A Knowledge Graph Construction Approach for Legal Domain

Biao DONG*, Haoze YU, Haisheng LI

Abstract: Considering that the existing domain knowledge graphs have difficulty in updating data in a timely manner and cannot make use of knowledge sufficiently in the construction process, this paper proposes a legal domain knowledge graph construction approach based on 'China Judgments Online' in order to manage the cases' knowledge contained in it. The construction process is divided into two steps. First, we extract the classification relationships of the cases from structured data. Then, we obtain attribute knowledge of cases from semi-structured data and unstructured data through a relationship extraction model based on an improved cross-entropy loss function. The triples describing knowledge of cases are stored through Neo4j. The accuracy of the proposed approach is verified through experiments and we construct a legal domain knowledge graph which contains more than 4K classification relationships and 12K attribute knowledge to prove its validity.

Keyword: China Judgments Online; Domain Knowledge Graph; Neo4j; Relationship Extraction

1 INTRODUCTION

As a platform for publishing effective judgment documents of the people's courts at all levels, the quantity of documents in 'China Judgments Online' has exceeded 100 million. At the same time, as the scale of the documents continues to grow, the content of the cases contained in them also explodes. Therefore, how to extract and manage the knowledge contained in it has become a key issue to be solved.

Knowledge graphs can manage a large scale of knowledge by describing relationships between concepts and instances as a kind of semantic knowledge base [1]. Therefore, it can be used as an effective tool to manage the knowledge in the judgment document. From the aspect of application area, knowledge graphs can be divided into general knowledge graphs and domain knowledge graphs. Many general knowledge graphs appeared after Google proposed the knowledge graph project and built an intelligent search engine based on it in 2012. Microsoft Research Asia [2] proposed a knowledgebase called 'Probase', which contains more than 1.6 billion web pages and 27 million concepts. The University of Leipzig and Mannheim [3] jointly constructed a multilingual knowledgebase called 'DBpedia'. The researchers of the Marcus Institute in German [4] constructed a comprehensive knowledgebase, 'YAGO'. Although general knowledge graphs can cover a large amount of knowledge and are widely used in academia and industry, it is difficult to supply relevant services for specific areas because of the lack of specific domain knowledge. Therefore, a large quantity of domain knowledge graphs has emerged. Wolfram Research [5] proposed a knowledge graph in the field of mathematics called 'WolframAlpha'. Amazon [6] proposed an online database for movies, TV shows and film productions called 'IMDB'. A bilingual knowledge graph called 'XLore' containing 663K concepts and 56K attributes was proposed by Wang et al. [7]. Yu et al. [8] constructed a food domain knowledge graph that contains 8,373 pairs of upper and lower relationships and 74,622 pairs of non-superordinate relationships. A knowledge graph construction approach for college internal policy control was proposed by Wang et al. [9]. These knowledge graphs contain a large amount of data in various domains, but in view that the knowledge in the construction processes is mainly extracted from existing databases or structured data, thus a great deal of knowledge is ignored. At the same time, although the general knowledge graphs contain a certain quantity of basic legal knowledge, there is still a lack of effective domain knowledge management methods for legal cases. Therefore, these existing works can be great references for the construction of legal domain knowledge graphs, but considering that legal domain data have the characteristic of real-time updating and exist in different structures, how to timely extract relationships from structured, semi-structured and unstructured data plays a significant role in the construction of legal knowledge graphs and will directly determine their quality and scale.

Research on the integration of computer science or information technology and law has become one of the focuses of the science of law. The prospect, specific practical requirements, and approaches etc. of applying artificial intelligence technology to adjudicate have become a hot topic in the field of law. Nevertheless, in terms of approach and technical means, the research of applying artificial intelligence in the judicial judgment field is just at the preliminary stage, most of which remains in conception, phenomenon explanation and imagination.

In this paper, the process of legal domain knowledge graph construction is divided into two parts: classification and attribute relationship extraction. The classification and attribute knowledge of cases in 'Chinese Judgments Online' is extracted through corresponding extraction approaches in view of their structures and to ultimately be stored by Neo4j. The contribution of this paper can be summarized as follows:
1) The classification relationship extraction algorithm is proposed to extract classification knowledge contained in structured data of web pages;
2) The Bi-GRU model based on the improved cross-entropy loss function is proposed to extract the attribute relationships from unstructured texts in web pages;
3) To prove the effectiveness of the proposed approach, a legal domain knowledge graph is constructed based on the 'China Judgments Online' and it can be optimized and expanded as the update of platform.

2 RELATED WORK

The existing relationship extraction approaches can be mainly divided into three types: unsupervised relationship extraction, weak-supervised relationship extraction and supervised relationship extraction.

Unsupervised relationship extraction approaches can automatically extract relationships without excessive manual intervention. Hasegawa et al. [10] proposed a fully connected clustering approach to cluster the contexts and then calculated the high-frequency words in each cluster as the marker of the relationship type based on the clustering results. Zhang et al. [11] regarded relationship extraction as a clustering problem on shallow parse trees and proposed a tree-similarity-based approach to extract relationships among name instances from a large raw corpus. Zhang et al. [12] present a fuzzy Bayesian network model to analyze the relationships between unloading level, warehouse inspection, warehouse monitoring and the quality in O2O ecommerce. Wang et al. [13] introduced the coclustering theory on the basis of k-means, not only clustering entity pairs but also relationship characteristics to make good use of the duality of datasets. A Bayesian network is used by Zhang et al. [14] to extract the relationships among factors such as personnel operation, equipment, information technology etc. and system reliability, and calculated the importance degree of each factor. However, because the relationship extraction results obtained through unsupervised approaches are difficult to be regularized, a large number of manual operations are still needed in the process of knowledge graph construction.

The weakly supervised relationship extraction approaches take a small number of instances and relationships as initial seeds and classify leftover instances by training classifiers. Zhu et al. [15] proposed a statistical extraction framework called 'StatSnowball'. Court et al. [16] automatically generated a database containing compounds and their Nell magnetic phase transition temperatures based on the weakly supervised relationship extraction approach. Although weakly supervised relationship extraction approaches can solve the problems caused by the wide relationship types, the number of seeds and how to choose them are still difficult problems to solve.

In recent years, with the development of deep learning, neural networks have been increasingly used in video processing [17], optimization of process parameters [18], relationship extraction tasks etc. Liu et al. [19] proposed a noise reduction approach that exploited semantic information from correctly labeled entity pairs to revise incorrect labels dynamically during training. Alzaidy et al. [20] addressed the relationship extraction task as sequence labeling and proposed a BiLSTM-CRF model. Zhang et al. [21] proposed a novel neural approach based on capsule networks with attention mechanisms to extract relationships in multi-instance and multilabel learning frameworks. Luo et al. [22] used dynamic transition matrices to characterize the noise of the training data in the process of relationship extraction. Yu et al. [23] proposed a remote supervised Bi-GRU model with multiple attention mechanisms in order to obtain non-classification relationships among concepts. Liu et al. [24] proposed a

relationship extraction network based on inner sentence noise reduction and transfer learning. Most of the existing neural networks ignore the temporal features among contexts in the learning process, and a large number of feature functions need to be manually defined to accurately complete the relationship extraction task if the conditional random field (CRF) is introduced. Therefore, this paper proposes a Bi-GRU network based on an improved cross-entropy loss function, which can consider the temporal characteristics of the legal domain texts as well as optimize the network parameter updating process.

3 RESEARCH APPROACH
3.1 Relationship Extraction Approach for Structured and Semi-structured Knowledge in 'China Judgments Online'

'China Judgments Online' is the most authoritative and influential website publishing case that was tried by the courts at all levels. It supplies massive texts of cases that provide conditions for intelligent adjudication. 'China Judgment Online' has a good classification system containing classification knowledge such as 'keyword', 'cause', and 'document types' for each case as shown in Fig. 1 and has high accuracy and credibility. In this paper, a classification relationship extraction algorithm shown in Algorithm 1 is proposed to obtain structured knowledge on the web page. We start with several keyword root nodes, the classification knowledge of cases is acquired layer by layer within the Max_Depth range and ultimately is obtained a hierarchy containing classification relationships among concepts and instances. At the same time, the classification hierarchy is able to acquire more concept nodes, instance nodes and classification relationships as the quantity of keyword root nodes and MaxDepth increases.

The attribute information of the cases, such as 'Legal Representative' and 'Respondent', are usually contained and explained in the 'Party' section of case pages. Therefore, this paper sets up a web crawler to crawl the information of the 'Party' to obtain a large scale of case-related semi-structured attribute information
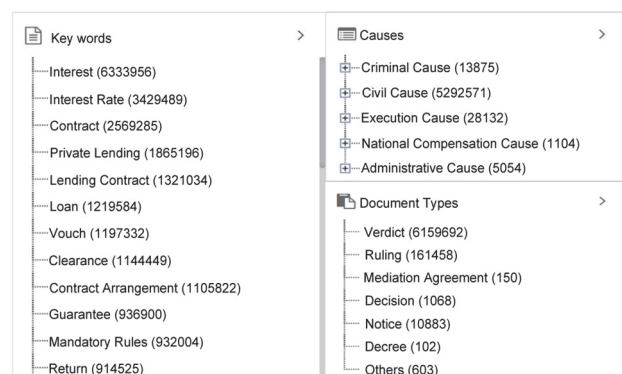

Key words	
Interest (6333956)	
Interest Rate (3429489)	
Contract (2569285)	
Private Lending (1865196)	
Lending Contract (1321034)	
Loan (1219584)	
Vouch (1197332)	
Clearance (1144449)	
Contract Arrangement (1105822)	
Guarantee (936900)	
Mandatory Rules (932004)	
Return (914525)	

Causes	
Criminal Cause (13875)	
Civil Cause (5292571)	
Execution Cause (28132)	
National Compensation Cause (1104)	
Administrative Cause (5054)	

Document Types	
Verdict (6159692)	
Ruling (161458)	
Mediation Agreement (150)	
Decision (1068)	
Notice (10883)	
Decree (102)	
Others (603)	

Figure 1 Classification knowledge of cases in 'China Judgments Online'

Algorithm 1 Classification relationship extraction algorithm

Input:	Keyword node set of web pages named *Concept*, maximum depth range of the crawler *Max_Depth*, current crawl depth *Depth*.
Output:	The hierarchy of keywords and cases in 'Chinese Judgments Online' named *Classification hierarchy* and *WPI table* which stores case web pages' information.
Procedure:	1) Initialize *Classification hierarchy*;
	2) Initialize the subconcept node set named *Temp*;
	3) *Depth* = 0;
	4) While (*Depth ≤ Max_Depth*):
	5) While '*Concept*' is not NULL:
	6) Obtain *Temp* and case node set named *Concept_Instance* contained in *Concept*;
	7) Store web pages corresponding to the cases in *Concept_Instance* into *WPI table*;
	8) *Depth* + = 1;
	9) Store *Temp* and *Concept_Instance* nto *Classification hierarchy*;
	10) Clear *Concept_Instance* contained in *Concept*;
	11) Replace *Concept* with *Temp*;
	12) Output *Classification hierarchy*, *WPI table*.

3.2 Bi-GRU Relationship Extraction Network Based on an Improved Cross-entropy Loss Function
3.2.1 Bi-GRU Network

This paper proposes a Bi-GRU network based on an improved cross-entropy loss function to extract the attribute relationships of cases from the unstructured texts in the case documents, and the network structure is shown in Fig. 2. The work of the input layer is divided into two procedures. First, we segment the document with the word segmentation tool 'Jieba'. Then, we use 'word2vec' to map the words in the sentences into word vectors.
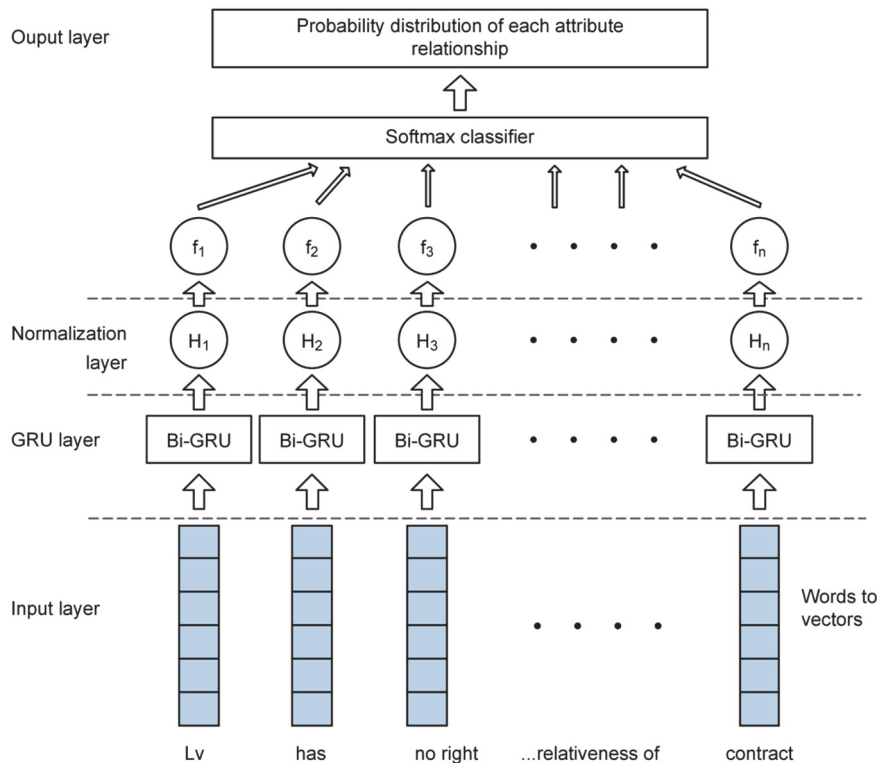

Figure 2 Attribute relationship extraction model based on an improved cross-entropy loss function

As a variant of LSTM, GRU only has an update gate and reset gate. Therefore, it greatly reduces the scale of the parameters and improves the training speed of the network, and its structure is also simpler than that of LSTM. Bi-GRU network can obtain the context information of the current word through the forward and reverse GRU to make relationship classification results more accurate. After the vector x_i obtained by mapping the *i*-th word in the

sentence is input into the Bi-GRU network, the new memory $\widetilde{h}_i$ of GRU is calculated from the past hidden state h_{i-1} and the new input x_i by Eq. (1), and the reset signal r_i is obtained through Eq. (2), which determines how much information of the hidden state of the last moment is needed to be forgotten.

$$\widetilde{h}_i = \tanh\left(W \cdot \left[r_i \cdot h_{i-1}, x_i\right]\right) \tag{1}$$

$$r_i = \sigma\left(W_r \cdot \left[h_{i-1}, x_i\right]\right) \tag{2}$$

The update gate z_i is calculated by Eq. (3), which determines how much information from the previous hidden layer state will be passed to the current hidden state h_i.

$$z_i = \sigma\left(W_z \cdot \left[h_{i-1}, x_i\right]\right) \tag{3}$$

The hidden state h_i is calculated from the past hidden state h_{i-1} and the new memory $\widetilde{h}_i$ is calculated through Eq. (4), where the parameters W, W_r, and W_z change continuously with the learning process of the GRU network. The forward feature and reverse feature of the i-th word are obtained through GRU network learning, and then the final output H_i of the i-th word in the Bi-GRU layer can be calculated through Eq. (5).

$$h_i = \left(1 - z_i\right) \cdot h_{i-1} + z_i \cdot \tilde{h}_i \tag{4}$$

$$H_i = \left[\vec{h}_i \oplus \bar{h}_i\right] \tag{5}$$

The feature vector $H\{H_1, H_2, ..., H_n\}$ corresponding to each word can be obtained from the output layer of the Bi-GRU network. After the normalization operation as shown in Eq. (6), $f\{f_1, f_2, ..., f_n\}$ are sequentially entered into the multiclassifier '*Softmax*', and the category with the highest score is used as the final attribute of the input word.

$$f = softmax(H) \tag{6}$$

3.2.2 Improved Cross-entropy Loss Function

Although each correctly classified sample has a slight effect on the update of network parameters, it still has a considerable impact on the extraction model and leads to worse convergence of the loss function due to the large number of training samples. This paper proposes an improved cross-entropy loss function to reduce the impact of correctly classified samples on parameter updating and focus on the misclassified word samples at the same time. The traditional cross-entropy loss function is shown in Eq. (7), where θ represents all parameters in the model, N represents the number of sentences, M represents the number of relationship classifications, $r_i \in \{0, 1\}$ represents the true value of the label, and $\hat{y}_i$ represents the

predictive value. Eq. (8) represents the calculated cross-entropy of a single sample in the traditional approach.

$$T_{CE}(\theta) = -\sum_{k=1}^{N}\sum_{i=1}^{M} r_{ki}\log(\hat{y}_{ki}, \theta) \tag{7}$$

$$L_{CE} = -\sum_{i=1}^{M} r_i\log(\hat{y}_i, \theta) \tag{8}$$

This paper proposes an improved cross-entropy loss function, which is shown in Eq. (9) and Eq. (10). α represents a variable focus parameter, P_s approaches 0 when the attribute relationship is correctly classified, and P_s approaches 1 when the attribute relationship is misclassified. Additionally, the value of the loss function can be adjusted by setting different focus parameters α on correctly classified and incorrectly classified samples.

$$p_s = 1 - r_i \cdot \hat{y}_i - \left(1 - r_i\right)\left(1 - \hat{y}_i\right) \tag{9}$$

$$T_{CE}(\theta) = \sum_{k=1}^{N} p_s^{\alpha} L_{CE} \tag{10}$$

4 EXPERIMENT RESULTS

This paper obtains 2,000 documents from 'China Judgment Online' and labels attribute information for more than 173k sentences. The attribute information is divided into twenty categories, including 'defendant', 'public prosecutor', 'legal provision', and 'irrelevant words', which are randomly divided into a training set and test set at a proportion of 7:3, and the parameters of the model are updated through the improved cross-entropy loss function above. We compare the proposed model with the original model and the existing approaches based on CRF and BiLSTM + CRF on accuracy rate P, recall rate R and F_1 values. The calculation of the F_1 value is shown in Eq. (11), and the result of the attribute relationship extraction in the documents is shown in Tab. 1. It can be seen that the approach proposed in our paper is superior to existing approaches in terms of the attribute relationship extraction task. Taking cases belonging to the category of 'illegal preservation compensation' as an example, more than 36,000 attribute relationships are extracted with an accuracy of 82.57% from 1,112 documents based on the proposed model.

$$F_1 = \frac{2PR}{P + R} \tag{11}$$

Table 1 Attribute relationship extraction results

Approaches	P / %	R / %	F_1
CRF	78.73	79.06	78.89
Bi-LSTM+CRF	82.38	83.49	82.93
Bi-GRU	81.70	82.41	82.05
Bi-GRU+ Our loss	82.85	83.79	83.32

Additionally, more than 4K classification relationships of the documents belonging to the category of 'illegal preservation compensation' are extracted from the

structured classification system of 'China Judgment Online' through the proposed classification relationship extraction algorithm. More than 12K attribute knowledge is extracted from the 'party' of each web page by a web crawler. We combine the classification relationships with the attribute knowledge extracted by the Bi-GRU network based on the improved cross-entropy loss function and stored them through Neo4j. Finally, the legal domain knowledge graph based on the China Judgments Online is constructed, and a fragment of it and part of the corresponding codes are shown in Fig. 3 and Fig. 4. The edges of the triples describe the classification or attribute relationships contained in cases, and the leaf nodes contain the attribute knowledge such as referee date, claimant, and principal agent etc.
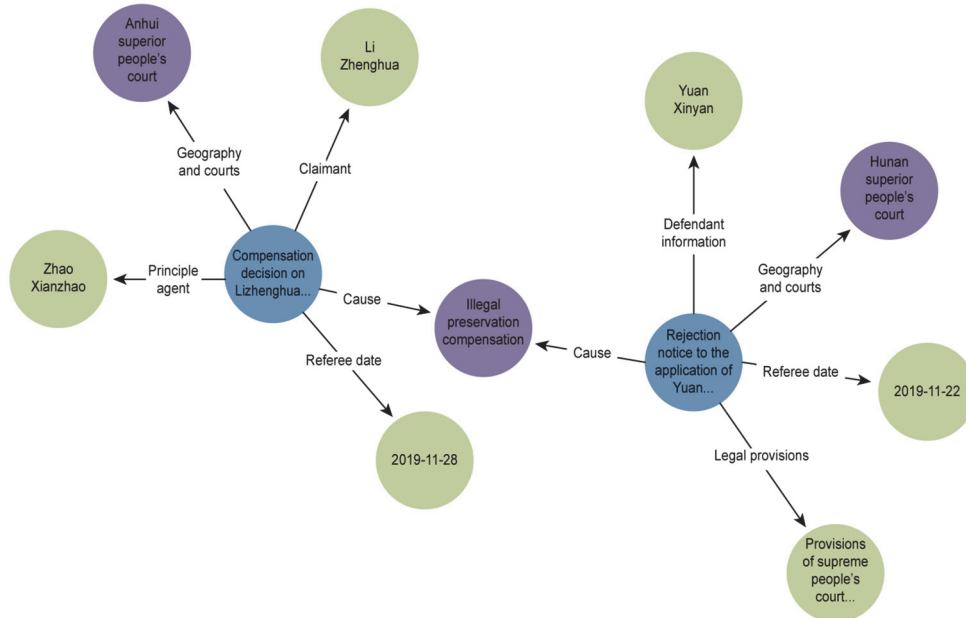

Figure 3 Part of the legal domain knowledge based on 'China Judgments Online'

```
2770  Lizhenghua = Node('Attributes',name='Lizhenghua')
2771  graph.create(Lizhenghua)
2772
2773  Anhui Superior People Court = Node('Ascription',name='Anhui Superior People Court')
2774  graph.create(Anhui Superior People Court)
2775
2776  Compensation decision on Lizhenghua illegal perservation = Node('Case',name='Compensation
      decision on Lizhenghua illegal perservation')
2777  graph.create(Compensation decision on Lizhenghua illegal perservation)
2778
2779  graph.create(Relationship(graph.data("MATCH(a:Case{name:'Compensation decision on Lizhenghua
      illegal perservation'}) return a ")[0]['a'],'Claimant',graph.data("MATCH(
      b:Attributes{name:'Lizhenghua'}) return b ")[0]['b']))
2780
2781  graph.create(Relationship(graph.data("MATCH(a:Case{name:'Compensation decision on Lizhenghua
      illegal perservation'}) return a ")[0]['a'],'Geography and courts',graph.data("MATCH(
      b:Ascription{name:'Anhui Superior People Court'}) return b ")[0]['b']))
```

Figure 4 Part of the codes corresponding to the legal domain knowledge graph

The legal domain knowledge graph is able to be further constructed and optimized with the updates of documents in the platform in order to maintain the timeliness of knowledge, and the corresponding adjustments can be made to the information such as quantity, progress and attributes of cases through revising nodes and sides.

5 CONCLUSION

This paper proposes relationship extraction approaches corresponding to structured, semi-structured and unstructured data to solve the problem of untimely updating of existing knowledge graphs and insufficient utilization of knowledge. We use 'China Judgments Online' as the data source and utilize the graph database Neo4j to store the constructed knowledge graph. In future work, we will continue to design corresponding relationship extraction models for the knowledge in different structures

of 'Chinese Judgments Online' to complement the classification and attribute knowledge in the legal domain knowledge graph. At the same time, we will also explore knowledge management methods applicable to other domains.

6 REFERENCES

[1] Wang, Z., Zhang, J. W., Feng, J. L., Chen, Z., & Aaai. (2014). Knowledge Graph Embedding by Translating on Hyperplanes. *Palo Alto: Assoc Advancement Artificial Intelligence*, *14*(2014), 1112-1119.

[2] Wang, Z., Huang, J., Li, H., Liu, B., Shao, B., Wang, H., & Zhu, K. Q. (2010). Probase: a Universal Knowledge Base for Semantic Search. *Microsoft Research Asia*.

[3] Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., & Hellmann, S. (2009). DBpedia - A crystallization point for the Web of Data. *Journal of Web Semantics*, *7*(3), 154-165.

https://doi.org/10.1016/j.websem.2009.07.002

[4] Amarilli, A., Galarraga, L., Preda, N., & Suchanek, F. M. (2014). Recent Topics of Research around the YAGO Knowledge Base. In L. Chen, Y. Jia, T. Sellis, & G. Liu (Eds.). *Web Technologies and Applications*, *8709*, 1-12. https://doi.org/10.1007/978-3-319-11116-2_1

[5] Necesal, P., Pospisil, J., & Iaeng. (2012). Experience with Teaching Mathematics for Engineers with the Aid of Wolfram Alpha. *World Congress on Engineering and Computer Science*, *I*, 271-274.

[6] IMDB official.(2016).IMDB. http://www.imdb.com. Accessed 27 Feb 2016.

[7] Wang, Z., Li, J., Wang, Z., Li, S., Li, M., Zhang, D., & Tang, J. (2013). XLore: A Large-scale English-Chinese Bilingual Knowledge Graph. *International semantic web conference (Posters & Demos)*, *1035*, 121-124.

[8] Yu, H. Z., Li, H. S., Mao, D. H., & Cai, Q. (2020). A relationship extraction method for domain knowledge graph construction. *World Wide Web-Internet and Web Information Systems*, *23*(2), 735-753. https://doi.org/10.1007/s11280-019-00765-y

[9] Wang, J. (2020). Knowledge Graph Analysis of Internal Control Field in Colleges. *Tehnicki Vjesnik-Technical Gazette*, *27*(1), 67-72. https://doi.org/10.17559/TV-20191004092659

[10] Hasegawa, T., Sekine, S., & Grishman, R. (2004). Discovering relations among named entities from large corpora. *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*, 415-422. https://doi.org/10.3115/1218955.1219008

[11] Zhang, M., Su, J., Wang, D. M., Zhou, G. D., & Tan, C. L. (2005). Discovering relations between named entities from a large raw corpus using tree similarity-based clustering. *Natural Language Processing - Ijcnlp 2005, Proceedings*, *3651*, 378-389. https://doi.org/10.1007/11562214_34

[12] Zhang, H., Feng, H., Cui, Y., & Wang, Y. (2020). A Fuzzy Bayesian Network Model for Quality Control in O2O e-Commerce. *International Journal of Computers Communications & Control*, *15*(1), 15. https://doi.org/10.15837/ijccc.2020.1.3783

[13] Wang, J., Jing, Y., Teng, Y., & Li, Q. (2012, August). A novel clustering algorithm for unsupervised relation extraction. *In Seventh International Conference on Digital Information Management (ICDIM 2012)*, 16-21. https://doi.org/10.1109/ICDIM.2012.6360156

[14] Zhang, H., Liu, Y., Zhang, Q., Cui, Y., & Xu, S. S. (2020). A Bayesian network model for the reliability control of fresh food e-commerce logistics systems. *Soft Computing*, *24*(9), 6499-6519. https://doi.org/10.1007/s00500-020-04666-5

[15] Zhu, J., Nie, Z., Liu, X., Zhang, B., & Wen, J. R. (2009). Statsnowball: a statistical approach to extracting entity relationships. *Proceedings of the 18th international conference on World wide web*, 101-110. https://doi.org/10.1145/1526709.1526724

[16] Court, C. J. & Cole, J. M. (2018). Data Descriptor: Auto-generated materials database of Curie and Neel temperatures via semi-supervised relationship extraction. *Scientific Data*, *5*, 12. https://doi.org/10.1038/sdata.2018.111

[17] Qin, L. L., Yu, N. W., & Zhao, D. H. (2018). Applying the Convolutional Neural Network Deep Learning Technology to Behavioural Recognition in Intelligent Video. *Tehnicki Vjesnik-Technical Gazette*, *25*(2), 528-535. https://doi.org/10.17559/TV-20171229024444

[18] Saric, T., Simunovic, G., Vukelic, D., Simunovic, K., & Lujic, R. (2018). Estimation of CNC Grinding Process Parameters Using Different Neural Networks. *Tehnicki Vjesnik-Technical Gazette*, *25*(6), 1770-1775. https://doi.org/10.17559/TV-20180419095119

[19] Liu, T., Wang, K., Chang, B., & Sui, Z. (2017). A soft-label method for noise-tolerant distantly supervised relation extraction. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1790-1795. https://doi.org/10.18653/v1/D17-1189

[20] Alzaidy, R., Caragea, C., & Giles, C. L. (2019). Bi-LSTM-CRF sequence labeling for keyphrase extraction from scholarly documents. *The world wide web conference*, 2551-2557. https://doi.org/10.1145/3308558.3313642

[21] Zhang, N., Deng, S., Sun, Z., Chen, X., Zhang, W., & Chen, H. (2018). Attention-based capsule networks with dynamic routing for relation extraction. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 986-992. https://doi.org/10.18653/v1/D18-1120

[22] Luo, B., Feng, Y., Wang, Z., Zhu, Z., Huang, S., Yan, R., & Zhao, D. (2017). Learning with Noise: Enhance Distantly Supervised Relation Extraction with Dynamic Transition Matrix. *Proceedings of the 55th AnnualMeeting of the Association for Computational Linguistics*, *1*, 430-439. https://doi.org/10.18653/v1/P17-1040

[23] Yu, H. Z., Li, H. S., Mao, D. H., & Cai, Q. A domain knowledge graph construction method based on Wikipedia. *Journal of Information Science*, *11*. https://doi.org/10.1177/0165551520932510

[24] Liu, T., Zhang, X., Zhou, W., & Jia, W. (2018). Neural relation extraction via inner-sentence noise reduction and transfer learning. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2195-2204. https://doi.org/10.18653/v1/D18-1243

Contact information:

Biao DONG
(Corresponding author)
Department of Law,
Beijing Technology and Business University,
No. 11 Fucheng Road, Beijing, China
E-mail: dongb@btbu.edu.cn

Haoze YU
Department of Computer,
Beijing Technology and Business University,
No. 11 Fucheng Road, Beijing, China
E-mail: 2628740743@qq.com

Haisheng LI
Department of Computer,
Beijing Technology and Business University,
No. 11 Fucheng Road, Beijing, China
E-mail: lihsh@btbu.edu.cn