Efficient Document-level Event Extraction via Pseudo-Trigger-aware Pruned Complete Graph

Supplementary Material

Abstract

In this supplementary material, we introduce more about the implementation details for better understanding our approach and reproducibility.

- Section I provides details of event detection component as used in Doc2EDAG (Zheng et al. 2019), GIT (Xu et al. 2021) and PTPCG;
- Section II lists the regular expressions used in entity extraction section in the main content;
- Section III introduces the pseudo trigger selection and pruned complete graph building process with an example;
- Section IV lists the pseudo trigger roles and corresponding importance scores of the ChFinAnn (Zheng et al. 2019) dataset;
- Section V introduces more about the DAG-based methods and differences compared with PTPCG;
- Section VI introduces the Bron-Kerbosch algorithm;
- Section VII lists all the hyper-parameters used in the experiments;
- Section VIII provides the results of Doc2EDAG, GIT and PTPCG on the development set during training;
- Section IX shows more inference speed test results;
- Section X provides additional PTPCG experimental results.

I Event Detection Details

As introduced in the main content, event detection is a multi-label classification task. In this section, we explain more details about the event type classification.

For all sentence representations $\mathcal{G} = \{g_i\}_{i=1}^{|\mathcal{D}|}$, we follow the same method as reported in Doc2EDAG and GIT. An event query $q_j \in \mathbb{R}^{1\times d_h}$ of type t_j is applied to an event-related document representation $\mathcal{P}_j$ via the multi-head attention (Vaswani et al. 2017) mechanism with query, key and value as inputs:

$$\mathcal{P}_j = \text{MultiHeadAttention}(q_j, \mathcal{G}, \mathcal{G}) \in \mathbb{R}^{1\times d_h} \tag{I.1}$$

$$p(\mathcal{P}_j|\mathcal{G}) = \text{sigmoid}(\mathcal{P}_j \times W_p^\top + b_p) \tag{I.2}$$

where $W_p \in \mathbb{R}^{d_h\times d_h}$ and $b_p \in \mathbb{R}^{d_h}$ are trainable parameters.

The cross entropy function is used to calculate the loss of event detection component. To get final prediction results, an argmax operation is applied for each type.

$$\mathcal{L}_{det} = -\sum_{j=1}^{|\mathcal{T}|} y_{type} \log p(\mathcal{P}_j|\mathcal{G}) \tag{I.3}$$

$$o_{type}^j = \text{argmax}\,(p(\mathcal{P}_j|\mathcal{G})) \tag{I.4}$$

where $|\mathcal{T}|$ denotes the number of event types, y_{type} is the golden type labels, and o_{type}^j is the event type prediction.

II Regular Expression for Additional Entities

In the entity extraction component, we add additional entities into the dataset for data augmentation. These additional entities are matched via regular expressions as listed in Table II.1.

Type	Regular Expression
money	`\d+(\.\d+)?元`
date	`\d{4}年\d{1,2}月\d{1,2}日`
percentage ratio	`\d+(\.\d+)?%`
shares	`(\d+股)[^票]`

Table II.1: Regular expressions for additional entities.

III Example of Pseudo Trigger Selection and Pruned Complete Graph Building

For better explaining the pseudo trigger selection and pruned complete graph building, we present a simple example in Figure X.4.

IV Pseudo Triggers and Importance Scores

All the pseudo trigger roles $\mathcal{R}$ and importance scores for each type in the ChFinAnn dataset are listed in Table X.3.

V Different Event Argument Combination Methods

Figure V.1 shows different event argument combination strategies.

With annotated triggers, it is easy to identify combinations by building trigger-centered trees as shown on the top of Figure V.1 (Chen et al. 2015; Nguyen, Cho, and Grishman 2016; Liu, Luo, and Huang 2018; Wadden et al. 2019; Lin et al. 2020). However, annotating the triggers and identifing the event records in documents are difficult and costly. To speedup the annotaion process, Chen et al. (2017) borrow the idea of distant supervision (DS) (Mintz et al. 2009) and use an event knowledge base to align raw texts and generate records automatically. However, it is difficult to match the automatically extracted triggers with DS-constructed event records, and the triggers are often not available.

DAG-based argument combination strategy is shown in the left bottom of Figure V.1. For each argument role in type t_i, all the predicted entities are classfied via a type-specific role classifier to determine whether such entities are arguments of this role. If there are multiple entities recognized as the same argument role, the DAG is split into two branches. Finally, each path represents a possible argument combination. This strategy makes significant progress in trigger-free document-level event extraction, but the resource consumption shortcoming is obvious. To build DAGs, devices have to store all the previous paths, and each node in the graph has different previous memories. This auto-regressive decoding strategy requires massive memories and huge computing resources, so the training and inference speed are both very slow.

We instead use the pruned complete graphs (right bottom of Figure V.1) to represent event argument combinations and further exploit the non-autoregressive decoding algorithm to extract final records. The experimental results show that our approach is fast in both training and inference.

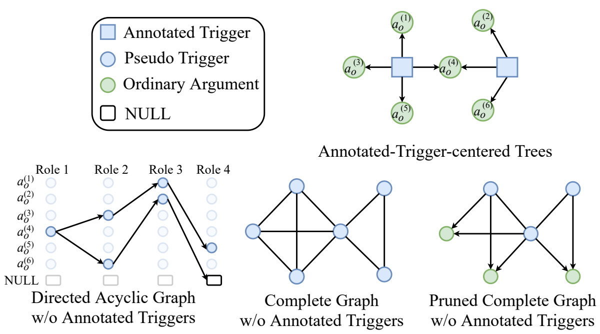


Figure V.1: Event argument combination comparison (best viewed in color).

VI Bron-Kerbosch Algorithm

In the combination extraction section of the main content, we introduce the non-autoregressive combination decoding algorithm. In this section, we list more details about the Bron-Kerbosch algorithm (Bron and Kerbosch 1973) used at line 17 of Algorithm 1 in the main content. To apply the recursive Algorithm VI.1, the input of adjacent matrix A must be symmetric, thus we performs the operation at line 16 of Algorithm 1. After that, the Algorithm VI.1 is applied to find possible pseudo trigger cliques.

Algorithm VI.1: Bron-Kerbosch

Input: Adjacent matrix A.
Output: Pseudo-trigger cliques $\mathcal{C}$.

for $c \in$ BK-Kernel$(\varnothing, \{1, 2, \ldots, |A|\}, \varnothing)$ **do**
  $\mathcal{C} \leftarrow \mathcal{C} \vee c$
end for

Algorithm VI.2: BK-Kernel

Input: Possible clique of arguments $\mathcal{Q}$, set of candidate vertices $\mathcal{S}$, set of excluded vertices $\mathcal{X}$
Output: A clique of vertices.

if $\mathcal{S} = \varnothing$ and $\mathcal{X} = \varnothing$ and $|Q| \geq 2$ **then**
  return $\mathcal{Q}$
end if
if $\mathcal{S} \cup X = \varnothing$ **then**
  $\mathcal{W} \leftarrow \mathcal{S}$
else
  $u \leftarrow$ random choice from $\mathcal{S} \cup X$
  // here $\mathcal{N}(u)$ denotes the neighbors of u
  $\mathcal{W} \leftarrow \mathcal{S} \backslash \mathcal{N}(u)$
end if
for $v \in \mathcal{W}$ **do**
  $\mathcal{F} \leftarrow \mathcal{S} \cap \mathcal{N}(v)$
  $\mathcal{H} \leftarrow \mathcal{X} \cap \mathcal{N}(v)$
  return BK-Kernel($\mathcal{Q} \vee v$, $\mathcal{F}$, $\mathcal{H}$)
  $\mathcal{S} \leftarrow \mathcal{S} \backslash v$
  $\mathcal{X} \leftarrow \mathcal{X} \vee v$
end for

VII Hyper-Parameters

We list all the hyper-parameters in Table VII.2 for reproducibility.

VIII Results on Development Set in Training

As Figure VIII.2 shows, PTPCG is the fastest to get convergence and reach the max micro-F1 score at the 57th epoch, while Doc2EDAG and GIT suffers from parameter updating in the first 20 epochs and get the max scores on the development set until the 82th and the 96th epoch.

IX Inference Speed Comparison

Figure IX.3 shows the inference speed test results on Doc2EDAG, GIT, TransPTPCG ($|\mathcal{R}| = 1$) and PTPCG with different bach sizes. Specially, GIT raises the “Out of Memory” error when batch size is 128. Our PTPCG outperforms Doc2EDAG and GIT in all batch size settings.

X Additional Experimental Results

We report additional performance results of PTPCG, including overall scores for each event type (Table X.4), performance on single & multiple record documents (Table X.5) and results of each seperate component (Table X.6).

We also provide a case predicted by PTPCG ($|\mathcal{R}| = 3$) in Figure X.5. This example is picked from ChFinAnn. To better show this example, we have reduced the document length.

Name	Setting
learning rate	5e-4
batch size	64
epoch	100
embedding dim d_h	768
mention type dim d_l	32
max sentence length	128
max number of sentence	64
minimal teacher prob	0.1
scheduled sampling start epoch	10
scheduled sampling end epoch	20
α_1	0.05
α_2	1.0
α_3	1.0
α_4	1.0
γ	0.5
#heads of multi-head att in event detection	1
layers of BiLSTM for token encoding	2
layers of BiLSTM for entitiy encoding	2

Table VII.2: Hyper-parameter details for implementation.

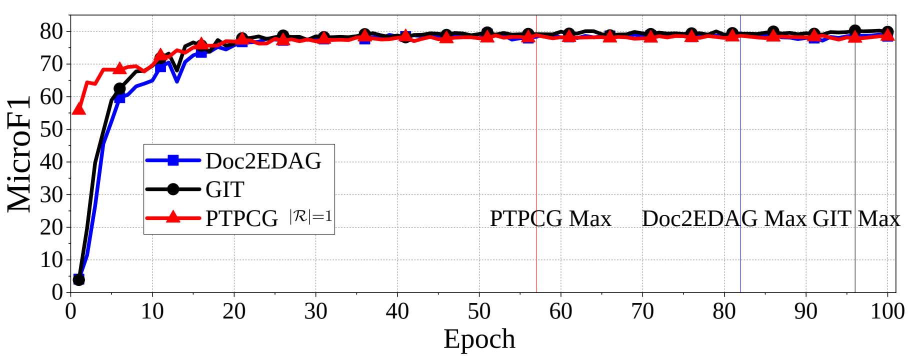


Figure VIII.2: Micro F1-scores on development set during training.

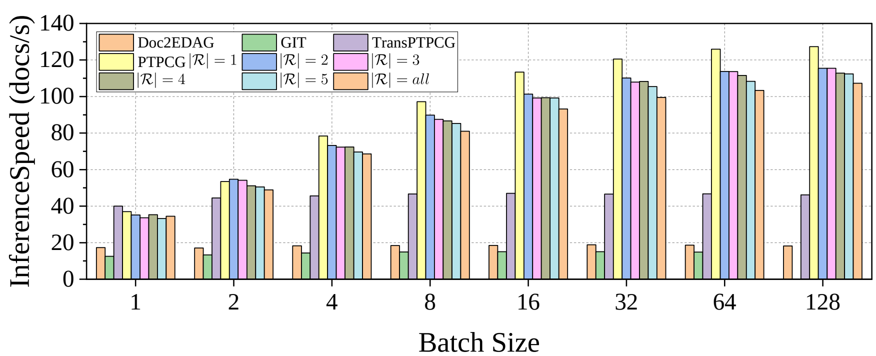


Figure IX.3: Inference speed comparison (best viewed in color).

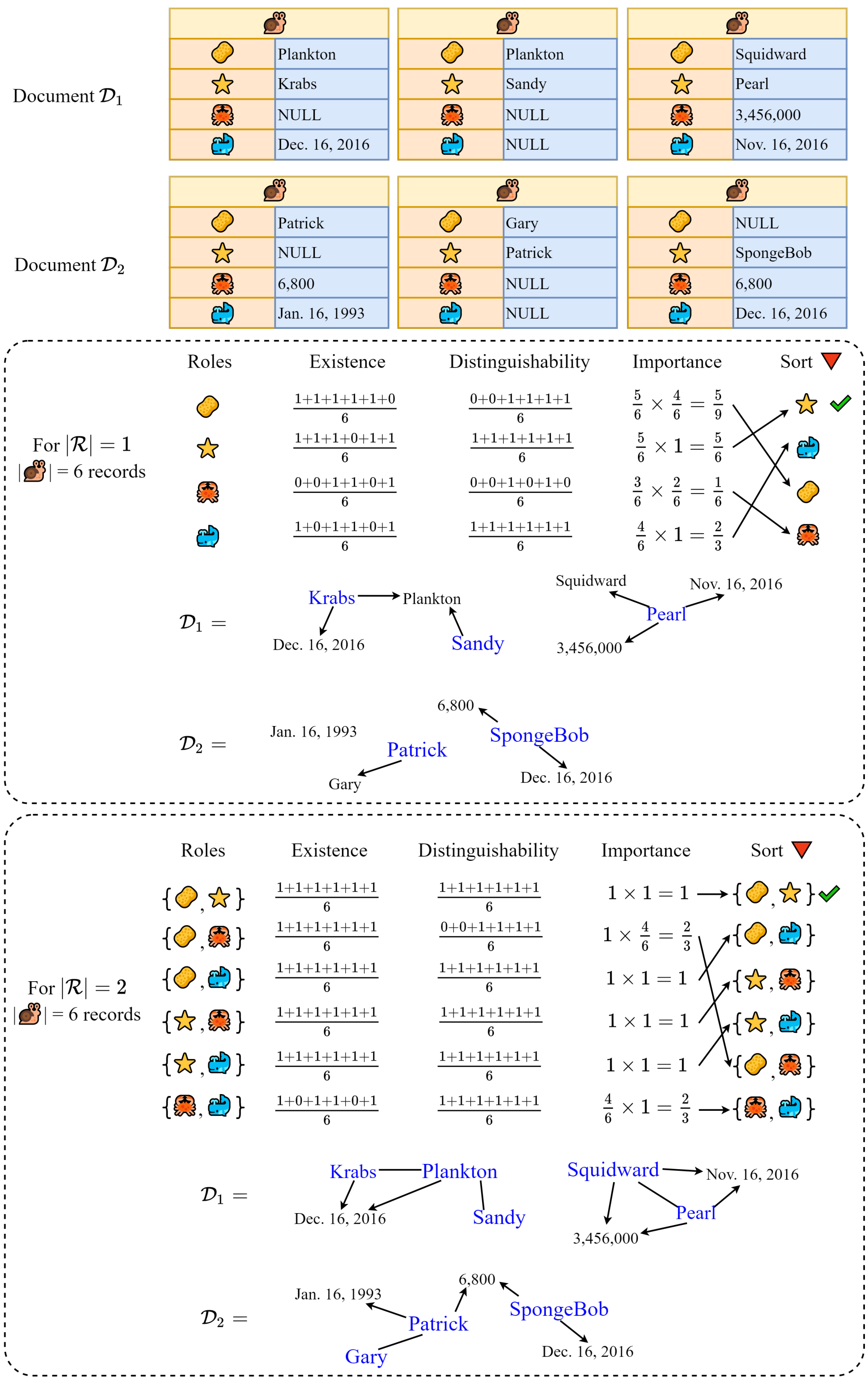


Figure X.4: Pseudo trigger selection and pruned complete graph building.

Event Type	$\lvert\mathcal{R}\rvert$	Importance (%)	Pseudo Trigger Roles
EquityPledge	1	88.47	PledgedShares
	2	92.39	PledgedShares, StartDate
	3	94.08	EndDate, PledgedShares, StartDate
	4	95.00	EndDate, PledgedShares, ReleasedDate, StartDate
	5	95.43	EndDate, PledgedShares, Pledgee, ReleasedDate, StartDate
	6	95.61	EndDate, PledgedShares, Pledgee, Pledger, ReleasedDate, StartDate
	7	95.65	EndDate, PledgedShares, Pledgee, Pledger, ReleasedDate, StartDate, TotalPledgedShares
EquityRepurchase	1	97.18	RepurchasedShares
	2	99.19	RepurchaseAmount, RepurchasedShares
	3	100.00	ClosingDate, RepurchaseAmount, RepurchasedShares
	4	100.00	ClosingDate, CompanyName, RepurchaseAmount, RepurchasedShares
	5	100.00	ClosingDate, CompanyName, HighestTradingPrice, RepurchaseAmount, RepurchasedShares
	6	100.00	ClosingDate, CompanyName, HighestTradingPrice, LowestTradingPrice, RepurchaseAmount, RepurchasedShares
EquityOverweight	1	93.25	TradedShares
	2	99.60	EquityHolder, StartDate
	3	99.80	EquityHolder, StartDate, TradedShares
	4	100.00	EquityHolder, LaterHoldingShares, StartDate, TradedShares
	5	100.00	EndDate, EquityHolder, LaterHoldingShares, StartDate, TradedShares
	6	100.00	AveragePrice, EndDate, EquityHolder, LaterHoldingShares, StartDate, TradedShares
EquityFreeze	1	70.61	LegalInstitution
	2	87.58	FrozeShares, LegalInstitution
	3	92.42	FrozeShares, LegalInstitution, StartDate
	4	93.03	EquityHolder, FrozeShares, LegalInstitution, StartDate
	5	93.33	EquityHolder, FrozeShares, LegalInstitution, StartDate, TotalHoldingRatio
	6	93.64	EquityHolder, FrozeShares, LegalInstitution, StartDate, TotalHoldingRatio, UnfrozeDate
	7	93.64	EquityHolder, FrozeShares, LegalInstitution, StartDate, TotalHoldingRatio, TotalHoldingShares, UnfrozeDate
	8	93.64	EndDate, EquityHolder, FrozeShares, LegalInstitution, StartDate, TotalHoldingRatio, TotalHoldingShares, UnfrozeDate
EquityUnderweight	1	91.88	TradedShares
	2	99.74	EndDate, EquityHolder
	3	99.74	EquityHolder, StartDate, TradedShares
	4	99.74	EndDate, EquityHolder, StartDate, TradedShares
	5	99.74	EndDate, EquityHolder, LaterHoldingShares, StartDate, TradedShares
	6	99.74	AveragePrice, EndDate, EquityHolder, LaterHoldingShares, StartDate, TradedShares

Table X.3: Pseudo trigger roles with highest importance scores for each type in ChFinAnn.

$\lvert\mathcal{R}\rvert$	EquityFreeze			EquityRepurchase			EquityUnderweight			EquityOverweight			EquityPledge			Average			Total (micro)		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
1	80.8	63.9	71.4	90.9	92.3	91.6	80.2	64.6	71.5	77.5	67.5	72.2	81.7	71.7	76.4	82.2	72.0	76.6	83.7	75.4	79.4
2	75.0	66.0	70.2	91.3	90.4	90.8	73.2	64.4	68.5	74.4	68.4	71.3	74.7	74.5	74.6	77.7	72.8	75.1	78.6	76.7	77.7
3	68.5	69.6	69.0	86.5	91.2	88.8	70.4	67.0	68.6	63.7	68.1	65.9	68.2	75.3	71.6	71.5	74.2	72.8	72.3	77.7	74.9
4	65.9	67.4	66.6	85.9	91.2	88.5	61.9	67.3	64.5	61.2	69.1	64.9	67.5	75.0	71.1	68.5	74.0	71.1	70.9	77.5	74.0
5	56.5	69.9	62.5	82.6	92.8	87.4	65.7	67.4	66.5	63.3	68.4	65.8	64.1	77.0	69.9	66.4	75.1	70.4	67.9	79.1	73.1
all	63.9	71.6	67.5	81.7	92.0	86.6	64.0	65.9	64.9	59.9	68.8	64.0	55.8	76.0	64.4	65.1	74.9	69.5	62.4	78.4	69.5

Table X.4: PTPCG overall scores on ChFinAnn.

$\lvert\mathcal{R}\rvert$	EquityFreeze		EquityRepurchase		EquityUnderweight		EquityOverweight		EquityPledge		Average		Total (micro)	
	S	M	S	M	S	M	S	M	S	M	S	M	S	M
1	83.6	59.9	93.7	73.8	77.3	63.6	79.7	62.8	86.1	70.5	84.1	66.1	88.2	69.1
2	77.4	63.7	93.2	70.8	76.7	57.3	76.9	64.1	83.7	69.3	81.6	65.0	86.3	68.1
3	73.6	64.7	91.0	69.4	74.9	60.4	74.7	55.6	79.3	67.2	78.7	63.5	83.1	65.9
4	71.2	62.3	90.9	68.3	70.7	56.3	69.9	58.9	80.3	65.8	76.6	62.3	82.6	64.7
5	63.2	61.8	89.6	70.3	71.4	60.0	71.9	58.3	79.7	64.6	75.2	63.0	81.6	64.0
all	69.5	65.7	88.6	69.9	69.4	58.8	69.4	57.4	73.6	59.3	74.1	62.2	78.7	60.1

Table X.5: PTPCG performance on documents with single and multiple records in ChFinAnn.

$\lvert\mathcal{R}\rvert$	EventDetection			EntityExtraction			Combination			Overall		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
1	99.0	98.8	98.9	97.5	99.4	98.4	40.9	40.9	40.9	83.8	75.4	79.4
2	99.0	98.9	99.0	97.3	99.5	98.4	35.4	40.2	37.7	78.6	76.7	77.7
3	99.2	98.9	99.0	97.4	99.4	98.4	32.5	41.4	36.4	72.3	77.7	74.9
4	99.2	98.9	99.0	97.4	99.5	98.4	31.1	40.7	35.2	70.9	77.5	74.0
5	99.1	98.9	99.0	97.5	99.4	98.4	31.6	43.8	36.7	67.9	79.1	73.1
all	99.1	99.0	99.1	97.3	99.5	98.4	26.4	41.4	32.2	62.4	78.4	69.5

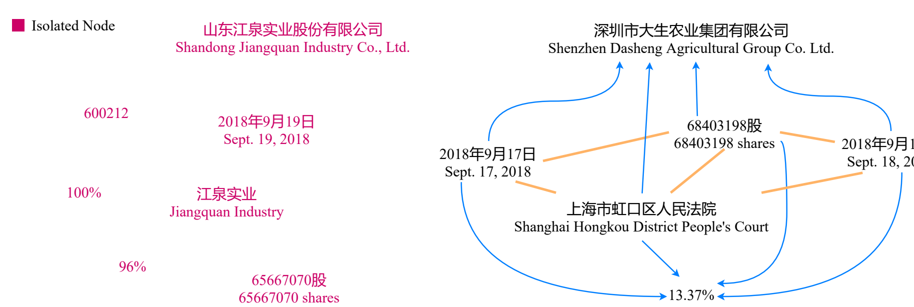
Table X.6: PTPCG performance of each component on ChFinAnn.

Entity Mention
Entity Mention w/ Multiple Appearances

Document

[1]	证券代码：600212证券简称：江泉实业编号：临2018 - 046 Securities code: 600212 Securities abbreviation: Jiangquan Industry No.: Lin 2018-046
[2]	山东江泉实业股份有限公司关于公司控股股东股份新增轮候冻结的公告 Announcement of Shandong Jiangquan Industry Co., Ltd. on the New Waiting List Freeze of Shares of the Company's Controlling Shareholder
[3]	本公司董事会及全体董事保证本公告内容不存在任何虚假记载、误导性陈述或者重大遗漏，并对其内容的真实性、准确性和完整性承担个别及连带责任。 The Board of Directors and all directors of the Company warrant that the contents of this announcement do not contain any false statements, misleading statements or material omissions, and shall bear individual and joint responsibility for the truthfulness, accuracy and completeness of its contents.
[4]	山东江泉实业股份有限公司（以下简称"公司"）于2018年9月19日收到控股股东深圳市大生农业集团有限公司（简称"大生农业集团"）通知，根据其于9月19日在中国证券登记结算有限责任公司的查询结果，大生农业集团所持公司股份新增轮候冻结，具体情况如下： Shandong Jiangquan Industry Co., Ltd. (hereinafter referred to as "the Company") received a notice from its controlling shareholder, Shenzhen Dasheng Agricultural Group Co. Ltd. ("Dasheng Agricultural Group") on Sept. 19, 2018, according to its query results in the China Securities Depository and Clearing Corporation on Sept. 19, the shares of the company held by Dasheng Agricultural Group were newly frozen on waiting list, as follows.
[5]	一、控股股东股份新增轮候冻结情况 I. Newly frozen shares of controlling shareholders
[6]	1、轮候冻结机关：上海市虹口区人民法院，轮候冻结起始日：2018年9月17日，轮候冻结数量：68403198股，轮候冻结期限：3年，自转为正式冻结之日计算。 1. Waiting freezing authority: Shanghai Hongkou District People's Court, starting date of waiting freeze: Sept. 17, 2018, number of waiting freeze: 68403198 shares, waiting freeze period: 3 years, calculated from the date of transfer to official freeze.
[7]	2、轮候冻结机关：上海市虹口区人民法院，轮候冻结起始日：2018年9月18日，轮候冻结数量：68403198股，轮候冻结期限：3年，自转为正式冻结之日计算。 2. Waiting freezing authority: Shanghai Hongkou District People's Court, starting date of waiting freeze: Sept. 18, 2018, number of waiting freeze: 68403198 shares, waiting freeze period: 3 years, calculated from the date of conversion into official freeze.
[8]	二、控股股东股份累计被质押、司法冻结及轮候冻结的情况 II. Accumulated pledge, judicial freeze and waiting freeze of controlling shareholders' shares
[9]	截至本公告披露日，大生农业集团持有本公司股份68403198股，占公司总股本的13.37%，处于质押状态的股份数量为65667070股，占其持有公司股份总数的96%；处于冻结状态的股份数量为68403198股，占其持有公司股份总数的100%。 As of the disclosure date of this announcement, Dasheng Agricultural Group holds 68403198 shares of the Company, accounting for 13.37% of the total share capital of the Company, the number of shares under pledge is 65667070 shares, accounting for 96% of the total number of shares held by it; the number of shares under freeze is 68403198 shares, accounting for 100% of the total number of shares held by it.
[10]	三、控股股东股份新增轮候冻结的影响及风险提示 III. Impact of the newly frozen shares of the controlling shareholder and risk indication
[11]	控股股东大生农业集团所持公司股份新增轮候冻结事项尚未对公司的正常运行和经营管理造成影响，但其冻结股份若被司法处置，可能导致公司实际控制权发生变更。 The freezing of shares held by the controlling shareholder Dasheng Agricultural Group has not yet affected the normal operation and management of the company, but if the frozen shares are disposed of judicially, it may lead to a change in the actual control of the company.
[12]	大生农业集团正积极协商处理股份冻结及轮候冻结等相关事项，争取尽快解除对本公司股份的冻结或轮候冻结。 The Dasheng Agricultural Group is actively negotiating the freezing and waiting list freezing of shares and other related matters, and will strive to lift the freezing or waiting list freezing of the Company's shares as soon as possible.
[13]	山东江泉实业股份有限公司董事会 Board of Directors of Shandong Jiangquan Industry Co. Ltd.
[14]	二〇一八年九月二十日 Sept. 20, 2018

Event Argument Combination



Event Type
Argument Role
Argument

Event Records

EquityFreeze	
EquityHolder	深圳市大生农业集团有限公司 Shenzhen Dasheng Agricultural Group Co. Ltd.
FrozeShares	68403198股 68403198 shares
LegalInstitution	上海市虹口区人民法院 Shanghai Hongkou District People's Court
TotalHoldingShares	68403198股 68403198 shares
TotalHoldingRatio	13.37%
StartDate	2018年9月17日 Sept. 17, 2018
EndDate	NULL
UnfrozeDate	NULL

EquityFreeze	
EquityHolder	深圳市大生农业集团有限公司 Shenzhen Dasheng Agricultural Group Co. Ltd.
FrozeShares	68403198股 68403198 shares
LegalInstitution	上海市虹口区人民法院 Shanghai Hongkou District People's Court
TotalHoldingShares	68403198股 68403198 shares
TotalHoldingRatio	13.37%
StartDate	2018年9月18日 Sept. 18, 2018
EndDate	NULL
UnfrozeDate	NULL

Figure X.5: PTPCG ($|\mathcal{R}| = 3$) prediction.

References

Bron, C.; and Kerbosch, J. 1973. Algorithm 457: Finding All Cliques of an Undirected Graph. *Commun. ACM*, 16(9): 575–577.

Chen, Y.; Liu, S.; Zhang, X.; Liu, K.; and Zhao, J. 2017. Automatically Labeled Data Generation for Large Scale Event Extraction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 409–419. Vancouver, Canada: Association for Computational Linguistics.

Chen, Y.; Xu, L.; Liu, K.; Zeng, D.; and Zhao, J. 2015. Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 167–176. Beijing, China: Association for Computational Linguistics.

Lin, Y.; Ji, H.; Huang, F.; and Wu, L. 2020. A Joint Neural Model for Information Extraction with Global Features. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7999–8009. Online: Association for Computational Linguistics.

Liu, X.; Luo, Z.; and Huang, H. 2018. Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 1247–1256. Brussels, Belgium: Association for Computational Linguistics.

Mintz, M.; Bills, S.; Snow, R.; and Jurafsky, D. 2009. Distant supervision for relation extraction without labeled data. In Su, K.; Su, J.; and Wiebe, J., eds., *ACL 2009, Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP, 2-7 August 2009, Singapore*, 1003–1011. The Association for Computer Linguistics.

Nguyen, T. H.; Cho, K.; and Grishman, R. 2016. Joint Event Extraction via Recurrent Neural Networks. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 300–309. San Diego, California: Association for Computational Linguistics.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. In *Advances in neural information processing systems*, 5998–6008.

Wadden, D.; Wennberg, U.; Luan, Y.; and Hajishirzi, H. 2019. Entity, Relation, and Event Extraction with Contextualized Span Representations. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 5784–5789. Hong Kong, China: Association for Computational Linguistics.

Xu, R.; Liu, T.; Li, L.; and Chang, B. 2021. Document-level Event Extraction via Heterogeneous Graph-based Interaction Model with a Tracker. In *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)*, 3533–3546. Online: Association for Computational Linguistics.

Zheng, S.; Cao, W.; Xu, W.; and Bian, J. 2019. Doc2EDAG: An End-to-End Document-level Framework for Chinese Financial Event Extraction. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 337–346. Hong Kong, China: Association for Computational Linguistics.