Event Relationship Analysis for Temporal Event Search

Yi Cai[1,⋆], Qing Li[2], Haoran Xie[2], Tao Wang[1], and Huaqing Min[1]

[1] School of Software Engineering, South China University of Technology, Guangzhou, China
ycai@scut.edu.cn

[2] Department of Computer Science, City University of Hongkong, Hongkong, China

Abstract. There are many news articles about events reported on the Web daily, and people are getting more and more used to reading news articles online to know and understand what events happened. For an event, (which may consist of several component events, i.e., episodes), people are often interested in the whole picture of its evolution and development along a time line. This calls for modeling the dependent relationships between component events. Further, people may also be interested in component events which play important roles in the event evolution or development. To satisfy the user needs in finding and understanding the whole picture of an event effectively and efficiently, we formalize in this paper the problem of temporal event search and propose a framework of event relationship analysis for search events based on user queries. We define three kinds of event relationships which are temporal relationship, content dependence relationship, and event reference relationship for identifying to what an extent a component event is dependent on another component event in the evolution of a target event (i.e., query event). Experiments conducted on a real data set show that our method outperforms a number of baseline methods.

1 Introduction

With the development of the Internet, news events are reported by many news articles in the form of web pages. People are getting more and more used to reading news articles online to know and understand what events happened. For a composite/complex event, it may consist of several component events, i.e., episodes. There are some interrelationships among these component events as they may be dependent on each other. For example, the event of "Toyota 2009-2010 vehicle recalls" contains several interrelated component events, e.g., the event "Toyota recall due to safety problems from 2009 to 2010" causes the happening of the event "NHTSA conduct investigations for Toyota recall" and the event "US congressional hearings hold for Toyota recall", and so on. Also, the event "US congressional hearings hold for Toyota recall" has a strong relationship with the event "Toyota's president to testify in US congressional hearings".

Quite often, what people interested in is not just a sole news article on an event, but also the related events reported by other news articles. Indeed, they are often interested in the whole picture of an event evolution or development along a timeline. This calls for modeling the dependence relationships between component events, and identifying

⋆ Corresponding author.

W. Meng et al. (Eds.): DASFAA 2013, Part II, LNCS 7826, pp. 179–193, 2013.

which component events play important roles in the entire event evolution or development. Unfortunately, the current news web sites do not facilitate people in finding out relevant news articles easily, and people may need to go through all these news articles in order to find out the interrelationships between component events. Current prevailing search engines (such as Google, Yahoo and so on) allow users to input event keywords as a query and return a list of news web pages related to the query. However, instead of organizing the result by events and relationships between events, these engines just provide users with a ranking list of news web pages. It is difficult and time consuming for users to view of the huge amount of news articles and to obtain the main picture of an event. Therefore, it is necessary to provide an effective way for users to efficiently search events they are interested in, and organize the search results in an easily understandable manner syntactically, so that users can obtain the main pictures of their interested events easily and meaningfully from the semantic perspective.

Although there have been some previous works attempting to find and link incidents in news [2] [3] or discover the event evolution graphs [17] [15], they only focus on time sequence and content similarity between two component events in identifying their dependence relationships. However, using these two factors only is inadequate in identifying dependence relationships among the component events in order to form the main picture of a big event evolution or development. For example, event "Toyota recall due to safety problems from 2009 to 2010" shares little similar content with events "NHTSA conduct investigations for Toyota recall" and "US congressional hearings hold for Toyota recall". It is obvious that the first event has a strong effect on the latter two events as it caused them to happen. Unfortunately, previous works do not analyze the event relationships well and cannot find out the dependence relationships between any two events which do not share enough similar content. As a result, the main picture of a "big event" discovered by previous works often is incomplete and several significant relationships are missing.

In this paper, to satisfy the user needs mentioned above in finding and understanding the whole picture of a complex event effectively and efficiently, we conduct an in-depth event relationship analysis for event search and propose a framework to search events based on user queries. The new characteristics of the proposed framework and the contributions of our work are as follows.

- In previous works, to discover dependence relationships between events, content similarity of events is measured by matching the keywords (terms) of events. However, there may be some keywords (in two events) which are actually related/ dependent but not identical. For example, "hospital" and "doctor" are dependent, but previous methods treat them as no relationship. To avoid this limitation, we adopt mutual information to measure the dependence between two terms (features), and then aggregate all mutual information between features in events to measure the content dependence degree between events. Such a process is named as *content dependence (CD) analysis* and the dependence relationship discovered based on dependence features of two events is named as *content dependence relationship*.
- As mentioned in paragraph 4, only content dependence analysis on events is inadequate to detect all event dependence relationships. According to the studies in Journalism [11] and our observation, it is not unusual for authors (reporters) to write

news articles on an event by referring to other events, when the authors consider there is a dependent relationship between them. For instance, some news articles about the event "Toyota recall due to safety problems from 2009 to 2010" refer to the event "NHTSA conduct investigations for Toyota recall". Motivated by this prevalent phenomena, We explore *event reference (ER) analysis* to detect whether there is an inter-event relationship specified by authors. The relationship between two events discovered by ER analysis is named as *event reference (ER) relationship*, which has not been explored by previous works.

- In contrast to previous works which only consider temporal relationship and content similarity, we adopt three kinds of event relationships (viz, temporal relationship, *CD* relationship obtained by *CD* analysis and *ER* relationship obtained by *ER* analysis) to identify the dependence relationship between two events. Note that *CD* and *ER* relationships are essentially event dependence relationships which are discovered by two different ways respectively. We name them by two different names with respect to the different ways for discovering them. *CD* relationships and *ER* relationships can be complementary to each other in identifying event dependence relationships.
- The search results are organized by a *temporal event map* (TEM) which constitutes the whole picture about an event's evolution or development along the timeline. Figure 1 shows an example TEM of the event "Toyota 2009-2010 vehicle recalls". A TEM provides a way to organize and represent the events search results by showing the interrelationships between/among the events. It provides an easier and more efficient means for users to know and understand their interested events in a comprehensive way.
- To evaluate the performance of our proposed approach, we conduct experiments on a real data set by comparing with a number of baseline methods. Experiment results show that our method outperforms baselines in discovering event dependence relationships.

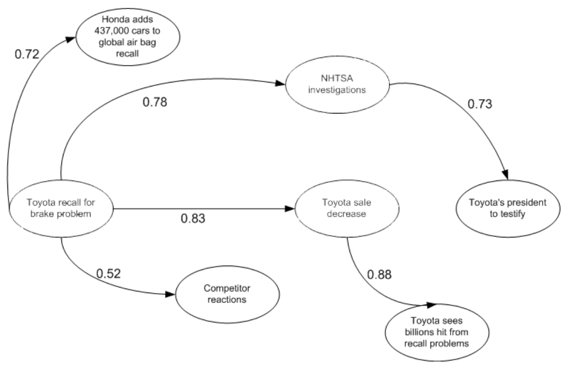

Fig. 1. An example of the *TEM* about the event "Toyota 2009-2010 vehicle recalls"

The rest of this paper is organized as follows. In section 2, we formulate the event search problem. Section 3 introduces the (temporal) event search framework. We conduct experiments on a real data set to evaluate the proposed methods for event search in section 4. In section 5, to further illustrate and evaluate our method, we study a query case about the event "SARS" happened in 2003. Related works are studied in section 6. We conclude the paper and introduce potential future works in section 7.

2 Problem Formulation

According to [9] and [17], an event is something that happens at some specific time and place. In reality, events often are reported by some documents, such as news articles in web pages. Formally, for an event a, there is a set of documents talking about a, and such a set of documents, denoted as $R_a = \{d_1^a, d_2^a, \cdots, d_n^a\}$, is named as **related document set** of a. Each document is about one event and an event can be reported by multiple documents. A document introducing an event includes the start time, place(s) and content of the event. Thus, for each document d_x^a, there should be a timestamp $\tau_{d_x^a}$, a set of place names $q_{d_x^a} = \{t_{x,1}, \cdots, t_{x,n}\}$ and a set of terms $h_{d_x^a} = \{f_{x,1}, \cdots, f_{x,n}\}$ about the event's content. We define an event as follows.

Definition 1. *An **event** a is a tuple (L_a, P_a, F_a) where L_a is the life cycle of a, P_a is the set of places where a happens, and F_a is the set of features describing a.*

The **life cycle** L_a of event a is the period (time interval) from the beginning time St_a to the end time Et_a of a, i.e., $L_a = [St_a, Et_a]$, where St_a is the earliest timestamp among all the timestamps of related documents of a, and Et_a is the latest timestamp among all that of related documents of a. The **place set** P_a of an event a is a set of terms denoted by P_a where $P_a = \{t_{a,1}, t_{a,2}, \cdots, t_{a,m}\}$ and each $t_{a,x}$ is a term which represents a place. For an event, it may consist of several component events, i.e., episodes.

For example, for the event of SARS epidemic which happened in 2002 among some 37 countries around the world, the life cycle of this event is from November 2002 to May 2006. The place of the event includes China, Canada, Singapore and so on. There are many reported news on the event on the Web. To describe the event, we can extract from the set of documents the set F_a of features (i.e., keywords), such as "SARS", "flu-like", "fever" and so on. The event of SARS epidemic consists of several component events such as "Experts find disease infect and SARS outbreaks", "China informs and cooperates with WHO", "SARS has great impact on economy" and so on.

Definition 2. *For each event a, it contains a set of **component events** denoted as $CE_a = \{a_1, a_2, \cdots, a_n\}$ where $1 \leq n$, a_x is a component event of a and $R_a = R_{a_1} \cup R_{a_2} \cup \cdots \cup R_{a_n}$.*

Definition 3. *Among all the component events of an event a, the **seminal component event** of a is the one whose start time is the same as that of a, i.e., the start time of the seminal component event is no later than those of the other component events of a.*

Definition 4. *Among all the component events of an event a, the **ending component event** of a is the one whose end time is as the same as that of a, i.e., the end time of the ending component event is no earlier than those of the other component events of a.*

For an event a and its component events, it is obvious that $St_a = Min_{i=1}^{n}(St_{a_i})$, $Et_a = Max_{i=1}^{n}(Et_{a_i})$, $P_a = P_{a_1} \cup P_{a_2} \cup \cdots \cup P_{a_n}$ and $F_a = R_{a_1} \cup F_{a_2} \cup \cdots \cup F_{a_n}$. We observe that there is a temporal requirement for two events to have a dependence relationship between them, as follows.

Observation 1. *If there is a dependence relationship from event a to event b, i.e., a is dependent on b, then there is a temporal relationship between a and b such that $St_b <= St_a$, i.e., b happens earlier than or at the same time as a.*

Definition 5. *A **Temporal Event Map** is a weighted directed graph, denoted by $TEM = (N, E, W_d)$, which consists of events as nodes, relations as edges, and weights on the edges as strength degrees of dependence relationships. In particular, each vertex $v \in N$ is an event, each edge $e_x \in E$ is a dependence relationship between two events, and $w_y \in W_d$ is a weight which indicates the strength degree of a dependence relationship.*

An example of temporal event map of the event "Toyota 2009-2010 vehicle recalls" is shown in Fig. 1.[1]

We formulate the problem of temporal event search as follows. The input of the search problem is a tuple (I_t, I_p, I_f) where I_t is a time interval, I_p is a set of terms of places, and I_f is a set of keywords about an event content. The event which is relevant to (corresponding to) the input is named as the **target event**, i.e., the event happens in the places in I_p during I_t, and the feature set of the target event contains I_f. The output of the search problem is a TEM constituting all the component events of the target event.

The problem of temporal event search can thus be regarded as a function ϕ:

$$\phi : I \times D \rightarrow T'$$

where I is the set of input, D is the set of documents and T' is the set of TEMs.

For the example of Fig. 1, we may have the following input:

$$I_t = [1/11/2009, 23/2/2010]; I_p = (USA); I_f = (Toyota, recall)$$

then the temporal event map for such a search task is the one shown in Figure 1.

3 Event Relationship Analysis

In this section, we propose a framework of event relationship analysis to support temporal event search. In our method, we first identify a set of related documents for the target event and extract component events from the related documents. We conduct content dependence (CD) and event reference (ER) relationship analysis to identify dependence between events.[2]

[1] We use the width of a line to indicate the strength of a dependence relationship.

[2] In the rest of the paper, we use the term "event" to denote "component event" for convenience wherever there is no ambiguity.

3.1 Preliminaries

A user query can be considered as search requirements corresponding to a target event which satisfies all the needs from the user. The related document set of the target event can be obtained by a function θ:

$$\theta : I \times D \rightarrow R$$

where I is the set of input, D is the set of documents and R is the set of related document sets.

In general, we consider (I_t, I_p, I_f) as three kinds of (not all are compulsory) user search requirements. In some cases, users may only input one or two of the (I_t, I_p, I_f). For such special cases, we only take the user input requirements into consideration, i.e., subset of (I_t, I_p, I_f).

For each target event a corresponding to an input I and its related document set R_a, we can detect several component events from R_a. All component events of a should happen during I_t, and their places are contained in I_p and features contained in I_f. The component event detection of a target event is a function φ:

$$\varphi : R \rightarrow E$$

where R is the set of related documents and E is the set of the component events.

For the problem of event detection, there have been many existing works published such as [1] [12] [14]. In this paper, we adopt the topic-model based method [14] as the preferred method to detect events.

3.2 Content Dependence Analysis

In analyzing content dependence (CD) relationships for temporal event search, we notice that features of an event a may have various degrees of importance in representing a. Some features are more representative than others for the event. An event can be represented by a *feature vector*, denoted by $\overrightarrow{F}_a$, which is a set of feature:value pairs.

$$\overrightarrow{F}_a = (f_{a,1} : v_{a,1}, f_{a,2} : v_{a,2}, \cdots, f_{a,n} : v_{a,n}), \forall i, 0 < v_{a,i} \leq 1$$

where $f_{a,i}$ is a feature and $v_{a,i}$ is the importance degree of $f_{a,i}$ for the event a. Hence, $v_{a,i}$ is the *NTF-IEF* (normalized term frequency-inverse event frequency) value of $f_{a,i}$, i.e.,

$$v_{a,i} = \frac{tf_{a,i}}{MAX_u(tf_{a,u})} \log \frac{N}{ef_i} \tag{1}$$

where $tf_{a,i}$ is the frequency of term i in R_a, N is the total number of component events, $MAX_u(tf_{a,u})$ is the maximal value among all $tf_{a,u}$ and ef_i is the number of component events containing term $f_{a,i}$.

As mentioned before, previous works use content similarity (most works adopt cosine similarity) to identify dependence relationships between events. However, two events may have some keywords which are dependent but not identical, which causes the previous works to be inadequate in measuring how relevant these two events are.

According to [10], variables (i.e., keywords) which are not statistically independent suggest the existence of some functional relation between them, and mutual information provides a general measure of dependencies between variables. Thus, we adopt mutual information to measure the dependence between features, and further use an aggregation of all mutual information between the feature sets in two events to measure the content dependence degree between them.

Formally, for two events a and b, the content dependence degree, denoted by $Cd(a, b)$, is an aggregation of all mutual information between all features in $\overrightarrow{F}_a$ and that in $\overrightarrow{F}_b$, as follows:

$$Cd(a,b) = \frac{\sum_{f_x \in \overrightarrow{F}_a} \sum_{f_y \in \overrightarrow{F}_b} I(f_x, f_y)}{|\overrightarrow{F}_a||\overrightarrow{F}_b|} \tag{2}$$

where $|\overrightarrow{F}_a|$ ($|\overrightarrow{F}_b|$) is the cardinality of the set $\overrightarrow{F}_a$ ($\overrightarrow{F}_b$), and $I(f_x, f_y)$ is the dependence degree between features f_x and f_y, measured as follows:

$$I(f_x, f_y) = P(f_x, f_y) \log \frac{P(f_x, f_y)}{P(f_x)P(f_y)} \tag{3}$$

where $P(f_x, f_y)$ is the probability of f_x and f_y co-occurring in the same document among all the related documents, and $P(f_x)$ is the probability of f_x occurring in a document among all documents, and $P(f_y)$ is the probability of f_y occurring in a document among all the documents.

By measuring all mutual information between two component events, we can obtain a **component content dependence matrix** of an event a, denoted as M_a^c, as follows:

$$M_a^c = \left\{ \begin{array}{c} Cd(1,1), Cd(1,2), \cdots, Cd(1,m) \\ \cdots \\ Cd(n,1), Cd(n,2), \cdots, Cd(n,m) \end{array} \right\}$$

where each entry is a content dependence degree between two component events.

3.3 Event Reference Analysis

Although content dependence measurement can address the limitation of content similarity measurement, it may still miss some dependence relationships between events. In particular, the existence of a dependence relationship between two events does not necessarily mean that there exists a content dependence relationship between them. In many cases, although the contents of two events are very different and even of different topics, people may still regard that there is a dependence relationship between them. For instance, “Experts find disease infect and SARS outbreaks” has an impact on “SARS has great impact on economy” and “SARS has a great impact on Tourism”. The latter two events are dependent on the first one even though their content dependence degree is indeed very small.

According to the studies of Journalism [11], when authors of news articles about an event a find and regard that there exists a dependence relationship between a and b (e.g., b triggers the happening of a, or a is evolved from b and so on), their articles may actually refer event b. This is in line with our observation on our collected data set. For instance, some news articles about the event "Toyota recall due to safety problems from 2009 to 2010" refer to the event "NHTSA conduct investigations for Toyota recall". Such an explicit reference relationship made by authors in their news articles reflect their viewpoints and consideration on the inter-event relationships [11]. Therefore, we may regard such event reference relationships as more meaningful and reliable than content dependence relationships, and ER relationship analysis provides a way to discover those event dependence relationships missed by CD analysis and obtain a more complete temporal event map (TEM).

We can also observe that when a news article of an event c refers to another event a, there are usually some phrases that identify event a in the documents of event c, and we name such phrases as *core features* of a. The definitions of core feature set of an event is defined below.

Definition 6. *The* ***core feature set*** *of an event a, denoted by F_a^c is a set of features which are salient in the event, distinguishable from those of other events, and jointly can identify the event.*

For two events a and b, if there exists a related document of b, denoted as d_x^b, such that $\exists f_i \in F_a^c, f_i \in d_x^b$, and $\tau(d_x^b) > St_a$ (i.e., a happens earlier than b), then we say there is a **reference relationship** from b to a, i.e., a **is a reference of** b or b **refers to** a. Such a reference relationship is a fuzzy relationship, and the more core features of a are mentioned in b, the more strength degree of the relationship. For example, for the event "US congressional hearings hold for Toyota recall" denoted by a and the event "Toyota's president to testify in US congressional hearings" denoted by b, we find that the core feature set of event a is $F_a^c = \{congress, hearing, safety\}$ while the core feature set of b is $F_b^c = \{Akio, Toyoda, testify, apologize\}$. For event a, some of its core features also exist in some documents (news articles) of b, (e.g., "congress", "hearing" and "safety" all appear in the news titled as "Toyota's president to testify before Congress" on Feb 19, 2010), so we say b refers to a.

The strength degree of a reference relationship from b to a is determined by a function $Cr(a, b)$ which is to be defined below. For event b referring to event a, it should follow the temporal restriction of Observation 1. For two events a and b, in order to find out whether b refers to a, we need to discover the core feature set of a first and then check whether the core features of a exist in the related documents of b.

According to our observation, the core feature set of an event has the following properties.

Property 1. The core features of an event a are the most salient and representative features of a, i.e., the features appear in the related documents of a with a high frequency.

Property 2. The core features of an event a are distinguishable from those of other events, i.e., the core features should facilitate us in identifying event a from all other events easily.

Based on the above properties, we propose the following function to select core features of an event a:

$$u(f_i, a) = p(f_i|a) \cdot p(a|f_i) \tag{4}$$

where $p(f_i|a)$ is the probability of feature f_i to exist in the related documents of event a, and $p(a|f_i)$ is the probability of a document (in which f_i is a feature) being on event a. Note that $p(f_i|a)$ and $p(a|f_i)$ reflect the properties 1 and 2 respectively.

We select top-*k* core features based on equation 4. For two events a and b, the more related documents of b refer to more core features of a, the stronger is the reference relation from b to a. We propose a function to measure the strength degree of a reference relation from a to b as follows:

$$Cr(a,b) = \frac{\sum_{i=0}^{N_b} M_{b,i}^a}{|F_a^c|} \times \frac{1}{N_b}, \forall M_{b,i}^a > 1 \tag{5}$$

where N_b is the number of related documents of b, $M_{b,i}^a$ is the number of core features of a existing in the document d_i^b, $|F_a^c|$ is the cardinality of F_a^c. Note that there is a restriction for $M_{b,i}^a$ in $Cr_(a,b)$ where $M_{b,i}^a > 1$, highlighting that a reference relationship from b to a should refer more than one core feature of a.

For the reason that the values of $Cr(a,b)$ and $Cr(b,a)$ may be greater than zero, a could refer to b and also b could refer to a, and the strength degrees of reference relationship from a to b could be different with that from b to a. An event can be refereed by many other events. Besides, one event can also refer to many other events.

By measuring all component event reference degree between any two events, we can obtain an **component event reference matrix** of a target event a, denoted by M_a^r, as follows:

$$M_a^r = \left\{ \begin{array}{c} Cr(1,1), Cr(1,2), \cdots, Cr(1,m) \\ \cdots \\ Cr(n,1), Cr(n,2), \cdots, Cr(n,m) \end{array} \right\}$$

Each entry is a reference degree between two events.

3.4 Temporal Event Map Construction

We adopt content dependence (CD) analysis and event reference (ER) analysis to identify event dependence relationships. In cases when users are only interested in the ER relationships between events, we can do a projection on the TEM and obtain an *event reference TEM*, which is a sub-graph of the entire TEM. Similarly, if users are only interested in the CD relationships between events, we also can do a projection on the TEM and obtain a *content dependence TEM*. Besides, it is easy to show all the CD, ER and event dependence relationships in a TEM. While there are many interesting issues related to the visualization of TEM, we omit further discussion here since are our focus in this paper is on event relationship analysis.

4 Evaluation

In this section, we conduct experiments on a real data set to evaluate our approach by comparing it with a number of baseline methods.

4.1 Experiment Setting

To evaluate our method for temporal event search, we collect 5063 English news articles (i.e., web page documents of news) from some mainstream news websites such as CNN News and BBC News. We select ten queries about major events to test our method, such as "Toyota 2009-2010 vehicle recalls", "2010 Copiap mining accident", "SARS in 2003" and so on. Among these, the event "SARS in 2003" contains the most number of related news articles (i.e., 231 articles), and the event "Christchurch Earthquake in 2010 in New Zealand" contains the fewest number of related news articles (i.e., 39 articles). According to our observation on the data set, when an event a refers to another event b, the number of referred core features of b is often around five.[3] Thus, we select top-5 core features to measure event reference relationships in our experiment.

To compare with our method, we adopt three baseline methods. The first one is the state-of-the-art method of discovering event evolution relationship proposed by Yang [17], which is similar to the method in [2] and we denote it as EEG. The second baseline, denoted as CDM, only considers content dependence analysis and does not use event reference analysis to judge event dependence relationships. Different from CDM, the third one, denoted as ERM, only considers event reference analysis instead of using content dependence analysis to judge event dependence relationships.

We have invited five human subjects to annotate the dependence relationships between events. All the annotated relationships are combined synthetically to obtain a set of relationships, i.e., the union of all the relationships annotated. Such a set of relationships given by the annotators is considered as a standard answer set (ground truth) of event dependence relationships. For the reason that different people may have different viewpoints on the event relationships due to, e.g., their knowledge and background, not every annotator came up with the same set of dependence relationships. Therefore, the standard answer set is an aggregation of the annotations given by all the annotators.

For the evaluation we use $Precision$, $Recall$ and $F - measure$ as the metrics. We denote the set of event dependent relationships (i.e., edges in a TEM) annotated by annotators as R_A, and the set of event dependence relationships discovered by machine as R_M. The metrics are defined as follows:

$$Precision = \frac{R_A \cap R_M}{R_M}; Recall = \frac{R_A \cap R_M}{R_A}$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

4.2 Experiment Results

In constructing TEM , there is a parameter α which is used to prune the "weak" event dependency relationships. So first, we test different values of α to evaluate the effect of α on *Precision*, *Recall* and *F-measure* for setting the best value of parameter α for the following experiment. In our testing, we use two query events, one is "SARS in 2003"

[3] Such an observation is only based on our collected data set. It could be different for other data sets.

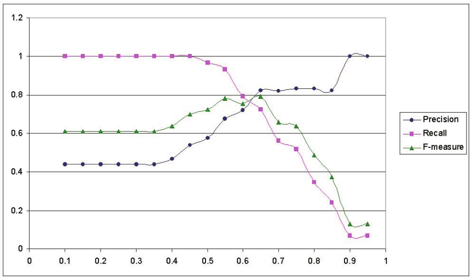

Fig. 2. The effect of α on Precision, Recall and F-measure

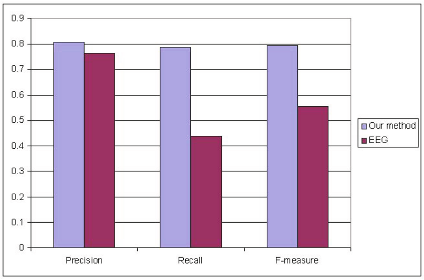

Fig. 3. Our method vs. EEG

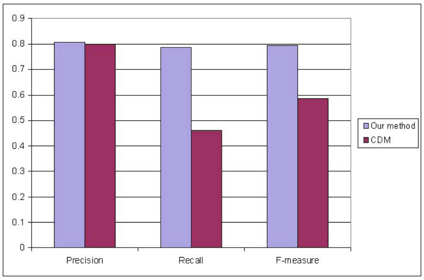

Fig. 4. Our method vs. CDM

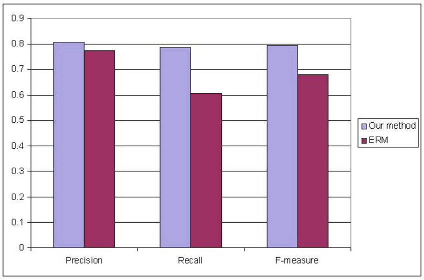

Fig. 5. Our method vs. ERM

which contains the most number of related news articles and the other is "Christchurch Earthquake in 2010 in New Zealand" which contains the fewest number of related news articles. Figure 2 shows the effect of α on *Precision*, *Recall* and *F-measure*. According to Fig. 2, we find that as α increases, the *Precision* and *F-measure* increase while *Recall* decreases. The reason is that when the value α is small, there are many event dependence relationships whose dependence degree is great than α (but the dependence relationship is still actually "weak"), so the $Recall$ is high and the $Precision$ is low. As α increases, more and more event dependence relationships of which dependence degree is lower than α are pruned, so the $Recall$ becomes lower and the $Precision$ becomes higher. When $\alpha = 0.65$, we obtain the highest value of *F-measure*. Thus, we set $\alpha = 0.65$ for all the test queries subsequently.

After setting the value of α, we conduct all test queries and average the results of them on different metrics. Figures 3-5 show the comparison of our method with all the three baseline methods on $Precision$, $Recall$ and $F - measure$. According to Figures 3-5, it is obvious that our method outperforms all the baseline methods on $Precision$, $Recall$ and $F - measure$. The $Precision$ and $Recall$ values of our method are around 0.8, meaning that not only most event dependence relationships discovered by our method are correct, but also our method can discover more event dependence relationships than the baselines. Our method's $F - measure$ score is also around 0.8 since it is a combination of $Precision$ and $Recall$. Note that CDM outperforms EEG a little on all the metrics, indicating that using mutual information to measure feature

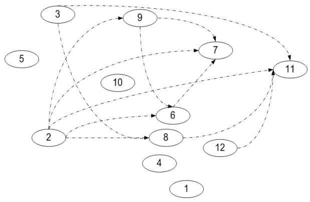

Fig. 6. The Result of EEG

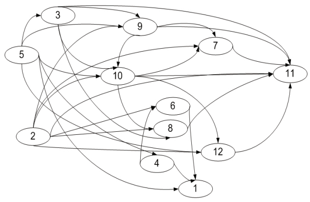

Fig. 7. The Result of Annotators

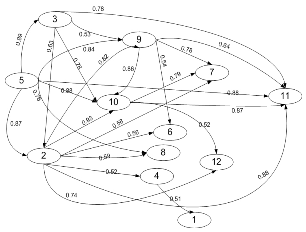

Fig. 8. The Result of Our Method

dependence is better than just matching keyword similarity (as done by previous works). ERM outperforms EEG and CDM on all three metrics. It indicates that using event reference analysis (i.e., ERM) to identify event dependence relationships is more effective than using content dependence relationship analysis (i.e., CDM) and content similarity analysis (i.e., EEG). Besides, it is quite interesting to see that the $Recall$ of CDM is greater than that of ERM, while both of them are smaller than that of our method. This means that the event dependence relationships identified by CDM and ERM are indeed different and complementary, and our method being a combination of CDM and ERM has the strength of both methods'. In other words, taking both content dependence and event reference analysis into consideration in identifying event dependence relationships can perform better than taking just one of these.

4.3 Case Study

To illustrate the performance of our proposed method more clearly, we further show a specific search case on the query event "SARS happened from 1/3/2003 to 30/6/2003 around the world" denoted by Q_{SARS}. The test query is $I_p = (China)$,$I_t = [1/3/2003, 30/6/2003]$, $I_f = (SARS)$.

Table 1. Component Events for the query about SARS event from 1/3/2003 to 30/6/2003

Component Event	Summary
1	SARS has great impact on Tourism
2	SARS cases are reported and updated regularly to reflect the disease seriousness
3	Experts treat patients with medicine in hospital
4	SARS has great impact on transportation especially airline
5	Experts find disease infect and SARS outbreaks
6	SARS has great impact on economy
7	Other countries donate and offer help for China for SARS
8	Scientists' find coronavirus and conduct animal test for vaccine
9	China informs and cooperates with WHO on fighting SARS
10	China makes effort on prevent disease spread
11	Beijing has made SARS under control
12	Quarantine probable cases and close schools for disinfecting

Table 2. Comparison on discovered event relationships of our method and EEG for Q_{SARS}

	Correct	Missed	Incorrect	New	Total
Our method	20	7	2	2	24
EEG	10	17	2	0	12

Table 1 shows all the component events which are related to Q_{SARS}. Table 2 shows the statistics of the discovered event relationships by our method and EEG based on the results of human annotators for this case. Figures 6-8 show the relationship graph (or TEM) obtained by EEG method, given by the human annotators and our method for Q_{SARS}, respectively. Our method can find more and miss less correct inter-event relationships. In addition, our method can discover not only the inter-event relationships but also the strength degrees of such relationships. More interestingly, our method can find some new relationships which were not found by EEG and even human annotators. Such new relationships are confirmed and approved by the annotators as meaningful ones (e.g., the relationship from event 5 to event 2 and the relationship from event 3 to event 2).

5 Background and Related Works

There are many works about processing events which may include news event or system events, although most of these work focus on news event.

To the best of our knowledge, there is no work on temporal event search before. A related work is done by Jin et al. [6] who present a temporal search engine supporting temporal content retrieval for Web pages called TISE. Their work supports Web pages search with temporal information embedded in Web pages, and the search relies on a unified temporal ontology of Web pages. TISE handles Web pages search only, and it cannot handle event search nor discover the event relationships.

Topic detecting and tracking (TDT) is a hot research topic related to our work. Given a stream of constantly generated new documents, TDT groups documents of the same topic together and tracks the topic to find all subsequent documents. There are several techniques on detecting news topics and tracking news articles for a new topic. For instance, Allan et al. [1] define temporal summaries of news stories and propose methods for constructing temporal summaries. Smith [12] explores detecting and browsing events from unstructured text. Some techniques are proposed to detect particular kinds of events. For example, Fisichella et al. [7] propose a game-changing approach to detect public health events in an unsupervised manner. Modeling and discovering relationships between events as generally out of the scope of current TDT research.

Mei and Zhai [8] study a particular task of discovering and summarizing the evolutionary patterns of themes in a text stream. A theme in an interval may be part of an event or a combination of several events that occur in the interval. Their work does not however capture the interrelationships of major events. Fung et al. [4] propose an algorithm named Time Driven Documents-partition to construct an event hierarchy in a text corpus based on a user query.

Some other works focus on discovering stories from documents and representing the content of stories by graphs. For example, Subasic et al. [13] investigate the problem of discovering stories. Ishii et al. [5] classify extracted sentences to define some simple language patterns in Japanese so as to extract causal relations, but their work cannot handle cases which are not defined in their patterns.

An event evolution pattern discovery technique is proposed by Yang et al. in [16]. It identifies event episodes together with their temporal relationships. They consider temporal relationships instead of evolution relationships. Although the temporal relationships can help organize event episodes in sequences according to their temporal order, they do not necessarily reflect evolution paths between events. An extended work of them occurs in [15]. Yang et al. [17] define the event evolution relationships between events and propose a way to measure the event evolution relationships. In their work, identifying an event evolution relationship between two events depends on the similarity of the features of the two events. Based on a small number of documents and events in a news topic, Nallapati et al. [9] define the concept of event threading. Their definition of event threading is a content similarity relationship from previous event to a later event. The event threading is organized as a tree structure rather than a graph. In order to identify event threading, they employ a simple similarity measure between documents to cluster documents into events and the average document similarity to estimate the content dependencies between events. Feng and Allan [2] extend Nallapati's work to passage threading by breaking each news story into finer granules, and propose a model called incident threading in [3].

6 Conclusions and Future Works

In this paper, we have defined three kinds of event relationships which are temporal relationship, content dependence relationship and event reference relationship, and have applied them to measure the degree of inter-dependencies between component events to support temporal event search. We have also formalized the problem of event search and proposed a framework to search events according to user queries. Experiments on a real data set show that our proposed method outperforms the baseline methods, and it can discover some new relationships missed by previous methods and sometimes even human annotators.

Admittedly, several possible future extensions can be made to our work. In our current method, only top-5 core features are selected for event reference analysis over the collected data set. How to choose the "right" number of core features for event reference analysis automatically for different data sets is an open issue for further study. Another potential extension is to implement a visualization tool with a sophisticated user interface based on our current method.

Acknowledgements. This work is supported by Foundation for Distinguished Young Talents in Higher Education of Guangdong, China (NO. LYM11019); the Guangdong Natural Science Foundation, China (NO. S2011040002222); the Fundamental Research Funds for the Central Universities,SCUT (NO. 2012ZM0077); and Guangdong Province University Innovation Research and Training Program (1056112107).

References

1. Allan, J., Gupta, R., Khandelwal, V.: Temporal summaries of new topics. In: SIGIR 2001: Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 10–18. ACM, New York (2001)
2. Feng, A., Allan, J.: Finding and linking incidents in news. In: CIKM 2007: Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management, pp. 821–830. ACM, New York (2007)
3. Feng, A., Allan, J.: Incident threading for news passages. In: CIKM 2009: Proceeding of the 18th ACM Conference on Information and Knowledge Management, pp. 1307–1316. ACM, New York (2009)
4. Fung, G.P.C., Yu, J.X., Liu, H., Yu, P.S.: Time-dependent event hierarchy construction. In: Proceedings of KDD 2007, pp. 300–309. ACM, New York (2007)
5. Ishii, H., Ma, Q., Yoshikawa, M.: Causal network construction to support understanding of news. In: Proceedings of HICSS 2010, pp. 1–10. IEEE Computer Society, Washington, DC (2010)
6. Jin, P., Lian, J., Zhao, X., Wan, S.: Tise: A temporal search engine for web contents. In: IITA 2008: Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application, pp. 220–224. IEEE Computer Society, Washington, DC (2008)
7. Fisichella, K.D.M., Stewart, A., Nejdl, W.: Unsupervised public health event detection for epidemic intelligence. In: CIKM 2010: Proceeding of the 19th ACM Conference on Information and Knowledge Management (2010)
8. Mei, Q., Liu, C., Su, H., Zhai, C.: A probabilistic approach to spatiotemporal theme pattern mining on weblogs. In: WWW 2006: Proceedings of the 15th International Conference on World Wide Web, pp. 533–542. ACM, New York (2006)
9. Nallapati, R., Feng, A., Peng, F., Allan, J.: Event threading within news topics. In: CIKM 2004: Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management, pp. 446–453. ACM, New York (2004)
10. Steuer, J.S.R., Daub, C.O., Kurths, J.: Measuring distances between variables by mutual information. In: Proceedings of the 27th Annual Conference of the Gesellschaft für Klassifikation: Cottbus, pp. 81–90. Springer (2003)
11. Rich, C.: Writing and Reporting News: A Coaching Method, 6th edn. Wadsworth Publishing Company (2009)
12. Smith, D.A.: Detecting and browsing events in unstructured text. In: SIGIR 2002: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 73–80. ACM, New York (2002)
13. Subasic, I., Berendt, B.: Web mining for understanding stories through graph visualisation. In: ICDM 2008: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, pp. 570–579. IEEE Computer Society, Washington, DC (2008)
14. Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., Su, Z.: Arnetminer: extraction and mining of academic social networks. In: KDD, pp. 990–998 (2008)
15. Yang, C.C., Shi, X.: Discovering event evolution graphs from newswires. In: WWW 2006: Proceedings of the 15th International Conference on World Wide Web, pp. 945–946. ACM, New York (2006)
16. Yang, C.C., Shi, X.-D., Wei, C.-P.: Tracing the Event Evolution of Terror Attacks from On-Line News. In: Mehrotra, S., Zeng, D.D., Chen, H., Thuraisingham, B., Wang, F.-Y. (eds.) ISI 2006. LNCS, vol. 3975, pp. 343–354. Springer, Heidelberg (2006)
17. Yang, C.C., Shi, X., Wei, C.-P.: Discovering event evolution graphs from news corpora. Trans. Sys. Man Cyber. Part A 39(4), 850–863 (2009)