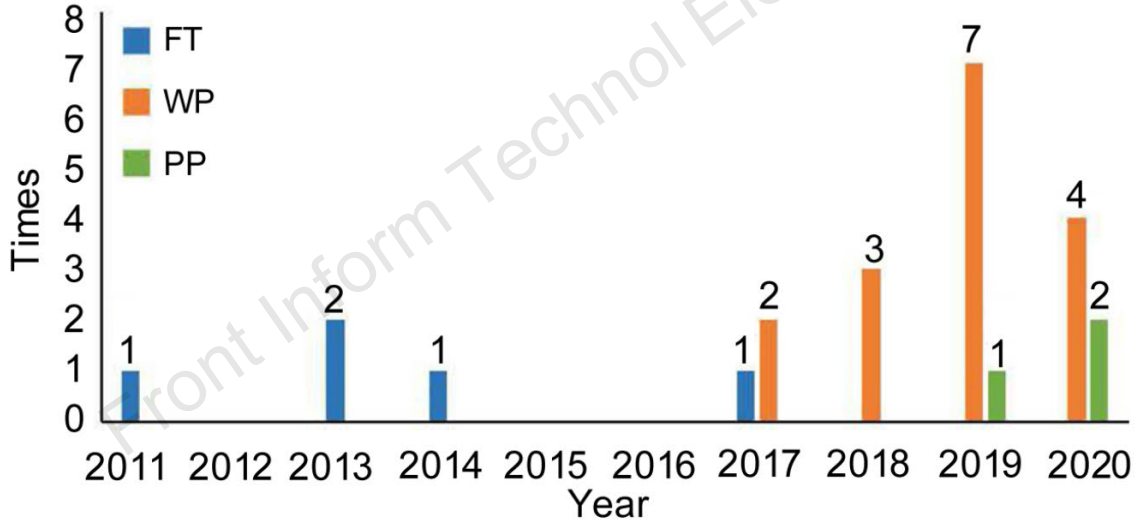

Fig. 4 Yearly occurrence of methods

FT: feature template; WP: word embedding paradigm; PP: pre-training paradigm. References to color refer to the online version of this figure

Resources available for cyber security NER

- Corpus

Table 6 Corpus for cyber security

Reference	Main entity	Type size	Language
Lal (2013)	Software, modifier, operating system, consequences, attack, means, file name, network, hardware	More than 45 000 tokens, 5000 tagged entities	English
Bridges et al. (2013)	Vendor, application, version, edition, OS, hardware file	853 560 tokens, 73 964 tags	English
Mazharov and Dobrov (2018)	Hacker, hacker_group, virus, device.tech, program	377 364 tokens, 13 076 tags	Russian
Kim et al. (2020)	Hash, malware, IP, URL	498 000 tokens, 15 720 tags	English

- Cyber security ontology
  - Cyber security ontology refers to the fusion and reasoning of a cyber security knowledge base using the ontology method to obtain a clear and standardized formal explanation.
- Evaluation criteria
  - The performance of an NER system is evaluated by comparing the output with manual annotations. Three common metrics are Precision, Recall, and F1-score.

Conclusions and future trends

- Exploring the application of unsupervised or semi-supervised technology. Such techniques leverage context patterns and have been quite successful in open named-entity extraction tasks.
- Developing a more comprehensive cyber security ontology. At present, there is no unified standard for entity classification in the field of cyber security. For supervised machine learning methods, switching to a different ontology means changing the entity categories, which is not feasible because it requires re-annotating the corpus.
- Developing a more comprehensive deep learning model. More deep learning technologies will be used to improve the effect of the model and its recognition ability, for example, multi-task learning (Caruana, 1997), active learning (Shen et al., 2017), transfer learning (Lee et al., 2018), and reinforcement learning (Kaelbling et al., 1996).

References

[1] Bridges RA, Jones CL, Iannacone MD, et al., 2013. Automatic labeling for entity extraction in cyber security.

[2] Caruana R, 1997. Multitask learning. Mach Learn, 28(1):41- 75.

[3] Devlin J, Chang MW, Lee K, 2018. BERT: pre-training of deep bidirectional transformers for language understanding.

[4] Dionísio N, Alves F, Ferreira PM, et al., 2019. Cyberthreat detection from Twitter using deep neural networks. Int Joint Conf on Neural Networks, p.1-8.

[5] Eddy SR, 1996. Hidden Markov models. Curr Opin Struct Biol, 6(3):361-365.

[6] Li W, et al., 2019. The way to apply machine learning to IoT driven wireless network from channel perspective. China Commun, 16(1):148-164.

[7] Gasmi H, Bouras A, Laval J, 2018. LSTM recurrent neural networks for cyber security named entity recognition. Proc 13th Int Conf on Software Engineering Advances, p.12-17.

[8] Georgescu TM, Iancu B, Zurini M, 2019. Named-entity- recognition-based automated system for diagnosing cybersecurity situations in IoT networks. Sensors, 19(15): 3380.

[9] Gu XM, Liu JY, Cheng PS, et al., 2020. Malware name recognition in tweets based on enhanced BiLSTM-CRF model. Comput Sci, 47(2):245-250 (in Chinese).

[10] Hearst MA, Dumais ST, Osuna E, et al., 1998. Support vector machines. IEEE Intell Syst Their Appl, 13(4):18-28.

References

[11] Joshi A, Lal R, Finin T, et al., 2013. Extracting cybersecurity related linked data from text. Proc 7th Int Conf on Semantic Computing, p.252-259.

[12] Kaelbling LP, Littman ML, Moore AW, 1996. Reinforcement learning: a survey.

[13] Kim G, Lee C, Jo J, et al., 2020. Automatic extraction of named entities of cyber threats using a deep Bi-LSTM-CRF network. Int J Mach Learn Cyber, 11(10):2341-2355.

[14] Lafferty JD, McCallum A, Pereira FCN, 2001. Conditional random fields: probabilistic models for segmenting and labeling sequence data. Proc 18th Int Conf on Machine Learning, p.282-289.

[15] Lal R, 2013. Information Extraction of Security Related Entities and Concepts from Unstructured Text. MS Thesis, University of Maryland, Baltimore County, Baltimore, USA.

[16] Lample G, Ballesteros M, Subramanian S, et al., 2016. Neural architectures for named entity recognition.

[17] LeCun Y, Bengio Y, Hinton G, 2015. Deep learning. Nature, 521(7553):436-444.

[18] Lee JY, Dernoncourt F, Szolovits P, 2018. Transfer learning for named-entity recognition with neural networks. Proc 11th Int Conf on Language Resources and Evaluation, p.4471- 4473.

[19] Li T, Guo YB, Ju AK, 2019. A self-attention-based approach for named entity recognition in cybersecurity. Proc 15th Int Conf on Computational Intelligence and Security, p.147-150.

[20] Liu WG, 2020. Network security entity recognition methods based on the deep neural network. In: Huang CC, Chan YW, Yen N (Eds.), Data Processing Techniques and Applications for Cyber-Physical Systems. Springer, Singapore, p.1687-1692.

References

[21] Long Z, Tan LZ, Zhou SP, et al., 2019. Collecting indicators of compromise from unstructured text of cybersecurity articles using neural-based sequence labelling. Int Joint Conf on Neural Networks, p.1-8.

[22] Lowd D, Meek C, 2005. Adversarial learning. Proc 11th ACM SIGKDD Int Conf on Knowledge Discovery in Data Mining, p.641-647.

[23] Ma PC, Jiang B, Lu ZG, et al., 2021. Cybersecurity named entity recognition using bidirectional long short-term memory with conditional random fields. Tsinghua Sci Technol, 26(3):259-265.

[24] Marrero M, Urbano J, Sánchez-Cuadrado S, et al., 2013. Named entity recognition: fallacies, challenges and opportunities. Comput Stand Interf, 35(5):482-489.

[25] Mazharov I, Dobrov BV, 2018. Named entity recognition for information security domain. Proc 20th Int Conf on Data Analytics and Management in Data Intensive Domains, p.200-207.

[26] McNeil N, Bridges RA, Iannacone MD, et al., 2013. PACE: pattern accurate computationally efficient bootstrapping for timely discovery of cyber-security concepts. Proc 12th Int Conf on Machine Learning and Applications, p.60-65.

[27] Mendes PN, Jakob M, García-Silva A, et al., 2011. DBpedia spotlight: shedding light on the web of documents. Proc 7th Int Conf on Semantic Systems, p.1-8.

[28] Mulwad V, Li WJ, Joshi A, et al., 2011. Extracting information about security vulnerabilities from web text. IEEE/WIC/ ACM Int Conf on Web Intelligence and Intelligent Agent Technology, p.257-260.

[29] Nadeau D, Sekine S, 2007. A survey of named entity recognition and classification. Lingv Investig, 30(1):3-26.

[30] Peters ME, Ammar W, Bhagavatula C, et al., 2017. Semi-supervised sequence tagging with bidirectional language models.

References

[31] Qin Y, Shen GW, Zhao WB, et al., 2019. A network security entity recognition method based on feature template and CNN-BiLSTM-CRF. Front Inform Technol Electron Eng, 20(6):872-884.

[32] Riloff E, 1993. Automatically constructing a dictionary for information extraction tasks. Proc 11th National Conf on Artificial Intelligence, p.811-816.

[33]Roy A, Park Y, Pan SH, 2017. Learning domain-specific word embeddings from sparse cybersecurity texts.

[34] Devlin J, Chang MW, Lee K, 2018. BERT: pre-training of deep bidirectional transformers for language understanding.

[35] Ruder S, 2016. An overview of gradient descent optimization algorithms.

[36] Shang HJ, Jiang R, Li AP, et al., 2017. A framework to construct knowledge base for cyber security. Proc IEEE 2nd Int Conf on Data Science in Cyberspace, p.242-248.

[37] Shen YY, Yun H, Lipton ZC, et al., 2017. Deep active learning for named entity recognition. Proc 2nd Workshop on Representation Learning for NLP, p.252-256.

[38] Simran K, Sriram S, Vinayakumar R, et al., 2020. Deep learning approach for intelligent named entity recognition of cyber security.

[39] Syed Z, 2010. Wikitology: a Novel Hybrid Knowledge Base Derived from Wikipedia. PhD Thesis, University of Maryland, Baltimore County, Baltimore, USA.

[40] Gu XM, Liu JY, Cheng PS, et al., 2020. Malware name recognition in tweets based on enhanced BiLSTM-CRF model. Comput Sci, 47(2):245-250 (in Chinese).

[41] Syed Z, Padia A, Mathews ML, et al., 2016. UCO: a unified cybersecurity ontology. AAAI Workshop on Artificial Intelligence for Cyber Security, p.14-21.

References

[42] Tikhomirov M, Loukachevitch N, Sirotina A, et al., 2020. Using BERT and augmentation in named entity recognition for cybersecurity domain. Proc 25th Int Conf on Applications of Natural Language to Information Systems, p.16-24.

[43] Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

[44] Wang XR, Xiong ZH, Du XY, et al., 2020. NER in threat intelligence domain with TSFL. Proc 9th Int Conf on Natural Language Processing and Chinese Computing, p.157-169.

[45] Weerawardhana S, Mukherjee S, Ray I, et al., 2014. Automated extraction of vulnerability information for home computer security. Proc 7th Int Symp on Foundations and Practice of Security, p.356-366.

[46] Wu H, Li XY, Gao YL, 2020. An effective approach of named entity recognition for cyber threat intelligence. Proc IEEE 4th Information Technology, Networking, Electronic and Automation Control Conf, p.1370-1374.

[47] Xiao ZF, 2018. Towards a two-phase unsupervised system for cybersecurity concepts extraction. Proc 13th Int Conf on Natural Computation, Fuzzy Systems and Knowledge Discovery, p.2161-2168.

[48] Zhang H, Guo YB, Li T, 2019. Multifeature named entity recognition in information security based on adversarial learning. Secur Commun Netw, 2019:6417407.

[49] Zhou SP, Long Z, Tan LZ, et al., 2018. Automatic identification of indicators of compromise using neural-based sequence labelling.