2015 8th International Symposium on Computational Intelligence and Design

Event Relations Extraction Based on Event Co-occurrence Network

Tao Liao[1], Peipei Sun[1], Zongtian Liu[2]
[1]School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, China
[2]School of Computer Engineering and Science, Shanghai University, Shanghai, China
E-mail: tliao@aust.edu.cn, ppsun@aust.edu.cn, ztliu@shu.edu.cn

***Abstract*—Cognitive scientists believe that humans memorize and understand the real world through "event". Event relations extraction has become increasingly important in some natural language processing applications. In this paper, we firstly consider the event as a basic semantic unit and present a new event co-occurrence network. Then, we study event relations extraction based on this event co-occurrence network. We used the association rule mining method to extract event co-occurrence pairs from event co-occurrence networks, and got the semantic relations between event classes after generalizing and analyzing these event co-occurrence pairs. The experimental results show that our event relations extraction method has good performance.**

Keywords—Event co-occurrence; Event relations extraction; Non-taxonomic relations; Chinese Emergency Corpus

I. INTRODUCTION

With the in-depth research for the event, more and more people pay attention to the analysis and knowledge mining of event relations in recent years. Event relations extraction has become increasingly important in some natural language processing applications, such as information retrieval, automatic summarization, and question answering and so on.

Studies of event relations are mostly focused on taxonomic relations and temporal relations at present, and they are also two kinds of common event relations in texts. The taxonomic relation reflects a static characteristic between events, and the dynamic relation between events is more represented by the temporal relation.

As the event taxonomic relation, there have been some typical researches. Vargas-Vera and Celjuska divided event into 41 different types and specified the corresponding slot for each event class, and they established the hierarchical relation between event classes at the same time [1]. VerbNet (an online dictionary) was built by Kipper Karin in 2005, and he studied the taxonomic relation between verbs and spread 5200 verbs over 237 top classes [2]. The ACE evaluation conference divided the event into 8 big classes and 33 sub classes [3].

The establishment of events in time order relation is the purpose of studying event temporal relation, which begins with the extraction of time expressions in the text. I. Mani introduced time expressions can be tagged, the temporal structure of events and other related time information extraction [4]. Wang used the method based on machine learning to determine the corresponding time expressions of events and the mapping relation between time information and event information in the text [5]. In recent years, research of events temporal relation mainly focused on TimeBank [6] and TempEval [7]. TimeBank is a corpus which uses TimeML as tagging language. TempEval is a task of SemEval-2007, and the temporal relations between events can eventually be determined in the document through the three sub tasks.

In addition, some scholars have studied event causal relation [8-12]. The event causal relation is also a special temporal relation. Khoo distinguished event causal relation for explicit and implicit relation [9]; Gan proposed a structure analysis method based on event causal relation [10]; Blanco defined causal relation between events and introduced the machine learning technique to extract the marked causal relation [11]. Besides, Fu presented a method for event causal relation extraction by using dual-layer Conditional Random Fields (CRFs) [12].

Clearly, the above studies can only cover part of event relations in the text. The relations between events can also be divided into taxonomic relations and Non-taxonomic relations. Non-taxonomic relations can be further subdivided into composition relation, causality relation, follow relation, and accompany relation and so on. They can not only reflect the time order relation between events, but also can very well reflect the deep semantic relations between events. This paper presents a new event co-occurrence network, and then extracts event relations on this basis.

II. RELATED WORK

A. Event Relations

In this paper, we divide event relations into two categories: taxonomic relations and Non-taxonomic relations, and give the following definitions.

Definition1 (*Taxonomic Relations*) Suppose two event classes, EC_1 and EC_2. There exists taxonomic relation between EC_1 and EC_2 if $EC_2 \subset EC_1$. We call EC_1 hypernymy event class and EC_2 hyponym event class.

Definition2 (*Non-taxonomic Relations*) There are four main types of Non-taxonomic relations of event.

(1) (***Composite Relation***) If a big event e can be decomposed to several small events e_i (i>0), and these small event have been finished means e is finished, then there exists composite relation between e and e_i (i>0). If each event instance of event class EC_1 is composed of one event instance of event class EC_2 and other event classes, then EC_1 composed of EC_2, denoted as $EC_2 < EC_1$.

(2) (***Causal Relation***) If an event instance of EC_1 happens, event in EC_2 is more likely to happen at above a

DOI 10.1109/ISCID.2015.125

specified probability threshold, there is a causal relation between EC_1 and EC_2. EC_1 is cause and EC_2 is effect, denoted as $EC_1 \Rightarrow EC_2$.

(3) (***Follow Relation***) Follow refers to events coming after in time or order, as a consequence or result, or by the operation of logic. If in a certain length of time, EC_2 follow EC_1 at above a specified probability threshold, there is a follow relation between EC_1 and EC_2, denoted as $EC_2 \rhd EC_1$.

(4) (***Accompany Relation***) If EC_1 concur with EC_2 in a certain period of time, and the occurrence probability is above a specified threshold, there is an accompany relation between EC_1 and EC_2, denoted as $EC_2 \parallel EC_1$.

B. Event Co-occurrence

Co-occurrence analysis has two theoretical bases: one is the adjacent theory of psychology; the other is knowledge structure and mapping principle of sociology. WordNet semantic dictionary developed by Princeton University gives the definition of co-occurrence: when used as a common noun, co-occurrence represents the simultaneous events, or events that are associated with each other [13]. Visible, the interconnection of things is the inherent reason of co-occurrence, and the co-occurrence phenomenon is the external manifestation of things' interconnection..

Definition 3 (*Event Co-occurrence Rate*): The ***event co-occurrence rate*** refers to the probability of events e_i and e_j co-occur in the same window unit and is denoted by $P(e_i, e_j)$, defined as follows:

$$P(e_i, e_j) = \frac{S(e_i, e_j)}{n(S)} \tag{1}$$

Where $S(e_i, e_j)$ represents the number of window unit that contains both events e_i and e_j in the text corpus, $n(S)$ represents the total number of window unit included in the text corpus.

C. Chinese Emergency Corpus

Definition 4 (*Event Denoter*): The ***event denoter*** is a word that can clearly express the happening of the event, i.e. the event action element [14].

<Event eid="e1">
<Time type="absTime">April 14, 2009</Time>,
<Denoter type="emergency"> traffic accident </Denoter>
occurred on
<Location>the Hangzhou-Shanghai expressway </Location>
</Event>,
<Event eid="e2">
<Participant>4 people</Participant>
<Denoter type="stateChange"> died </Denoter>
</Event>,
<Event eid="e3">
<Participant>2 people</Participant>
<Denoter type="stateChange"> injured </Denoter>
</Event>.

Figure 1. The effect diagram of labeled events

III. Event Co-occurrence Network

The event co-occurrence network is an undirected graph, which is triples of the form $G = (N, E, W)$.

Taking into account that the length of a single news text is shorter, this paper uses respectively each ***CEC*** event topic class corpus to build the corresponding event co-occurrence network.

This paper extracts events and their action factors from the labeled event topic class corpus ***D*** to get the event features set $EV_D = \{e_1, ..., e_i, e_j, ..., e_k\}$ (Events with the same action factor are mapped to the same event feature), and constructs the event co-occurrence network on this basis.

Construction of the event co-occurrence network includes the following steps:

Step 1: Initializing the event co-occurrence network, the nodes set $N_D = \{\}$, the undirected edges set $E_D = \{\}$, and the weights set of undirected edges $W_D = \{\}$;

Step 2: Mapping event features of the event features set $EV_D = \{e_1, ..., e_i, e_j, ..., e_k\}$ to the nodes of the event co-occurrence network graph structure, and getting the nodes set $N_D = \{n_1, ..., n_i, n_j, ..., n_k\}$;

Step 3: Taking any two nodes n_i and n_j from N_D, and adding one undirected edge e_{ij} between n_i and n_j if their corresponding event features (e_i and e_j) occur in the same window unit (the window unit is set to a sentence in this paper), and then getting the undirected edges set $E_D = \{..., e_{ij}, ...\}$;

Step 4: Counting the weight of each undirected edge in the event co-occurrence network to get the weights set of undirected edges $W_D = \{..., w_{ij}, ...\}$, $w_{ij} = P(e_i, e_j)$.

Follow the steps above, we construct the event co-occurrence network of the event topic class corpus ***D***.

We select the traffic accident class corpus from the ***CEC***, and specify the construction process of the event co-occurrence network. By counting event features in the traffic accident class corpus, we found that event features are very much and sparse, and this is not conducive to analysis and learning of the event co-occurrence network. Therefore, we artificially constructed the synonymous event table to realize event class clustering of the synonymous event features. That is, the synonymous event features are mapped to the same event feature.

According to statistics the traffic accident class corpus contains 49 texts, 265 sentences and 105 event features. We generate the undirected edges by utilizing the event co-occurrence network construction method proposed in this paper, thus obtain 102 nodes (Removal of 3 isolated nodes) and 366 undirected edges. We apply the visualization software ***NetDraw*** to visualize the traffic accident class corpus, and each node is labeled by the corresponding event feature. The event co-occurrence network of the traffic accident class corpus is shown in Fig.2.

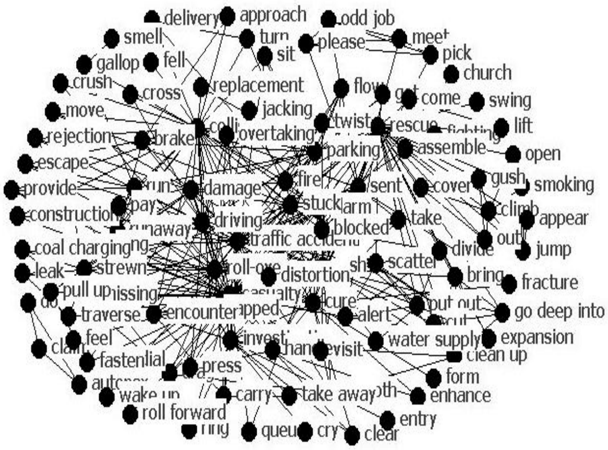

Figure 2. Event co-occurrence network of the traffic accidents corpus

IV. EVENT RELATION EXTRACTION

This paper made the extraction process of event co-occurrence pairs as a process of extracting event fixed semantic relation rules and used the association rule mining method to extract event co-occurrence pairs from the event co-occurrence network, and got the semantic relations between event classes after generalizing and analyzing these event co-occurrence pairs. Since this paper mainly excavates the relations among events, so here we only analyze the co-occurrence of phenomenon between any two event items, wherein support is expressed by Equation 2, confidence is expressed by Equation 3.

$$Support(e_i, e_j) = P(e_i, e_j) \tag{2}$$

$$Identify(e_i, e_j) = \max\left(\frac{P(e_i, e_j)}{P(e_i)}, \frac{P(e_i, e_j)}{P(e_j)}\right) \tag{3}$$

Support reflects the co-occurrence frequency of event co-occurrence pairs, which reflects the importance of event co-occurrence pairs in the text, but confidence reflects the association strength between the co-occurrence events. When we extract event co-occurrence pairs (event relations), it is required that support and confidence should be greater than the corresponding thresholds.

Here we still select the traffic accident class corpus as experimental data (the window unit is set to a sentence). Through observation, with support diminishing from 0.02, the number of event co-occurrence pairs also grows slowly, but when the support is 0.005, the number of event co-occurrence pairs increases rapidly. Similarly, when the confidence changes from 0.05 to 0.15, the number of event co-occurrence pairs decreases slowly, but when the confidence is 0.20, the number of event co-occurrence pairs reduces significantly, therefore we choose support 0.01 and confidence 0.15 as thresholds.

TABEL I. THE RELATION BETWEEN EVENT CLASSES (PART)

No.	Event co-occurrence pair		Relation type
	Event feature e_i	Event feature e_j	
1	collide	casualty	Casual relation
2	traffic accident	casualty	Casual relation
3	casualty	cure	Follow relation
4	traffic accident	collide	Taxonomic Relation
5	roll-over	casualty	Casual relation
6	send	cure	Follow relation
7	casualty	damage	Accompany relation
8	collide	roll-over	Casual relation
9	collide	Fire	Casual relation
10	out of control	collide	Casual relation
11	collide	rolling	Follow relation
12	traffic accident	blocking	Casual relation

According to the above thresholds, we can extract 37 event co-occurrence pairs, and We have summarized and analyzed them and found that there have 34 event relations among event classes as shown in Table I (In the construction processes of event co-occurrence network, every event has clustered into event classes), and then constructing the event relation table about event classes.

V. EXPERIMENTS AND DISCUSSION

Based on the CEC corpus, we further annotated event relations for five emergency classes. Each text was annotated and checked by two group annotators, and then we got the detailed information of various event relations, as shown in Table II.

TABEL II. EVENT RELATION INFORMATION OF THE CORPUS

Event relation type		Amount	Rate (%)
Taxonomic Relation		157	5.7
Non-taxonomic Relation	Composite	145	5.3
	Causal	1028	37.5
	Follow	876	32.0
	Accompany	533	19.5
Overall		2739	100

From Table II, it can be seem that there are a large number of causal relation (37.5%) , follow relation (32%) and accompany relation (19.5%), but taxonomic and composite relation are very few. Considering the sparseness problem of the corpus, we mainly extracted causal, follow and accompany relation in the experiment. The experimental results are measured by Precision, Recall and F-Measure and summarized in Table III.

TABEL III. EXPERIMENTAL RESULTS OF EVENT RELATION EXTRACTION

Event relation type	***P***(%)	***R***(%)	***F***(%)
Causal Relation	91.6	86.8	89.1
Follow Relation	88.3	83.2	85.7
Accompany Relation	89.2	84.7	86.9
Average	89.6	85.2	87.3

In addition, this paper compared the experimental results with some other research results at home and abroad, and our comparison results are listed in Table IV. As the use of different experimental corpus and the extraction type of event relation is not the same, this paper inconvenience direct compares our experimental results to domestic and foreign automatic summarization methods, but also achieves good results.

TABEL IV. THE RESULTS OF OUR METHOD COMPARED WITH OTHER STUDIES

Author	Experimental Data/ Event relation type	Experimental Results		
		P(%)	***R***(%)	***F***(%)
Kolya [16]	TempEval-2007 training set./ Temporal	56.9	56.9	56.9
Yang X R[17]	30 new articles collected from internet, and Each new article contains an average of 32 events./ Logical	59.5	67.4	63.2
Ding X S [18]	A corpus containing 1268 passages, 3008 sentences in total./ Causal	75	96	84.2
Our method	The CEC corpus, a total of 500 texts./ Causal , Follow and Accompany	89.6	85.2	87.3

VI. CONCLUSION

In this paper, we consider the event as a basic semantic unit and present a new event co-occurrence network structure based method for text representation. Then we made the extraction process of event co-occurrence pairs as a process of extracting event fixed semantic relation rules and used the association rule mining method to extract event co-occurrence pairs from event co-occurrence networks, and got the semantic relations between event classes after generalizing and analyzing these event co-occurrence pairs. The experimental results show that event relations extraction based on the event co-occurrence network has good performance.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their valuable comments. This paper is supported by the Natural Science Foundation of China (No. 60975033, 61273328).

REFERENCES

[1] Vargas-Vera M, Celjuska D. "Event recognition on news stories and semi-automatic population of an ontology ". Selected Advanced Knowledge Technology Papers, 2004,pp.615-618.

[2] Kipper K,Korhonen A,Ryant N,et al. "Extending VerbNet with novel verb classes". In: Proc. The Fifth International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy, 2006.

[3] Levin B. "English verb classes and alternations: A preliminary investigation". University of Chicago press Chicago, 1993.

[4] I. Mani. "Recent developments in temporal information extraction". In: Proc. The International Conference on Recent Advances in Natural Language Processing (RANLP'03), 2004, pp. 45-60.

[5] Wang Yun,Yuan Chunfa."A Time-Event Mapping Method Based Transformation". Journal of Chinese Information Processing, 2004, vol.18, pp.23-30.

[6] J. Pustejovsky, J. Castano, R. Ingria, et al. "TimeML: Robust specification of event and temporal expressions in text". In: Proc. New Directions in Question Answering, 2003, pp.28-34.

[7] M. Verhagen, R. Gaizauskas, F. Schilder, et al. "Semeval-2007 task 15: Tempeval temporal relation identification". In: Proc. The 4th International Workshop on Semantic Evaluations (SemEval-2007), Prague, June 2007, pp. 75-80.

[8] C Hashimoto, K Torisawa, J Kloetzer, et al. "Toward Future Scenario Generation: Extracting Event Causality Exploiting Semantic Relation, Context, and Association Features". In: Proc. The 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, Maryland, USA, 2014, pp.987-997.

[9] Khoo C., Kornfilt J., Oddy R., et al. "Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing". Literary and Linguistic Computing, 1998, vol.13, pp: 177-186.

[10] Gan Honghua, Pan Yunhe. "A New Analysis of the structure of Event Causation". PR & AI, 2003, vol.16, pp.56-62.

[11] Blanco E., Castell N., Moldovan D. "Causal relation extraction". In: Proc. The 6th International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco, 2008, pp. 310-313.

[12] Jianfeng Fu, Zongtian Liu, Wei Liu, and Qiang Guo. "Using dual-layer CRFs for event causal relation extraction". IEICE Electronics Express, 2011, vol.8, pp. 306-310.

[13] Co-occurrence.http://www.wordreference.com/definition/Co-occurrence.

[14] Fu, J., W. Liu, and Z. Liu. "A study of Chinese event taggability". In: Proc. The 2010 International Conference on Communication Software and Networks (ICCSN 2010), Singapore, 2010, pp.400-404.

[15] Liu Z T, Huang M L, Zhou W, et al. "Research on Event-oriented Ontology Model". Computer Science, 2009, vol.36, pp.189-192, 199.

[16] Kolya, Anup Kumar; Ekbal, Asif; Bandyopadhyay, Sivaji. "Event-event relation identification: A CRF based approach". In: Proc. The 6th International Conference on Natural Language Processing and Knowledge Engineering, NLP-KE, 2010.

[17] Yang Xuerong, Hong Yu, Ma Bin, Yao Jianmin, Zhu Qiaoming. "Event relation Recognition by Event term and Entity Inference". Journal of Chinese Information Processing, 2014, vol.28, pp.100-108.

[18] Ding Xiaoshan, Li Fang, Zhang Dongmo. "Causal Relation Recognition between Sentence-based Events". In: Proc. The 2011 Chinese Control and Decision Conference, 2011, pp. 1688-1693.