Peer-Reviewed Structured Abstract

Doc2Event: Extracting Chinese Document-Level Events into Generation Models

Guangshuai Ding[1,3], Xiongtao Zhang[2,*], Bin Lin[1], Xiaomin Zhu[1], Li Ma[1], Ji Wang[2], XiaoGuang Ren[3]

[1]Strategic Assessments and Consultation, Academy of Military Sciences, Beijing, 10000, P. R. China. Email: 18774067101@163.com, bingo186@126.com, xmzhu@nudt.edu.cn, 18874857546@163.com.

[2] Laboratory for Big Data and Decision, National University of Defense Technology, Changsha, 361000, P. R. China. Email: zhangxiongtao14@nudt.edu.cn, wangji@nudt.edu.cn.

[3] Academy of Military Sciences, Beijing, 10000, P. R. China. Email: 18774067101@163.com, rxg_nudt@126.com.

*Corresponding author

Summary

Current methodologies for document-level event extraction, particularly within Chinese textual data, face significant challenges, such as the extraction of isolated events and the imprecise delineation of event interrelationships. The advent of large language models offers a promising frontier for enhancing event extraction capabilities. Considering this, this study proposes an innovative framework for the extraction of multiple events that effectively mitigates these limitations. The proposed framework employs the Entity-based Directed Acyclic Graph (EDAG) to accurately model and articulate serialization relationships among events in Chinese documents. Furthermore, it incorporates prefix prompt, thereby broadening the applicability of generative models to document-level multi-event extraction tasks. Empirical evaluations of the proposed framework were conducted using the mT5 model on challenging datasets, including the DuEE-Fin dataset from the financial domain and the FNDEE dataset from the military domain.

Background

In recent years, the alignment of pre-trained models with downstream tasks through the "pre-training + prompt learning" paradigm has proven to be a successful approach[1][2], offering promising solutions to the numerous challenges[3-7] faced in document-level event extraction. However, the direct application of this methodology to Chinese document-level event extraction remains fraught with bottlenecks, primarily due to the intrinsic complexity of the Chinese language. These challenges encompass the modeling of extracted events, the design of appropriate prompt templates, and the selection of optimal fine-tuning strategies. In this paper, we

- Propose an innovative framework that addresses the challenges of multi-event extraction by introducing generative models and prompt-based learning.

- Present a document-level event extraction method that utilizes a small number of samples by leveraging empirical knowledge from domain generalized models.
- Validate our method through widely used Chinese document-level event extraction datasets DuEE-Fin and FNDEE. The experimental results demonstrate that our method outperforms state-of-the-art (SOTA) models.

Methodology

The proposed method (presented in Fig. 1.) constructs prompt templates for document-level event extraction combined with generative macrolanguage models. It represents document-level event records using directed acyclic graphs, serializes them for representation, and ultimately formulates the document-level event extraction task as a sequence-to-sequence language generation task. To adapt the pre trained generative model to the document-level event extraction task, Lora-tuning is performed.

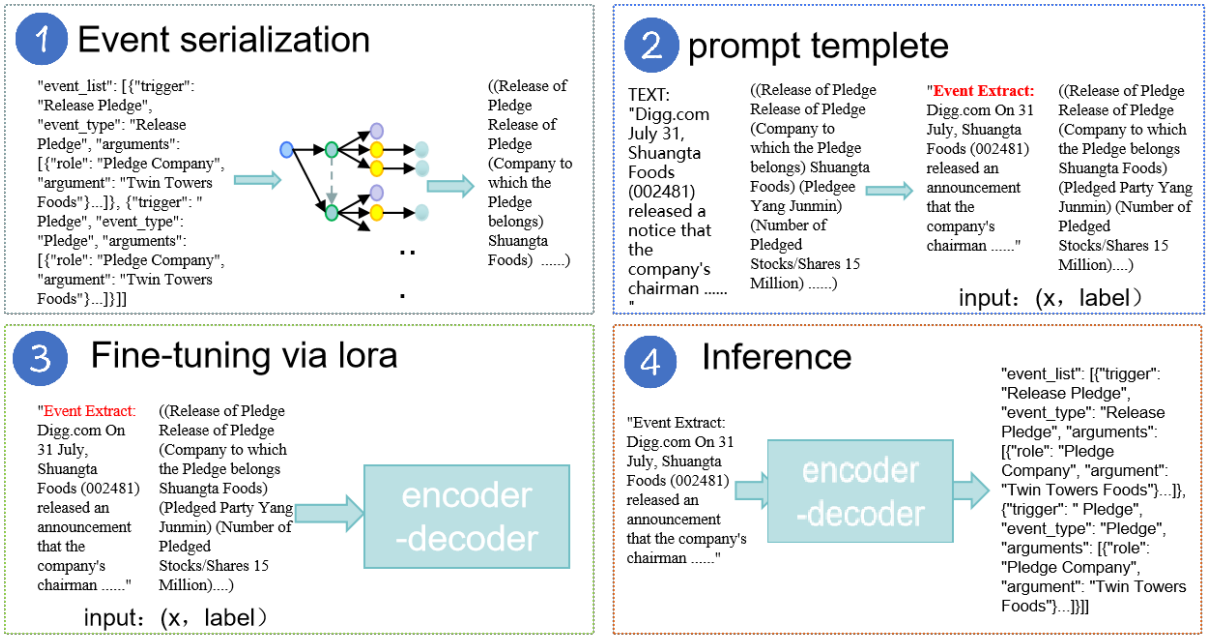

Fig. 1. Overview (Translated from Chinese)

Prompt-Learning based event extraction model:

The modeling process employs the transfer learning approach, with Prefix="Event Extraction:", x′ =['Event Extraction:′ x].The output is produced by applying the template function to the raw text input x.

A directed acyclic graph for event serialization:

The generative language model operates on sequences as input and output. Thus, the output text needs to be transformed into sequence form. We leverage directed acyclic graphs as a descriptive format for representing the event record table. As shown in Fig. 2, the graph-based representation is then serialized using the depth-first traversal method, with dotted lines denoting the event relationships.

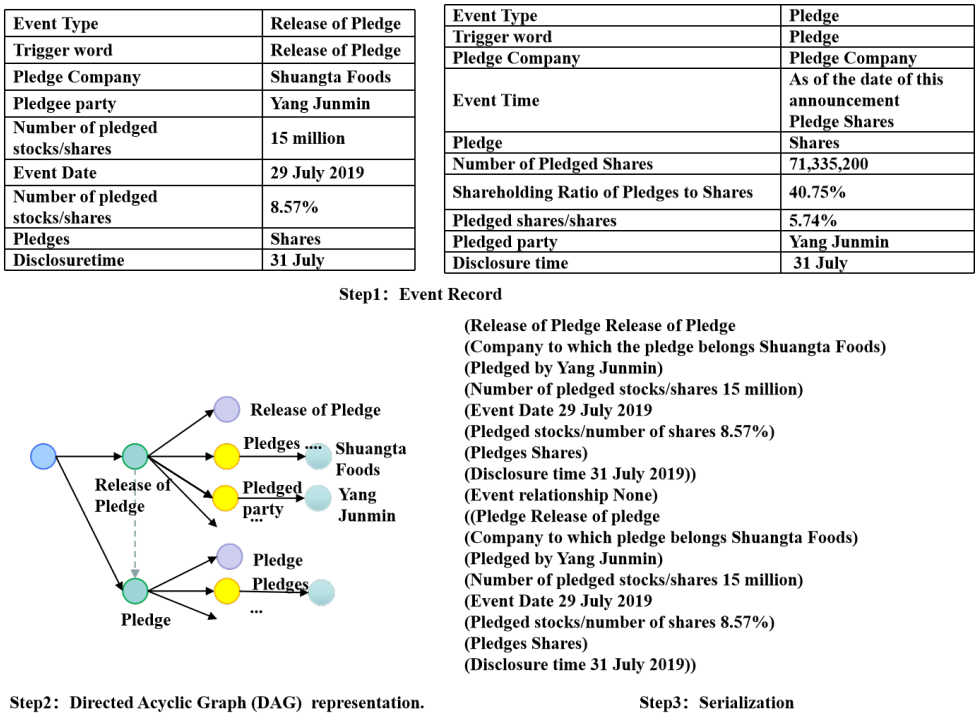

Fig. 2. Example of the event serialization method (Translated from Chinese)

Results and Observations

This paper conducts four sets of experiments as follows:

Experimental Results Basic Experiments:

We conducted basic experiments (shown in Tables 1 to 6) on the DuEE-Fin [8] and FNDEE [11] using the proposed method in this paper. A thorough comparison with the four approaches [3][9][4][10]. F1(S.) represents the F1 score for a single sentence, while F1(M.) represents the F1 score for an event with thesis elements distributed across multiple sentences.

Table 1. Experimental results on DuEE-Fin

Model	Dev				
	P	R	F1	F1(S.)	F1(M.)
Doc2EDAG	73.7	59.8	66.0	69.1	58.2
GIT	75.4	61.4	63.4	73.7	63.8
PTPCG	70.5	62.3	66.2	-	-
ReDEE	77.0	72.0	74.4	78.9	70.6
Doc2Event$_{\text{T5-base(Full-Tuning)}}$	79.7	75.9	77.8	79.5	76.3
Doc2Event$_{\text{T5-base(LoRA-tuning)}}$	80.7	75.5	80.7	81.2	77.3

Table 2. Experimental results on DuEE-Fin

Model	Online test		
	P	R	F1
Doc2EDAG	67.1	51.3	58.1
GIT	70.3	46.0	55.6
PTPCG	66.7	54.6	60.0
ReDEE	66.7	54.6	62.8
Doc2Event$_{\text{T5-base(Full-Tuning)}}$	67.9	69.4	68.7

$Doc2Event_{T5\text{-}base(LoRA\text{-}tuning)}$	71.1	69.5	70.3

Table 3. Experimental results on FNDEE

Model	**FNDEE(Dev)**		
	P	**R**	**F1**
Doc2EDAG	39.5	37.3	38.2
GIT	41.2	37.5	39.3
ReDEE	42.1	38.3	40.1
$Doc2Event_{T5\text{-}base\ (Full\text{-}Tuning)}$	51.5	58.0	54.5
$Doc2Event_{T5\text{-}base\ (LoRA\text{-}tuning)}$	54.6	59.2	56.8

Domain Migration Capability Analysis Experiments:

To evaluate the domain migration capability of our method, we conducted experiments by training the model on the DuEE-Fin and then fine-tuning it on NFDEE with varying sample sizes.

Table 4. Experimental results on DuEE-Fin ($Doc2Event_{mT5\text{-}base}$)

Sample Ratio	**$FNDEE(Dev)_{w/o}$**			**$FNDEE(Dev)_{w}$**		
	P	R	F1	P	R	F1
100%	51.5	58.0	54.5	54.3	56.2	55.2
50%	53.5	51.6	52.6	54.3	52.0	53.2
25%	46.8	52.0	49.3	50.4	52.0	51.2
10%	43.1	43.8	43.4	46.9	46.2	46.5

Fine-Tuning Methods Analysis Experiments:

We explored four fine-tuning schemes [12-14] within our proposed method. These experiments were conducted on both the DuEE-Fin and FNDEE.

Table 5. Experimental results on DuEE-Fin

Method	**Dev**			**Online test**		
	P	**R**	**F1**	**P**	**R**	**F1**
$Doc2Event_{T5\text{-}base(Full\text{-}Tuning)}$	79.7	75.9	77.8	67.9	69.4	68.7
$Doc2Event_{mT5\text{-}base(P\text{-}Tuning)}$	74.2	72.1	73.1	62.1	66.3	64.0
$Doc2Event_{mT5\text{-}base(Prefix\text{-}Tuning)}$	77.1	73.2	74.9	63.4	67.2	65.2
$Doc2Event_{mT5\text{-}base(LoRA)}$	**80.7**	**75.5**	**80.7**	**71.1**	**69.5**	**70.3**

Table 6. Experimental results on FNDEE

Method	**FNDEE(Dev)**		
	P	**R**	**F1**
$Doc2Event_{T5\text{-}base(Full\text{-}Tuning)}$	51.5	58.0	54.5
$Doc2Event_{mT5\text{-}base(P\text{-}Tuning)}$	50.1	51.5	51.0
$Doc2Event_{mT5\text{-}base(Prefix\text{-}Tuning)}$	51.1	53.2	52.1
$Doc2Event_{mT5\text{-}base(LoRA)}$	**54.6**	**59.2**	**56.8**

Analysis: The proposed method alleviates the performance degradation caused by entity overlap and argument scattering. Its effectiveness surpasses the SOTA algorithms on two domain-specific document-level event extraction datasets. It demonstrates good domain transferability, as it achieves favourable results with minimal training samples on new datasets. Moreover, after adapting to the task, retraining the model on new vertical domain datasets leads to even better performance than before. It effectively utilizes the extensive knowledge

accumulated by the generative model during the pre-training phase, even in low computational scenarios. It allows for efficient training within a shorter time frame.

Conclusion

This paper outlines a generic paradigm for accomplishing document-level event extraction tasks employing generative models and prompt learning, which can be applied to any generative model. Our experimental results demonstrate its superiority in small sample scenarios, opening new avenues for document-level event extraction.

References

1. Li, Q. et al. (2022) 'A survey on Deep Learning event extraction: Approaches and applications,' IEEE Transactions on Neural Networks and Learning Systems, pp. 1–21. https://doi.org/10.1109/tnnls.2022.3213168.
2. Liu, P. et al. (2023) 'Pre-train, prompt, and predict: A systematic survey of prompting methods in Natural Language Processing', ACM Computing Surveys, 55(9), pp. 1–35. doi:10.1145/3560815.
3. Zheng, S. et al. (2019) 'Doc2EDAG: An end-to-end document-level framework for Chinese financial event extraction', Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) [Preprint]. doi:10.18653/v1/d19-1032.
4. Zhu, T. et al. (2022) 'Efficient document-level event extraction via pseudo-trigger-aware pruned complete graph', Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, pp. 4552–4558. doi:10.24963/ijcai.2022/632.
5. Du, X., Rush, A.M. and Cardie, C. (2021) Grit: Generative role-filler transformers for document-level event entity extraction, arXiv.org. Available at: https://arxiv.org/abs/2008.09249 (Accessed: 20 October 2024).
6. Lu, S. et al. (2022) 'Explainable document-level event extraction via back-tracing to sentence-level event clues', Knowledge-Based Systems, 248, p. 108715. doi:10.1016/j.knosys.2022.108715.
7. Lu, Y. et al. (2021) 'Text2event: Controllable sequence-to-structure generation for end-to-end event extraction', Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) [Preprint]. doi:10.18653/v1/2021.acl-long.217.
8. Han, C. et al. (2022) 'Duee-fin: A large-scale dataset for document-level event extraction', Lecture Notes in Computer Science, pp. 172–183. doi:10.1007/978-3-031-17120-8_14.
9. Xu, R. et al. (2021) 'Document-level event extraction via heterogeneous graph-based interaction model with a tracker', Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) [Preprint]. doi:10.18653/v1/2021.acl-long.274.
10. Liang, Y. et al. (2022) 'Raat: Relation-augmented attention transformer for relation modeling in document-level event extraction', Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies [Preprint]. doi:10.18653/v1/2022.naacl-main.367.
11. Zhu, M. et al. (2024) CMNEE: A large-scale document-level event extraction dataset based on open-source Chinese military news, arXiv.org. Available at: https://arxiv.org/abs/2404.12242 (Accessed: 20 October 2024).
12. Hu, E.J. et al. (2021) Lora: Low-rank adaptation of large language models, arXiv.org. Available at: https://arxiv.org/abs/2106.09685 (Accessed: 20 October 2024).
13. Li, X.L. and Liang, P. (2021) 'Prefix-tuning: Optimizing continuous prompts for generation', Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) [Preprint]. doi:10.18653/v1/2021.acl-long.353.
14. Liu, X. et al. (2023) GPT understands, too, arXiv.org. Available at: https://arxiv.org/abs/2103.10385 (Accessed: 20 October 2024).