Information Processing and Management 57 (2020) 102319

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Extracting temporal and causal relations based on event networks

Duc-Thuan Vo[a], Feras Al-Obeidat[b], Ebrahim Bagheri[a,*]

[a] Laboratory for Systems, Software and Semantics (LS^3), Ryerson University, Toronto, Canada
[b] Zayed University, United Arab Emirates

ARTICLE INFO

Keywords:
Event extraction
Open information extraction
Event network
Temporal event
Causal event

ABSTRACT

Event relations specify how different event flows expressed within the context of a textual passage relate to each other in terms of temporal and causal sequences. There have already been impactful work in the area of temporal and causal event relation extraction; however, the *challenge* with these approaches is that (1) they are mostly *supervised* methods and (2) they rely on syntactic and grammatical structure patterns at the *sentence-level*. In this paper, we address these challenges by proposing an *unsupervised* event network representation for temporal and causal relation extraction that operates at the *document level*. More specifically, we benefit from existing Open IE systems to generate a set of triple relations that are then used to build an event network. The event network is bootstrapped by labeling the temporal disposition of events that are directly linked to each other. We then systematically traverse the event network to identify the temporal and causal relations between indirectly connected events. We perform experiments based on the widely adopted TempEval-3 and Causal-TimeBank corpora and compare our work with several strong baselines. We show that our method improves performance compared to several strong methods.

1. Introduction

Learning temporal and causal relationships between events that have been mentioned in a text is an important task in information extraction towards deeper language understanding (Qian, Deng, Ye, Ma & Yuan, 2019; (Zhang and Lv, 2018); (Navarro-Colorado and Saquete, 2016, (Zheng and Suri, 2019, (Ye and Luo, 2019, Fan, Fan, Smith & Garner, 2019). Temporal and causal relations can happen both between events within the same sentence and between events across sentences in a document. Understanding relations between events in a document is beneficial to various Natural Language Processing applications such as question answering (Abacha et al., 2016; Qian et al., 2019), document summarization (Ji et al., 2013; Verhagen, Sauri, Caselli & Pustejovsky, 2010; (Kim and Kim, 2018) and textual entailment (UzZaman et al., 2013; Verhagen et al., 2010). Moreover, identifying the causal and temporal relation between events is an important step in predicting occurrence of future events, and can be beneficial in risk analysis as well as proactive decision making (Hogenboom, Frasincar, Kaymak, Jong, & Caron, 2016).

An increasing amount of recent work has focused on recognizing temporal and causal event relations within a document, but mostly limited to identifying intra sentences causal relations with explicit causal indicators. (Mirza & Tonelli, (2014a) and Laokulrat, Miwa and Tsuruoka (2015) have reported that incorporating temporal information can improve the performance of a causal relation classifier. These authors have developed and annotated a causal corpus (Causal-TempBank) based on the TempEval-3 corpus (Mirza & Tonelli, 2014a) for the task of causal event relation. Mirza and Tonelli (2016a) built both a rule-based multi-sieve approach and a feature-based classifier to recognize causal relations in Causal-TimeBank. These approaches to temporal and causal relation

* Corresponding author.
E-mail addresses: thuanvd@ryerson.ca (D.-T. Vo), Feras.al-obeidat@zu.ac.ae (F. Al-Obeidat), bagheri@ryerson.ca (E. Bagheri).

https://doi.org/10.1016/j.ipm.2020.102319
Received 29 October 2019; Received in revised form 18 April 2020; Accepted 31 May 2020
Available online 20 June 2020
0306-4573/

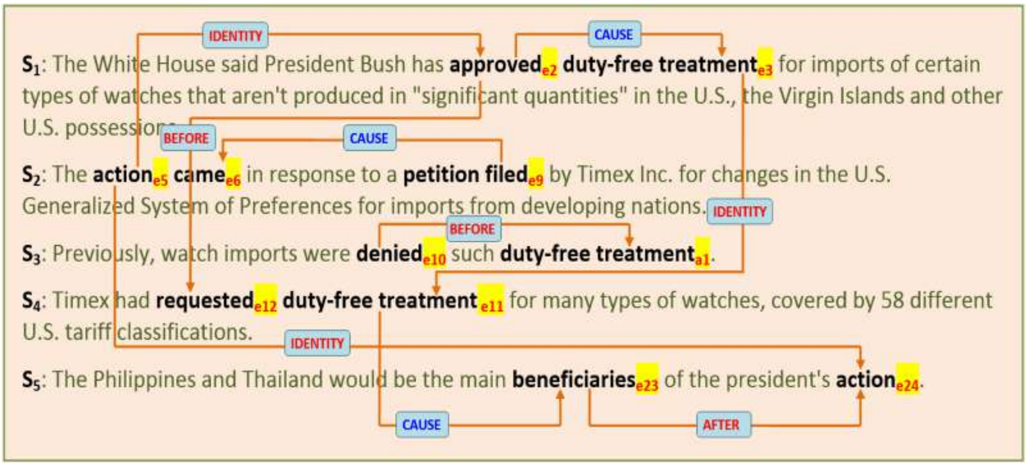


Fig. 1. Examples of direct event relations and indirect event relations in a textual document.

classification use machine-learning-based classifiers (Laokulrat et al., 2015; Laokulrat, Miwa, Tsuruoka & Chikayama, 2013; (Mirza & Tonelli, (2014b)); (Mirza and Tonelli, (2016a))) that are trained based on a predefined, finite and fixed schema of relation types. The common strategy of these techniques is to generate linguistic features based on syntactic, dependency, or shallow semantic structures of the text. Based on these features, supervised learning methods are used to identify pairs of events that are related to each other and can be classified into pre-defined relation types. However, the state-of-the-art approaches (Laokulrat et al., 2015; (Mirza and Tonelli, 2016a) suffer from two key drawbacks. First, they are focused on a limited subset of features, which might not, in many cases, be present in every sentence or be sparsely available. Second, training on linguistic structures such as the output of syntactic and dependency parsers does not necessarily identify all possible types of event relations when they are presented in different sentences or different documents.

One of the challenges for obtaining comprehensive temporal and causal relations is that such event relations are possible in pair of events from sentences that are not similar in a document. The challenge is especially true for identifying cross-sentence temporal and causal event relations and they have no clear indicators in their relationships. For instance, considering the three sentences shown Fig. 1, the events in both $\langle e_2\text{-}e_3\rangle$, $\langle e_5\text{-}e_8\rangle$ are related to each other by the "BEFORE" temporal relation type. Here, while e_2 and e_3 are present in the same sentence, events e_5 and e_8 are in different sentences. As such, sentences that rely on features based on grammatical parsers can fail to identify correct relation types. We observe that events occur in action flows that happen throughout the documents. Therefore, for improving event relation identification, we consider an event network representation that involves the connections between event flows of a document and can capture cross-sentence event relations in a document. We address these challenges by adopting an Open Information Extraction (Open IE) approach (Corro et al., 2013; Vo & Bagheri, 2018), which is able to extract relations and their arguments without the need to restrict the search to predefined relation types, in order to build the event network for event relation extraction. We consider and incorporate all identified Open IE patterns that consist of at least one event instance in the event network, which is then systematically traversed for identifying temporal relations. As an example in Fig. 1, both $\langle e_2\text{-}e_3\rangle$ and $\langle e_5\text{-}e_3\rangle$ relations can be extracted from two Open IE patterns, namely ("*President Bush*", "*has approved*", "*duty-free treatment for imports*") and ("*Timex*", "*had requested*", "*duty-free treatment*"), respectively. The event network is built based on such patterns identified and extracted by Open IE where each triple will be represented as two vertices connected by an edge. The traversal of this event network allow us to identify indirect relations through event coreference relations in the network. Furthermore, we employ a shortest path strategy to determine the event flow between two events based on the constructed event network.

1.1. Research objective and contributions

The main objective of this paper is to address the main two limitations of earlier work that rely on supervised methods at the sentence level for extracting event relations. To this end, our work is completely unsupervised and operates at the document level given its exploitation of Open IE techniques. We propose an event network representation that will be used for systematically extracting causal and temporal event relations. Our most significant contributions can be enumerated as follows:

- We propose an event network structure primarily based on information from Open IE systems for extracting temporal and causal relations between events. The event network is the basis for systematically exploring the event flow between events by considering how events can be reached from one another.
- We present a method to measure the weight of event flows to detect shortest paths between events. Through the weight, shortest paths are identified that can ultimately lead to higher confidence event relation extractions.
- We present algorithms for temporal and causal relation extraction that make it possible to reason over the set of relations observed in the event flow to make a determination about the relation between two events. Indirect relations are inferred through transitive

inferences over direct relations.

- Our approach works at the document level and hence can identify event relations between events that have been expressed across different sentences. This is made possible due to the linking of different events in the proposed event network whose linking transcends individual sentences and forms a representation of events in the document.

The rest of this paper is organized as follows: literature related to the concept of event relation extraction is presented in Section 2. Section 3 offers detailed description of our proposed approach to construct an event network for event extraction. Following that, we present the methods used for causal and temporal relation extraction. Section 4 offers an in-depth analysis of our experiments for temporal and causal event extraction where the results obtained from our proposed approach are compared to strong baseline methods. In the last section, we draw conclusions about the merits of our work and offer ways to advance our work in the future.

2. Related work

In the recent years, several approaches for temporal and causal relation extraction have been proposed that include those that employ linguistic patterns (Girju et al., 2009; Hashimoto et al., 2014; Laylavi, Rajabifard & Kalantari, 2017), statistical measures (Beamer & Girju, 2009; Do, Chang & Roth, 2011; Hashimoto et al., 2015; Hu & Walker, 2017; Hu, Rahimtoroghi, & Walker, 2017; Ozdikis, Oğuztüzün & Karagoz, 2016), lexical semantics of events (Riaz & Girju, 2013, Hashimoto et al., 2014) and supervised classifiers (Mirza & Tonelli, 2014b; Mirza and Tonelli, 2016a; Fei, Ren & Ji, 2019; Zhao, Jin & Wang, 2018; Li and Mao, 2019), primarily with the goal of identifying event knowledge from a text corpus. In this study, we propose to build an event network based on Open IE in order to extract event relations on both (1) events that are directly related to each other, and (2) those that can be indirectly linked to each other through links in the event network. In this section, we present several studies that are related to our work.

2.1. Event extraction

Grishman, Huttunen and Yangarber (2002), Tanev, Piskorski and Atkinson (2008) and Fukui, Okada, Satoh and Numao (2019) present pattern-based approaches for event extraction that identify events in free text such as news, emails and health reports and derive detailed information about who did what to whom, where and when. Grishman et al. (2002) also built a complete system that automatically extracts and updates an event database of infectious disease outbreaks. The system has an extraction engine that is based on text zones including headline, date, text body to extract events using predefined event patterns. Tanev et al. (2008) identified syntactic patterns to extract events on violent and natural disaster news. Their method used several components in a pipeline to learn patterns through a semiautomatic pattern acquisition technique consisting of a small corpus that was manually annotated with event information. Regarding supervised learning, several approaches (Abebe, Tekli, Getahun, Chbeir & Tekli, 2019; Ahn, 2006; Bethard & Martin, 2006; Fan et al., 2019; Hardy, Kanchakouskaya & Strzalkowski, 2006; Ji & Grishman, 2008; Ye and Luo, 2019, Zhang, Boons & Batista-Navarro, 2019) train one or more classifiers that can extract events from textual documents. Ahn (2006) proposed a series of classification sub-tasks such as event anchor identifier, argument identifier, attribute assigner and event core reference module for event extraction. Bethard and Martin (2006))) designed an event extraction method using sequence tagging with the BIO (Begin-Inside-Outside) schema. The schema is used to trained features based on Support Vector Machines (SVM) over the TimeBank corpus (Pustejovsky et al., 2003). Furthermore, Ji and Grishman, (2008) and Wang, Zhang and Chang (2017) train event relation classifiers based on the words and phrases of relation arguments that evoke events in the text. Li, Ji, Deng and Han (2011) used topic modeling to improve the process of event extraction. Their method operates by recognizing topic clusters for event extraction. (Ye and Luo, (2019) proposed a general ranking based multi-label learning framework combined with convolutional neural networks to seek latent connection between relation types for event extraction. Zhang et al. (2019) presented a hybrid approach by combining semantic role labeling frames and lexicon of event nouns and Levin's verb classes for event extraction over 20 event types related to socioeconomic phenomena.

2.2. Temporal event extraction

Temporal event relations specify how different events expressed within the context of a textual passage relate to each other in terms of time sequence. Allen (1983) first introduced a transitivity table of temporal intervals for reasoning about events, e.g., the table determines if event e_1 happens before event e_2 and e_2 happens before e_3 that e_1 happens before e_3. Supervised methods (Chambers et al., 2008; Mani, Wellner, Verhagen & Pustejovsky, 2007) have focused on formulating temporal features ordering around event pairs to formulate a classification task. Mani et al. (2007) as well as Tatu and Srikanth (2008) have curated training data by exploiting external temporal resources to improve supervised learning for temporal relation extraction. (Chambers and Jurafsky, (2008a), UzZaman and Allen (2010) take advantage of external information to enhance temporal logic transitivity constraints for pariwise classification. These methods face limitations related to the sparseness of event relations in a corpus (Chambers & Jurafsky, 2008).

Furthermore, Amigo, Artiles, Li and Ji (2011) and Denis and Muller (2011) presented an inference method over time graph to address limitations of constraint space to improve the identification of temporal relations. The time graph is built for representing temporal entities with edges as relations to predict the relationship between entity nodes. Moreover, UTTime (Laokulrat et al., 2013) and NavyTime (Chambers, 2013) exploit data-driven analysis using syntactic information and lexical semantic information for classifying temporal relation types. These systems only exploited features at the sentence-level such as predicate-argument structure of relations in a sentence. (Mirza & Tonelli, (2014b) and (Mirza and Tonelli, (2016a) have proposed hybrid classification models (Chambers, Cassidy, McDowell & Bethard, 2014; D'D'Souza & Ng, 2013; (Navarro-Colorado and Saquete, (2016)) for temporal event classification. They integrate rule-based and data-driven classifiers in a sieve-based architecture for temporal ordering. The classifiers are ordered by their individual precision. After each classifier proposes a label, the architecture infers transitive links from the new labels, adds them to the temporal label graph and informs the next classifier about this decision. (Hoang and Mothe, (2018)) have presented a method to predict a location and time in a tweet by enhancing tweet representation with prepositions before a proper noun. Regarding deep learning methods, Han et al. (2019) have proposed a deep structured learning framework to learn scoring functions for pair-wise relations and employed a structured support vector machine for temporal relation extraction. (Zhang et al., 2019b) have introduced a method to incorporate word and character level information by using features extracted from a deep neural model for event relation extraction. Junuthula, Haghdan, Xu and Devabhaktuni (2019) present a block point process model (BPPM) for continuous-time event-based dynamic networks to analyze temporal relations between events in on-line social networks.

2.3. Causal event extraction

Recognizing causal relations between any two events is as challenging as detecting their temporal relation. In this context, (Bethard et al., 2007) have collected conjoined event pairs from the Wall Street Journal corpus which described cause-effect relations for causal classification. Rink, Bejan and Harabagiu (2010) improved Bethard et al.'s work by identifying relation types as a feature and proposed textual graph classification for causal event relation classification. Do et al. (2011) and Peng, Song and Roth (2016) have presented a method to measure causality on event triggers, e.g., relations on verb-verb, verb-noun and noun-noun between two events by using pointwise mutual information. They also used discourse information from trained corpus, e.g., connective types on Penn Discourse TreeBank (Do et al., 2011), as discourse connectives towards identifying causality between events. (Ittoo and Gosse Bouma, (2011)) proposed causal patterns based on syntactic structure analysis in order to extract explicit and implicit causal relations. On other hand, Riaz and Girju (2013) extracted causal relations based on analyzing verbs between events. Recently, (Mirza & Tonelli, (2014a)) and (Mirza, (2016b)) presented a data-driven approach with rules to extract (explicit) causal relations between events from a text. The rule-based system relies on an algorithm that, given a term *w* belonging to affect, link, causative verbs or causal signals, looks for specific dependency where term *w* is connected to the two observed events. If such dependencies are found, a causal link (CLINK) is automatically set between the two events. Negi, Pavuri, Patel and Jain (2019))) have trained syntactic and dependency features in order to classify the causal relationship between drugs and medical conditions such as suspect drug or non-suspect drug. (Li and Mao, (2019)) have used a Knowledge-oriented Convolutional Neural Network (K-CNN) model that incorporates human knowledge to capture the causal relationship, and a causal-oriented direction in order to learn cause-effect features for causal relation extraction. Fei et al. (2019) have proposed a method to use a recursive neural network that automatically learns syntactic features from dependency trees in order to represent global dependencies for biomedical event detection.

2.4. Open information extraction

Existing Open IE systems (Corro et al., 2013; Mausam, Schmitz, Bart & Soderland, 2012; Etzioni et al., 2012; Vo and Bagheri, 2017) are able to automatically extract a set of relatively stable relations from a textual corpus based on some heuristic patterns such as through the analysis of the grammatical structure of sentences using dependency or syntactic parsing or part of speech tagging. In this paper, we exploit triple relations extracted by Open IE in order to construct an event network without the need to restrict the search to pre-specified semantic relations. We will cover a broader literature that is immediately related to Open IE system in this section.

The emergence of a pioneering Open IE system such as TextRunner following the seminal work of Banko, Cafarella, Soderland, Broadhead and Etzioni (2007) brought a myriad of techniques to the fore in recent years with the stated aim of improving upon the original TextRunner. Currently, the majority of Open IE systems use a shallow syntactic representation (Banko et al., 2007; Wu and Weld, 2010; Fader, Soderland & Etzioni, 2011) or dependency parsing (Wu and Weld, 2010; Mausam et al., 2012; Corro et al., 2013) in the form of verbs or verbal phrases and their arguments. TextRunner (Banko et al., 2007) and ReVerb (Fader et al., 2011) use automatically generated training data and syntactic analysis while WOE^{pos} (Wu and Weld, 2010) trains the corpus automatically by procuring infoboxes from Wikipedia. TextRunner trains a Bayes classifier in an offline phase and consequently applies it for the efficient extraction of propositions in the online phase. To address the problem of incoherent and uninformative extractions, Fader et al. (2011) developed ReVerb, which effectively applies syntactical and lexical constraints. In the unique approach adopted by ReVerb, verbal relation sequences are initially extracted based on a set of POS patterns as opposed to extracting entities. The entities are subsequently identified around the relation sequences, sanctioning the system to extract all potential relation tokens between two entities. Using dependency parsing, $\text{WOE}^{\text{parse}}$ (Wu and Weld, 2010) expands on this and uses automatically generated training data to learn extraction patterns on dependency parsing. Mausam, Schmitz et al. (2012) present OLLIE based on various heuristics to obtain propositions from dependency parsers. Additionally, hand-labeled data is used to create a training set consisting of millions of relations extracted by ReVerb (Fader et al., 2011). OLLIE learns relation patterns from the dependency path and lexicon information engendering a space where relations that match the identified patterns can be extracted. On other hand, OIE approaches (Mirrezaei, Martins & Cruz, 2015; Gamallo, Garcia, & Fernandez-Lanza, 2012; Schmidek & Barbosa, 2014) generally tend to exploit the information available on Wikipedia, DBpedia or Freebase for Open IE. (Mirrezaei et al., (2015)) address the failure of systems like ReVerb or OLLIE to extract noun mediated relations by proposing a promising approach based on collecting patterns from texts by using Wikipedia, DBpedia, and Freebase to build sets of sentences illustrating the occurrence of particular relations. (Gamallo, Garcia, & Fernandez-Lanza, (2012)) adopt a weakly-supervised rule-based approach with the goal of simplifying the linguistic structure of a sentence through partial dependency parsing. This is followed by the application of generic semantic extraction rules over the results which have been obtained via the distant supervision strategy. Schmidek and Barbosa (2014) propose to simplify sentences leveraging clauses, which are NER tagged. Through the use of the Naive Bayes classifier, the authors determine existing relationships between chunks for the purpose of connecting or disconnecting them at a lower computational cost in comparison to parsing.

More recent Open IE systems, ClausIE (Corro et al., 2013) and Vo and Bagheri (2018) use dependency parsing and a small set of domain-independent lexica without any post-processing or training data. At the outset, the approaches exploit linguistic knowledge about the grammar of the English language to first detect clauses in an input sentence and to subsequently identify each clause type based on the grammatical function of its constituents. As a result, these methods are able to generate high-precision extractions and can be flexibly customized to adapt to the underlying application domain.

3. Overview of the proposed approach

Traditional methods for event relation extraction are primarily developed at the sentence level (Mirza et al., 2016a); Zhou, Qian & Fan, 2010) and hence do capture longer event relations that span over multiple sentences or across a document. In order to address this issue, we build our work on Open IE systems that are able to automatically extract triples in the form of (arg1, rel, arg2) representing basic propositions or assertions from text. The main benefit of using Open IE systems is that they are not limited to a sentence level restriction and can operate over a document. As such, we intend to address the limitation of earlier work by exploiting the characteristics of Open IE systems for identifying event relationships. To this end, we benefit from Open IE to build an event network that can capture temporal and causal event relations across different sentences at the document level. The proposed event network aims at identifying the temporal and causal event relations in a document, both within a sentence with direct links and across sentences with indirect links. Our method captures event relations from event flows by detecting paths in the event network. We present algorithms to analyze event flows in the proposed event network that will be used to identify temporal and causal event relations. An overview of our proposed approach is illustrated in Fig. 2. As shown in the figure, we first exploit an Open IE system to extract relation triples in a document in order to build an event network (Fig. 2a). For each triple extracted by the Open IE system, each of the two entities will form a node in the event network and an edge will be connecting the two nodes to each other (Fig. 2b). Once the network is fully built, we identify potential temporal or causal relations between two nodes of the event network by performing specific forms of traversal on the event network (Fig. 2c).

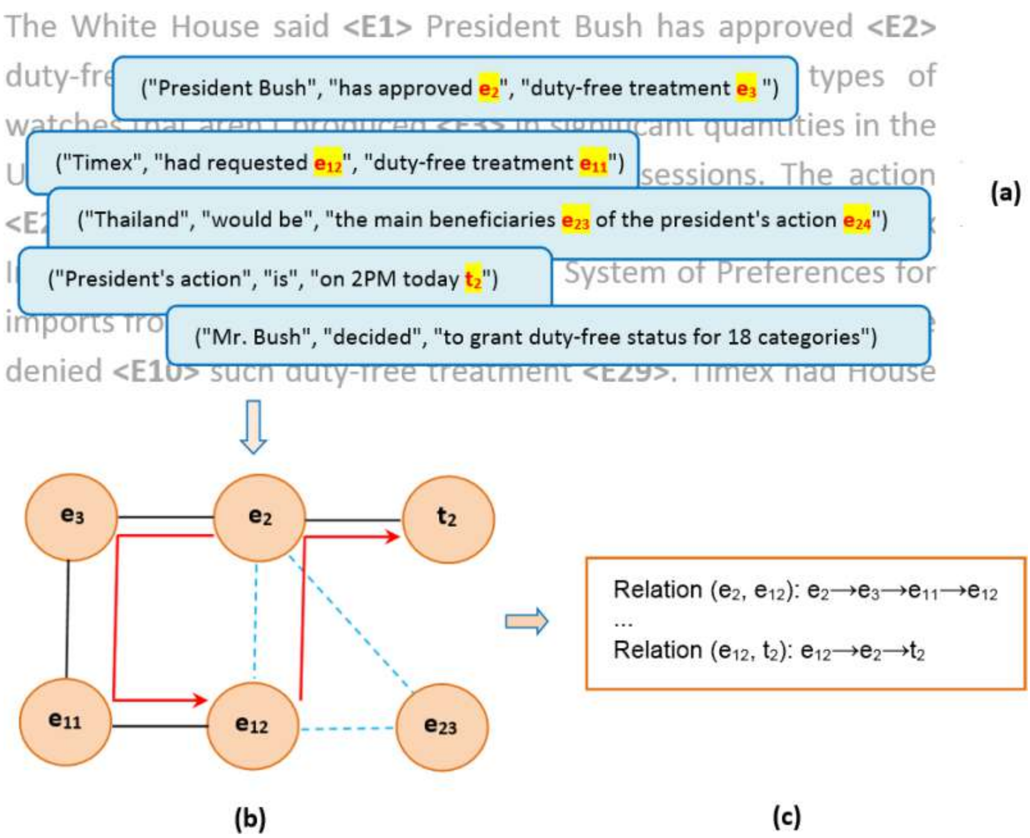


Fig. 2. An overview of the proposed approach with (a) Open IE extraction, (b) event network construction, and (c) event flow extraction.

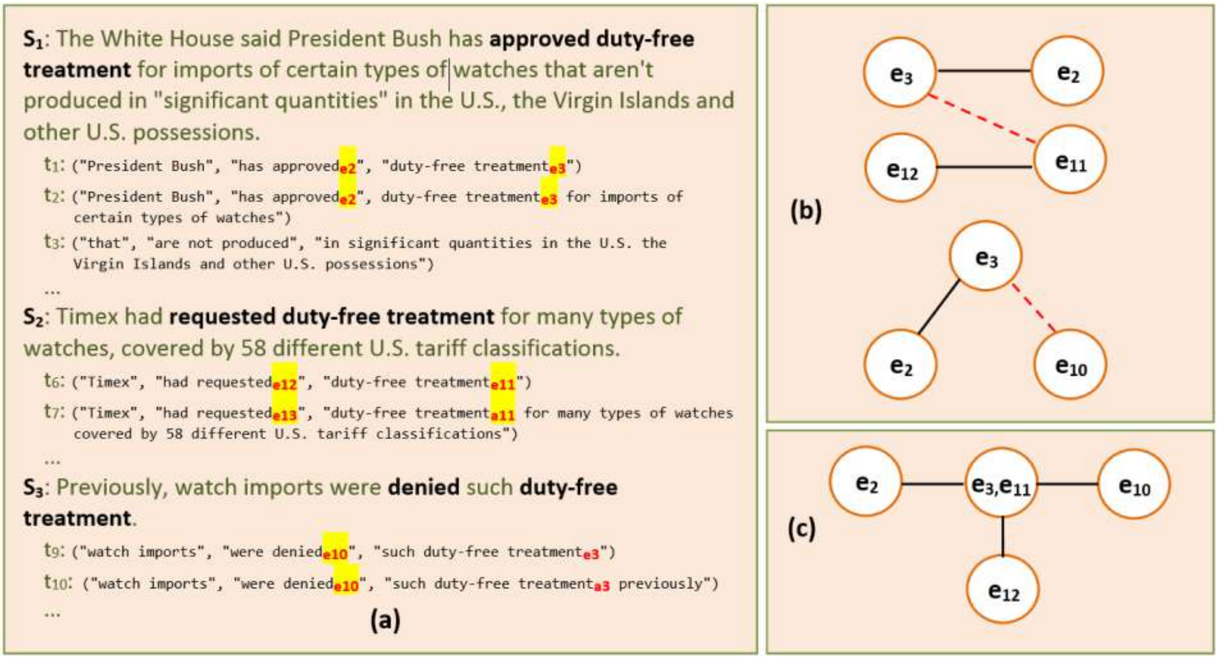


Fig. 3. Building a sample event network with events e_2, e_3, e_{11}, e_{12} and e_{10}. (a) Open IE relation triples; (b) constructing the event network with matching arguments; (c) finalizing the construction of the event network after matching.

3.1. Graph-based event network

Our proposed graph-based event network is built directly from triples generated by Open IE systems. Two events that are present in the same extracted pattern are considered as two event nodes in the event network that are directly connected to each other with an edge. The collections of all the extracted Open IE triple patterns are used to complete the event network. Let us denote an event network as a graph $G(V, E)$ where each vertex denotes an event and each edge denotes a relation. Based on the event network, the objective is to determine the existence and type of relation between two event nodes such as (X, Y) expressed as $R(X, Y)$. In the event network, identifying temporal and causal event relations will involve the consideration of event flows in the form of network paths. For example, given a network with set of relation nodes (e_2, e_3), (e_3, e_{11}), (e_{11}, e_{12}), (e_2, t_1) shown in Fig. 2b, in order to determine the relation (e_2, e_{12}), we need to identify the network path ($e_2 \rightarrow e_3 \rightarrow e_{11} \rightarrow e_{12}$) between e_2 and e_{12}. It can be observed that the relations between (e_2, e_3), (e_3, e_{11}) and (e_{11}, e_{12}) are informative indicators for predicting the relation (e_2, e_{12}). Based on such a network path and if we know that event e_2 began concurrently with event e_3, and e_{12} started before the occurrence of event e_{11} and e_3, then we can determine that event e_2 happened after e_{12}.

Let us first cover the process through which we build the event network. The basic idea behind building the event network is that it operates based on a set of Open IE triples and a set of events representing the arguments of these triples. Summarily, we process each triple produced by the Open IE system by creating a matching vertex for each of the triple's arguments in the event network. The two vertices corresponding to the triple arguments would then be linked together through an edge in the event network. It should be noted that a *matching* method is used in order to ensure that events already identified in previous triples do not represent new vertices and are matched to existing vertices in the event network, e.g. events e_3 and e_{11} shown in Fig. 3a and b are matched with each other as they both reference “duty-free treatment”. After processing each individual triple and creating corresponding vertices and edges, as well as ensuring that similar vertices are folded into one to avoid duplicates, the resulting graph forms our intended event network.

Now given this event network and by considering the shortest path between each pair of nodes in the network (Olya and Hessam, 2014), it is possible to infer the type of the event relation between source and target nodes based on the set of event relations observed on the shortest path. Indirect event relations can be inferred through transitivity of relations (Allen, 1983; Laokulrat et al., 2015) on direct relations as shown in Fig. 4. For instance, consider events e_1 and e_4 in Fig. 4a. Given the fact that the shortest path between events e_1 and e_4 passes through e_2, it is possible to infer that because event e_1 happened before e_2 and e_2 occurred prior to event e_4 that e_1 also happened before e_4.

3.2. Temporal relation identification

According to the TempEval-3 task description (Mirza et al., 2016a and Laokulrat et al., 2015), two pairs of temporal events can be related to each other through one of four groups, namely Timex-Timex, Event-DCT, Event-Timex and Event-Event where DCT denotes *Document Creation Time* and Timex denotes *Temporal Expressions*. A pair of temporal entities, including events and temporal

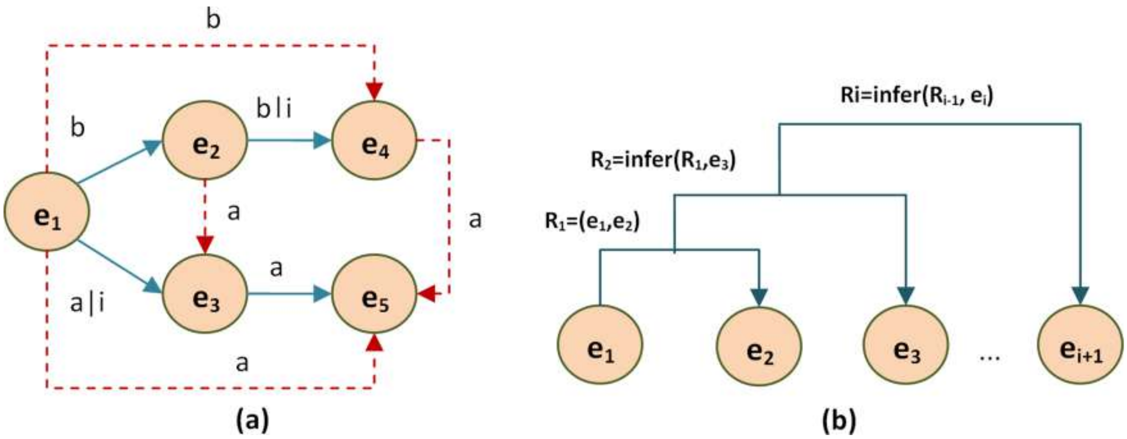


Fig. 4. Inferred relations; (a) Inferred sample relations in a three node path (e_1-e_4, e_2-e_3, e_1-e_5, e_4-e_5) with *b*: before; *i*: includes; *a*: after; (b) Recursively inferred relations.

Algorithm 1
The outline of the process for identifying temporal relations.

Input: Graph $G = (V, E)$
Pair of events $\{X, Y\}$
Output: Type of relation $R(X, Y)$
if Direct$(X, Y) \in G$ **then**
$R(X, Y) \leftarrow$ Rules(E$\{X, Y\}$, G)
Update(G)
else if Indirect$(X, Y) \in G$ **then**
R$(X, Y) \leftarrow$ Time$\{X,Y\} \oplus$ Tense$\{X,Y\}$
Update(G)
if R(X, Y) = NULL **then**
Event-flow$\{X,Y\} \leftarrow$ Shortest-path($\{X, Y\}$, G))
Rules(Event-flow$\{X,Y\}$, G)
$R(X, Y) \leftarrow$Infer(Event-flow$\{X$,Y$\}$)
Update(G)
end if
return $R(X, Y)$

expressions, annotated as a temporal relation, is called a TLINK. Temporal relation recognition is the task of classifying TLINKs into temporal relation types. It uses a complete set of TLINK relations, which consists of 13 types of temporal relations shown in Table 1.

In this paper, we present a method to extract both direct and indirect temporal event relations in the event network. The objective is to determine the existence and type of temporal relation between two event nodes such as (X, Y) expressed as $R(X, Y)$ in the event network. Algorithm 1 presents the pseudo-code for our proposed algorithm for identifying temporal relations. The basic idea behind our algorithm is that it takes a Graph G(V, E) and a pair of events (X, Y) and subsequently determines relation R(X, Y). Our algorithm distinguishes between relations whose nodes are directly connected to each other in the network, denoted by Direct(X, Y), and those events whose nodes are not directly connected, shown as Indirect(X, Y). If the relation is concerned with two directly connected events, we employ a rule-based model to determine the type of the relationship.

Determining temporal relation types for direct relations

Inspired by the work in Chambers et al. (2014) and Mirza et al. (2016a), the temporal rule-based model relies on hand-crafted rules. In the rule-based model, we separate temporal constructs containing temporal features of event relations. Some of the relations between the events are captured based on morpho-syntactic information from their textual expression. Several relations are based on semantic information such as typical event duration while other relations are computed independently based on dependency paths. We also include rules for predicting relations between reported events (Timex and DCT events) and other events based on their time characteristics and syntactically captured tense. We introduce four categories of rules as follows:

Timex-Timex (T-T). For Timex-Timex relations, we take temporal expressions such as DATE, TIME and DURATION into consideration and then identify the relation types based on their normalized values. For example, the day after last Monday will be normalized into 15-Jan-2019 if last Monday was 16-Jan-2019.

Event-DCT (E-DCT). To predict relation types of Event-DCT event pairs, the rules are based on tense and aspect of entity events. For instance, the event entity "will attack$_E$", which is in the future tense will be recognized as "AFTER". We define separate rules as

Table 1
Temporal relation types.[1]

Relation	Description
BEFORE	One before the other
AFTER	One after the other
BEGINS	One being the beginning of the other
BEGUN_BY	One being begun by the other
DURING	One holding during the duration of the other
END_BY	One being ended by the other
ENDS	One being the ending of the other
IBEFORE	One immediately before the other
IAFTER	One immediately after the other
IDENTITY	Referring to the same event
INCLUDES	One including the other
IS_INCLUDES	One being included in the other
SIMULTANEOUS	One is simultaneous with the other

follows:

if $e \in$ Tense(PAST) ***then*** label(BEFORE)
if $e \in$ Tense(PRESENT) ***and*** Tense_Progressive (E, D) ***then*** label(INCLUDES)
if $e \in$ Tense(FUTURE) ***then*** label(AFTER)

Event-Timex (E-T). The rules built for Event-Time exploit the senses of prepositions (Litkowski, 2014). We consider extracting the dependency path between an event *e* and a timex *t* via modifiers. Six types of time prepositions are defined for constructing the rules such as STARTTIME (e.g., from, since), ENDTIME (e.g., until), DURATION (e.g., during), FOLLOWTIME (e.g., after), PRECEDINGTIME (e.g., before), and POINTTIME (e.g., in, on, at). For instance, when *t* is presented in the POINTTIME form and its dependency modifier is `compound`, the event relation will be recognized as 'IS INCLUDED'. These rules are defined as follows:

if DP NMOD(*E, T*) $\in$ STARTTIME ***then*** label(BEGUN_BY)
if DP NMOD(*E, T*) $\in$ ENDTIME ***then*** label(END_BY)
if DP NMOD(*E, T*) $\in$ DURATIONTIME ***then*** label(DURING)
if DP NMOD(*E, T*) $\in$ FOLLOWTIME ***then*** label(AFTER)
if DP NMOD(*E, T*) $\in$ PRECEDINGTIME ***then*** label(BEFORE)
if DP NMOD(*E, T*) $\in$ POINTTIME ***then*** label(IS_INCLUDED)

Event-Event (E-E). The rules for event pairs (e_1,e_2) are constructed in two sets based on dependency paths and the verbs surrounding the events. We extract the dependency path between event e_1 and event e_2 to build the first set of rules. For the second set, the rules are built based on the tense aspects of events e_1 and e_2. For instance, if event e_2 is the logical subject of e_1 as in "…the chain reaction touched$_{e1}$ off by the collapse$_{e2}$ of Soviet Union", events e_1 and e_2 are connected by an AFTER relation. The following rules are based on the tense aspect of each event as follows:

if DP LGS-PMOD(e_1, e_2) and VERB-AUXPASS(e_1) ***then*** label(AFTER)
if DP LOC-PMOD(e_1, e_2) and AVERB-LOC(e_1) ***then*** label(IS_INCLUDED)
if DP XCOMP(e_1, e_2) and PROGRESSIVE(e_1) ***then*** label(SIMULTANEOUS)
if DP XCOMP(e_1, e_2) and VERB-TERMINATION(e_1) ***then*** label(ENDS)
if DP XCOMP(e_1, e_2) and VERB-INITIATION (e_1) ***then*** label(BEGIN)
if DP XCOMP(e_1, e_2) and VERB-CONT(e_1) ***then*** label(INCLUDES)
if DP XCOMP(e_1, e_2) ***then*** label(BEFORE)

Determining temporal relation types for indirect relations

Once the type of relations for the events that are directly connected to each other in the event network are determined based on the introduced rule sets, we determine the relation types for indirect relations based on their event flow. Algorithm 1 proceeds to detect types of each event relations primarily based on transitivity between direct relations in the shortest path between any given two events as shown in Fig. 4a. It is possible to reason over the set of temporal relations observed on the shortest path to make a determination about the type of temporal relation between the two source and target events. In our work, indirect relations are

[1] http://www.timeml.org

Table 2
Transitive inference for determining of indirect temporal relation types. UNDEF denotes undefined relation.

	BEFORE	AFTER	IBEFORE	IAFTER	IDENTIFY	INCLUDES	IS_INCLUDED	DURING	SIMULTANEOUS	BEGINS	BEGUN_BY	END	END_BY
BEFORE	BEFORE	UNDEF	BEFORE	UNDEF	BEFORE	BEFORE	UNDEF	BEFORE	BEFORE	BEFORE	BEFORE	UNDEF	BEFORE
AFTER	UNDEF	AFTER	UNDEF	AFTER	AFTER	AFTER	UNDEF	UNDEF	AFTER	UNDEF	AFTER	AFTER	AFTER
IBEFORE	BEFORE	UNDEF	BEFORE	UNDEF	IBEFORE	BEFORE	UNDEF	UNDEF	BEFORE	BEGIN	BEGIN	UNDEF	BEFORE
IAFTER	UNDEF	AFTER	UNDEF	AFTER	IAFTER	AFTER	UNDEF	UNDEF	AFTER	UNDEF	AFTER	IAFTER	IAFTER
IDENTIFY	BEFORE	AFTER	IBEFORE	IAFTER	INDENTIFY	INCLUDES	IS_INCLUDED	DURING	SIMULTANEOUS	BEGINS	BEGUN_BY	END	END_BY
INCLUDES	UNDEF	UNDEF	UNDEF	UNDEF	INCLUDES	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF
IS_INCLUDED	BEFORE	AFTER	BEFORE	AFTER	IS_INCLUDED	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF
DURING	BEFORE	AFTER	BEFORE	AFTER	DURING	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF
SIMULTANEOUS	UNDEF	UNDEF	UNDEF	UNDEF	SIMULTANEOUS	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF
BEGINS	BEFORE	AFTER	BEFORE	IAFTER	BEGINS	UNDEF	UNDEF	UNDEF	UNDEF	BEGINS	UNDEF	UNDEF	UNDEF
BEGUN_BY	UNDEF	AFTER	UNDEF	END_BY	BEGUN_BY	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	BEGUN_BY	UNDEF	UNDEF
ENDS	BEFORE	AFTER	IBEFORE	AFTER	END	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	END	UNDEF
END_BY	BEFORE	UNDEF	IBEFORE	UNDEF	END_BY	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	UNDEF	END_BY

Table 3
Dependency paths between two events with causal verbs.

Types	Dependency path
Subject(e_1) → Verb	
+ e_1 is subject of v	`SBJ`
+ v is predicative complement of e_1	`PRD-IM`
+ v is modifier of e_1 (nominal)	`NMOD`
+ v is apposition of e_1	`APPO`
+ v is general adverbial of e_1	`ADV`
+ v is adverbial of purpose/reason of e_1	`PRP-IM`
Verb → Object(e_2)	
+ e_2 is object of v	`OBJ`
+ e_2 is logical subject of v (passive verb)	`LGS-PMOD`
+ e_2 is predicative complement of v (raising/control verb)	`OPRD, OPRD-IM`
+ e_2 is general adverbial of v	`ADV-PMOD`
+ e_2 is adverbial of direction of v	`DIR-PMOD`
+ e_2 is modifier of v (adjective or adverbial)	`AMOD-PMOD`
Verb(e_1) → Verb(e_2)	
+ v_2 is predicate of v_1	`XCOMP`

inferred through transitivity of temporal relations (Allen, 1983; Laokulrat et al., 2015). The possible transitive inferences for temporal relations are shown in Table 2. The table shows how a transitive inference will be made given the temporal relations shown on the columns and rows of the table. For instance, if event X occurs BEFORE event E and event E happens immediately BEFORE event Y, we can infer a new temporal relation "X happens BEFORE Y" based on transitive inference in Table 2 (the cell at the intersection of the BEFORE column with the BEFORE row).

3.3. Causal relation identification

Similar to temporal relations, we propose an algorithm for identifying causal relations (CLINK/CLINK-R) based on the proposed event network shown in Algorithm 2. The process for identifying causal relations follows the same flow as that of determining temporal relations. We distinguish between direct and indirect relations.

Algorithm 2
The process for identifying causal event relations.

Input: Graph $G = (V, E)$
Pair of events $\{X, Y\}$
List of CA Verbs K
Output: Type of relation $R(X, Y)$
if Direct(X, Y) $\in G$ **then**
$R(X, Y)$ ←Causal-Rules(K, $E\{X, Y\}$, G)
Update(G)
else if Indirect(X, Y) $\in G$ **then**
Event-flow$\{X,Y\}$ ← Shortest-path($\{X, Y\}$, G))
Causal-Rules(K, Event-flow$\{X, Y\}$, G)
$R(X, Y)$ ←Infer(Event-flow$\{X, Y\}$)
Update(G)
end if
return $R(X, Y)$

Determining causal relation types for direct relations

In the case of direct relations, we identify terms that belong to affect, link or causative verbs and find specific dependent structures in which such terms are associated with two directly related events. For instance, one could determine a causal relation between a pair of events (e_1, e_2) when there exists a causal verb v, for which e_1 is the subject of verb v, and e_2 is either object of v or complement of v. Such relations between events and causal verbs are usually expressed syntactically; therefore, we define rules for identifying pairs of events being related to a causal verb in a causal construct by looking at their dependency paths. If such dependencies are found, a CLINK/CLINK-R is automatically set between the two events. Table 3 shows the details of dependency paths for this purpose. We consider three types of interactive causal verbs between events as Subject → Verb, Verb → Object and Verb → Verb in order to define the rules. For example, in the case of Subject → Verb and Verb → Object, an event e_1 is predicate complement of cause verb v whose object is event e_2 showing dependencies `PRD-IM` and `OBJ`, respectively. For instance, in the sentence "*The assessment*$_{\mathbf{e1}}$ *was to stop*$_{\mathbf{v}}$ *their investment*$_{\mathbf{e2}}$*…*", 'e_1 causes e_2' because `PRD-IM`(assessment, stop) and `OBJ`(stop, investment).

Our work differentiates itself from Mirza et al. (2016a), which uses a predefined set of causal verbs that could face limitations when used in an *open* domain. In our work, we define the causal verbs for the rules presented in Table 3 by measuring the *causality*

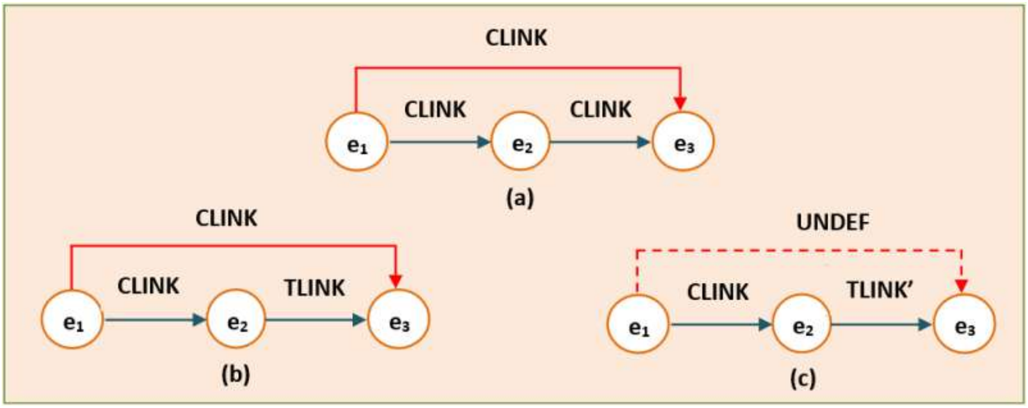


Fig. 5. Transitivity process of causal relation. a) Causal relation (CLINK) → causal relation (CLINK); b) Causal relation (CLINK) → temporal reverse dependent relation (TLINK); c) Causal relation (CLINK) →temporal dependent relation (TLINK').

association (cause-effect) between a pair of events e_1 and e_2 (Do et al., 2011). In order to calculate cause-affect association, we measure pointwise mutual information (*PMI*) using two separate components between events with verbs, where we denote P_{pp} by measuring the association between event predicates, and P_{pa} by measuring the association between the predicate of an event and the arguments of the other event. Specifically, we measure cause-affect association (*CA*) as follows:

$$CA(e_1,\ e_2) = P_{pp}(e_1,\ e_2) + P_{pa}(e_1,\ e_2) \tag{1}$$

where $P_{pp}(e_1, e_2)$ is defined as:

$$P_{pp}(e_1,\ e_2) = PMI(p_{e_1},\ p_{e_2}) \times IDF(p_{e_1},\ p_{e_2}) \tag{2}$$

which takes into account the *PMI* between verbs p_{e_1} and $p_{e_2}$of event e_1 and $e_{2,}$ respectively. In Suppes' Probabilistic theory of Causality (Do et al., 2011; Suppes, 1970), the author showed that the first event e_1 is a possible cause of the second event e_2, if e_1 happens more frequently with e_2 than by itself, i.e. $\mathrm{P}(e_1) > P(e_2)$. This can be easily rewritten as $\mathrm{P}(e_1, e_2)/\mathrm{P}(e_1)\mathrm{P}(e_2) > 1$, similar to the definition of PMI:

$$PMI(p_{e_1},\ p_{e_2}) = \log\frac{P(p_{e_1},\ p_{e_2})}{P(p_{e_1})P(\ p_{e_2})} \tag{3}$$

and,

$$IDF(p_{e_1},\ p_{e_2}) = \log\frac{T}{1+N} \tag{4}$$

where T is the total triple tuples and N is the number of triple tuples that contains $p_{e_1},\ p_{e_2}$.

We further define $P_{p,\ a}(e_1,\ e_2)$ as follows:

$$P_{pa}(e_1,\ e_2) = \frac{1}{|A_{e_2}|}\sum_{a\in A_{e_2}} PMI(p_{e_1},\ a) + \frac{1}{|A_{e_1}|}\sum_{a\in A_{e_1}} PMI(p_{e_2},\ a) \tag{5}$$

where A_{e_1} and A_{e_2} are the list of arguments from T that contain e_1 and $e_{2,}$ respectively.

Determining causal relation types for indirect relations

Now for indirect causal relations, on the basis of the transitivity process proposed by Allen (1983), we construct a set of inference relations in order to determine the transitivity of causal relations on both *causal relation → causal relation* (CLINK→CLINK) and *causal relation → temporal relation* (CLINK→TLINK) as shown in Fig. 5. Regarding the transitive process for CLINK→CLINK (Fig. 5a), we base it on the characteristics of causal relations such that if event e_1 causes event e_2 and event e_2 causes event e_3, then e_3 will be considered to depend on e_1. Therefore, a transitive relation will be made between e_1 and e_3 as e_1 causes e_3. Regarding CLINK→TLINK relations, the first part of the relation represents a causal dependency between two events while the second part of the relation expresses a temporal relation. In such cases (Fig. 5b), similar to CLINK→CLINK, a new transitive causal relation can be generated between events in causal and temporal relations. For instance, when event e_1 causes event e_2 and event e_3 has a temporal relation with event e_2, this could, in the case of *reverse temporal dependent relations*, mean that event e_3 is dependent on event e_1. Consequently, event e_1 will cause event e_3. Similar to Mirza et al. (2016b), we classify these reverse temporal dependent relations where e_3 is dependent on e_2 as IDENTITY, DURING, INCLUDES, BEGINS, ENDS and SIMULTANEOUS. For the case of *temporal dependent relation* in CLINK→TLINK' (Fig. 5c), although e_2 is dependent on e_1 but e_3 is not dependent on e_2. Therefore, a relation, expressed as "UNDEF", is defined between e_1 and e_3. The list of *temporal dependent relation* where a clear causal relation cannot be inferred from them include IAFTER,

Table 4
Transitive inference of causal relation types. UNDEF specifies undefined relation.

	CLINK	CLINK-R
Causal link		
CLINK	CLINK	UNDEF
CLINK-R	UNDEF	CLINK-R
Reverse temporal dependent link		
IDENTITY	CLINK	CLINK-R
DURING	CLINK	CLINK-R
INCLUDES	CLINK	CLINK-R
BEGINS	CLINK	CLINK-R
ENDS	CLINK	CLINK-R
SIMULTANEOUS	CLINK	CLINK-R
Temporal dependent link		
IAFTER	UNDEF	UNDEF
IBEFORE	UNDEF	UNDEF
IS_INCLUDED	UNDEF	UNDEF
BEGUN_BY	UNDEF	UNDEF
ENDED_BY	UNDEF	UNDEF

Table 5
Details of the TempEval-3 corpus.

Corpus	Documents	Open IE 5 triples	LS3RyIE triples	#relations
Training dataset	183	2398	2537	2191
Testing	20	896	968	1173

IBEFORE, IS_INCLUDED, BEGUN_BY and ENDED_BY. In summary, the details of the transitive inference process is presented in Table 4. We exploit our proposed event network to detect CLINK between e_1 and e_2 and TLINK between e_2 and e_3 for predicting a causal relation between e_1 and e_3 by following the set of transitive inference rules that are listed in Table 4, e.g., if e_1 CLINK e_2 and e_2 IDENTITY e_3, then we infer a new causal relation 'e_1 CLINK e_3'.

4. Performance evaluation

We have proposed methods to extract event relations based on the proposed event network aiming at two tasks: 1) Temporal event extraction and 2) Causal event extraction. In this section, to carry out evaluations for these tasks, we conduct experiments on several benchmark datasets and compare the performance of our proposed work with strong baseline methods. Particularly, in the first task, we use TempEval-3 corpora. Furthermore, Causal-TimeBank corpora, developed by Mirza & Tonelli, (2014a), will be used for the second task. We employ LS3RyIE (Vo & Bagheri, 2018) and Open IE 5,[2] which are Open IE systems, to extract relation instances for constructing the event network:

- **Open IE 5.** This IE system, which is a combination of the OLLIE (Mausam et al., 2012), RelNoun (Pal & Mausam, 2016), CALMIE (Saha & Mausam, 2018) and SRLIE (Christensen, Mausam, Soderland, & Etzioni, 2011) methods, performs deep analysis on the identified verb-phrase relations and then extracts all relations mediated by verbs, nouns, and adjectives, among others.
- **LS3RyIE.** This IE system is an extension of the work by ClausIE (Corro et al., 2013) and exploits linguistic knowledge about the grammar of the English language to detect clauses based on the grammatical function of its constituents.

Note that, we only consider triples which contain event entities for building the network. Note that, reference mapping was also applied to enhance node matching in the event network. We calculated the context from the Open IE patterns using cosine similarity on term frequency vector then merged those nodes with a score ≥ 0.5 We will show how our proposed work can identify temporal and causal relations and reduces the number of erroneous relation types compared to previous baseline approaches.

4.1. Temporal event relation extraction

4.1.1. Experimental dataset

For benchmarking our approach in this task, we have conducted experiments on the Tempeval-3 corpora task C. TempEval-3 (Pustejovsky, Littman, Saurí & Verhagen, 2006) consists of an annotated corpus of temporal relations that was created following the

[2] https://github.com/dair-iitd/OpenIE-standalone

Table 6
Extraction samples on Open IE 5 and LS3RyIE systems.

Sentence 1: "Imports of the types of watches$_{e20}$ that now will be eligible$_{e21}$ for duty-free treatment$_{e25}$ totaled$_{(event3)}$ about \$ 37.3 million in 1988$_{txm2}$, a relatively small share of the \$1.5 billion in U.S. watch imports that year$_{tmx3}$, according to an aide to U.S. Trade Representative Carla Hills".
Open IE 5:
$t_{1\text{-}1}$: (the \$ 1.5 billion in U.S.; watch; imports; T:that year)
$t_{1\text{-}2}$: (Imports of the types of watches; totaled; about \$ 37.3 million; T:in 1988; according to an aide to U.S. Trade Representative)
$t_{1\text{-}3}$: (watches; will be; eligible for duty-free treatment; T:now)
$t_{1\text{-}4}$:8 (Carla Hills; [is] Trade Representative [from]; United States)
LS3RyIE:
$t_{1\text{-}5}$: ("Imeports of the types of watches ", "will be", "eligible now")
$t_{1\text{-}6}$: ("Imports of the types of watches ", "will be", "eligible for duty-free treatment")
$t_{1\text{-}7}$: ("Imports of the types of watches ", "will be", "eligible")
$t_{1\text{-}8}$: ("Imports of the types of watches", "totaled", "about \$ 37.3 million in 1988 a relatively small share of the \$ 1.5 billion in U.S. watch imports that year according to an aide to U.S. Trade Representative Carla Hills")
$t_{1\text{-}9}$: ("Imports of the types of watches", "totaled", "about \$ 37.3 million a relatively small share of the \$ 1.5 billion in U.S. watch imports that year according to an aide to U.S. Trade Representative Carla Hills")
$t_{1\text{-}10}$: ("Imports of the types of watches", "totaled", "about \$ 37.3 million")
Sentence 2: "Polls have shown$_{e31}$ public support$_{e32}$ for a republic increasing$_{e33}$, rising$_{e34}$ from about 35 percent several years ago$_{tmx10}$ to about 51 percent this year$_{tmx11}$, as pro-monarchist sentiment diminishes."
Open IE 5:
$t_{2\text{-}1}$: (Polls; have shown; public support for a republic increasing)
LS3RyIE:
$t_{2\text{-}2}$: ("Polls", "have shown", "public support for a republic increasing rising from about 35 percent several years ago to about 51 percent this year as pro-monarchist sentiment diminishes")
$t_{2\text{-}3}$: ("a republic increasing", "be rising", "from about 35 percent several years ago to about 51 percent this year")
$t_{2\text{-}4}$: ("a republic increasing", "be rising", "from about 35 percent several years ago as pro-monarchist sentiment diminishes")
$t_{2\text{-}5}$: ("a republic increasing", "be rising", "from about 35 percent several years ago")
$t_{2\text{-}6}$: ("pro-monarchist sentiment", "diminishes")

TimeML specification. It contains news articles with 183 training documents and 20 testing documents, with just over 61,000 non-punctuation tokens, coming from a variety of news reports, specifically from outlets such as ABC, CNN, PRI, VOA, Wall Street Journal and newswire from AP and NYT. The task of TempEval-3 corpus has been designed by UzZaman et al. (2013) at SemEval-2013, with a complete set of 13 TLINK types. We use the same training and test data released in the context of Tempeval-3 for evaluation. The distribution of the relation types in training and test datasets is shown in Table 5.

We compare our method with several strong baseline approaches designed for Task C of TempEval-3 named UTTime (Laokulrat et al., 2013), TRelPro (Mirza & Tonelli, 2014b), Laokulrat et al. (2015) and CATENA (Mirza et al., 2016a). UTTime (Laokulrat et al., 2013) exploited features at the sentence-level such as predicate-argument structure of relations in a sentence using syntactic information and lexical semantic information to classify temporal relation types by SVM. TRelPro (Mirza & Tonelli, 2014b) employed an SVM classifier based on event linguistic features such as POS tags, chunking, and dependency paths. Laokulrat et al. (2015) extracted temporal event relations using time graphs and stacked learning. Staked learning is a supervised learning framework that performs temporal inference based on local and non-local features for temporal relation classification. CATENA (Mirza et al., 2016a) integrated rule-based models and data-driven classifiers using linguistic features based on the analysis of the syntactic features, dependency features, and shallow semantic structure of text. The system predicts temporal relation types using three supervised classification models on E-D, E-T, E-E using SVM.

4.1.2. Evaluation results

We evaluate our algorithm on both training and testing datasets for predicting event relations because our approach is unsupervised in nature. We first extracted triple patterns from Open IE systems to construct the proposed event network in each document. We built the event network based on the generated Open IE triples extracted by the LS3RyIE and Open IE 5 systems. As a result of this process, 968 and 2537 triples were generated by LS3RyIE while Open IE 5 generated 2398 and 916 triples that were then used to build the event network for the testing and training sets, respectively. Several sample triples extracted from similar sentences using Open IE 5 and LS3RyIE are shown in Table 6. In light of the fact that LS3RyIE explores the clause structure of the sentence, the system has extracted a higher number of triples. In the process of exploring the clause structure, the adverbials in a clause are considered in addition to the verb or verbal phrases adverbials for extraction. LS3RyIE generated 4 triples and 6 triples in Sentence 1 and Sentence 2, respectively. Regarding Open IE 5, the system extracts verb phrase-based relations building a set of syntactic and dependency constraints to identify relations based on verb phrases then finds a pair of arguments for each identified relation phrase. Open IE 5 extracted 4 triples and 1 triples in Sentence 1 and Sentence 2, respectively. We observed that number of extractions of LS3RyIE and Open IE are seemly similar in Sentence 1. However, Open IE 5 focuses on the main verb and the information surrounding it that could limit the identification of the subject and object when faced with a long sentence. For example, the system fails to identify the subject and object in the structure between "republic" and "rising" in Sentence 2. Consequently, several events that describe entity events or temporal events, e.g., "rising", "year ago", "this year" and "diminishes", are missing from the extracted triples from Open IE 5. In the process of building the event network, an event network is effectively constructed when most of events are extracted and fully presented in the network based on extracted triples from Open IE systems. One of the important considerations

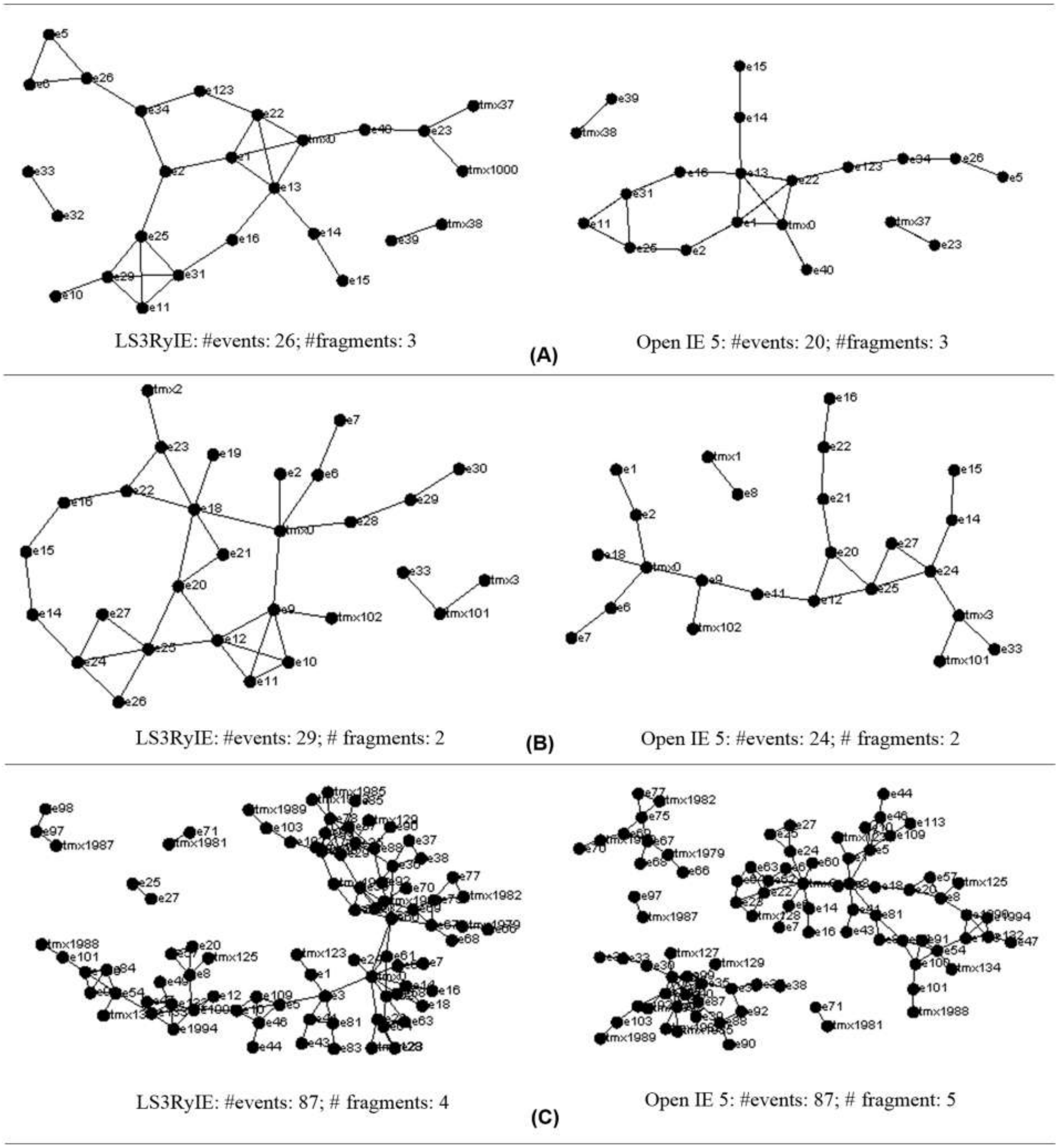


Fig. 6. Samples of event networks from LS3RyIE and Open IE 5; a) Event network on document wsj_0026; b) Event network on document nyt_20130322; and c) Event network on document APW19980301.0720.

that impacts the quality of relations between events is the coherence in the network. An effective network will show a strong coherence of events when it has fewer fragments with a high number of extracted events. Figs. 6 shows some samples of event networks based on the extracted triples from LS3RyIE and Open IE 5. LS3RyIE has fewer fragments with a higher number of extracted events compared to Open IE 5 when building an event network for the same three documents. The networks have (#events = 26; #fragments = 3), (#events = 29; #fragments = 2), and (#events = 87; #fragments = 4) for documents wsj_0026, nyt_20130322 and APW19980301.0720 based on LS3RyIE, respectively. Using Open IE 5, the networks have with (#events = 20; #fragments = 3), (#events = 24; #fragments = 2), and (#events = 87; #fragments = 5) for documents wsj_0026, nyt_20130322 and APW19980301.0720, respectively.

Table 7
Experimental results on four relation types on the TempEval-3 corpus.

Categories	Open IE 5			LS3RyIE		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Training set						
BEFORE	58.51	44.44	50.51	75.74	62.77	68.65
AFTER	55.97	36.95	44.51	76.56	59.80	67.15
BEGINS	21.05	14.29	17.02	62.50	52.63	57.14
BEGUN_BY	20.00	12.50	15.38	75.00	50.00	60.00
DURING	18.42	13.72	15.73	75.00	64.29	69.23
END_BY	15.79	11.11	13.04	36.36	36.36	36.36
ENDS	25.00	20.00	22.22	55.55	55.55	55.55
IBEFORE	100.00	60.00	75.00	83.33	50.00	62.50
IAFTER	60.00	30.00	40.00	85.00	70.00	76.77
IDENTITY	43.75	26.42	32.94	66.67	50.54	57.50
INCLUDES	49.71	37.83	42.96	57.89	49.25	53.22
IS_INCLUDED	64.95	49.44	56.14	72.65	59.44	65.38
SIMULTANEOUS	30.98	19.13	23.65	57.35	48.75	52.70
Testing set						
BEFORE	69.80	45.82	55.32	77.54	60.17	67.76
AFTER	62.93	39.25	48.35	74.03	59.07	65.71
BEGINS	66.66	66.66	66.66	50.00	50.00	50.00
BEGUN_BY	50.00	33.33	40.00	80.00	75.50	77.68
DURING	50.00	50.00	50.00	66.66	25.00	36.36
END_BY	100.00	50.00	66.67	50.00	33.33	40.00
ENDS	50.00	33.33	40.00	100.00	66.66	80.00
IBEFORE	50.00	33.33	40.00	66.66	66.66	66.66
IAFTER	60.00	50.00	54.55	70.33	43.25	53.56
IDENTITY	09.09	09.09	9.09	33.33	40.66	36.63
INCLUDES	48.07	34.25	40.00	64.52	53.34	58.40
IS_INCLUDED	72.80	52.30	60.87	83.01	69.39	75.59
SIMULTANEOUS	24.39	14.93	18.52	40.81	28.57	33.61

Table 8
Experimental results on four types of temporal event relations.

Categories	Open IE 5			LS3RyIE		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Training set						
E-E	37.28	22.46	28.03	60.63	41.81	49.49
E-T	64.12	42.25	50.94	84.82	73.21	78.58
E-D	66.25	61.92	64.01	88.24	87.50	87.86
T-T	41.30	25.00	31.15	62.50	62.50	62.50
Overall	53.32	37.28	43.88	69.58	57.18	62.77
Testing set						
E-E	46.06	26.29	33.47	60.51	47.21	53.04
E-T	81.43	60.17	69.20	82.37	60.06	69.47
E-D	68.48	62.43	65.32	70.63	68.75	69.68
T-T	42.86	30.00	35.30	77.77	68.62	72.91
Overall	59.75	40.47	48.26	72.92	57.54	64.32

Table 9
Performance comparison.

	Precision	Recall	F-measure
UTTime (Laokulrat et al., 2013)	55.60	57.40	56.50
TRelPro (Mirza & Tonelli, 2014b)	58.48	58.80	58.17
Laokulrat et al. (2015)	57.60	57.90	57.80
CATENA (Mirza et al., 2016a)	62.60	61.30	61.90
Proposed method (Open IE 5)	57.51	39.36	46.73
Proposed method (LS3RyIE)	70.82	57.31	63.35

S_{10}: He's not telling$_{\mathbf{e14}}$ the truth.
OIE triples: ("He", "is not telling", "the truth")
S_{12}: She was shot$_{\mathbf{e20}}$ with a rifle.
OIE triples: ("She", "was shot with", "a rifle")

Fig. 7. Sample triples obtained through Open IE.

It should be noted that unlike the baselines that are supervised temporal relation extraction methods, our work is completely unsupervised and as such we do not require separate training and testing datasets. For this reason, we report the performance of our work on the data available in both sets. Table 7 shows the detailed performance of our proposed approach over 13 relation types. Among the relation types, our approach obtains high performance in BEFORE, IBEFORE, and IS_INCLUDED temporal relation types. The BEFORE relation type yielded 68.65% and 67.76% of F-measures based on LS3RyIE and 50.51% and 55.32% of F-measures based on Open IE 5 on the two parts of the dataset and the IS_INCLUDED relation type obtained 65.38% and 75.59% of F-measures in LS3RyIE and 56.14% and 60.87% of F-measures in Open IE 5 in training and testing datasets, respectively. Our method is successful in recognizing event relation types due to the use of transitivity between direct and indirect links in the event network. We observed that several relations are affected or dependent on other relations in the event flow, e.g., it is possible to predict BEFORE($e_1 \rightarrow e_3$) in the path $e_1 \rightarrow e_2 \rightarrow e_3$ when it is presented as BEORE($e_1 \rightarrow e_2$), and IS_INCLUDED($e_3 \rightarrow e_2$). This explains also the large number of positively predicted temporal relation types in our work.

However, our method obtained lower results in a few relations such as DURING, END_BY, and SIMULTANEOUS. Upon further exploration of these relations, we found that the Open IE systems used in our experiments struggles to identify the required triples for these relation types primarily due to the sparseness of the DP link in these events. For instance, when predicting the SIMULTANEOUS relation between e_1 and e_2 in the following sentence "… *refer*$_{\mathbf{e1}}$ *to U.N. resolution 425 in a speech …that would not be attacked, Yitzhak Mordechai said*$_{\mathbf{e2}}$ *at that time*.", the Open IE system could not extract the direct triple (e_1, e_2), ("*a speech*", "*said*") or ("*a speech*", "*time*") because there were no dependency parsing (DP) links between them. As such, these events were not directly connected to each other in the event network and therefore the relation type was not subsequently identified.

We have also performed a more detailed evaluation by dividing temporal relations into four categories of temporal event types as shown in Table 8. In the training set, our approach achieved F-measures of 49.49%, 78.58%, 87.86%, and 62.50% based on LS3RyIE and F-measures of 28.03%, 50.94%, 64.01%, and 31.15% based on Open IE 5 for Event-Event (*E-E*), Event-Time (*E-T*), Event-DCT (*E-DCT*), and Time-Time (*T-T*), respectively. On the training set, our work obtained F-measures of 53.04%, 69.47%, 69.68% and 72.91% based on LS3RyIE and F-measures of 33.47%, 69.20%, 65.32%, and 35.30% based on Open IE 5 for Event-Event, Event-Time, Event-DCT and Time-Time, respectively. Particularity, *E-T* yielded the highest results while *E-E* obtained the weakest results compared to the others. This confirms that the structure of the *E-E* relation is more complex than other categories because events in such relations can be expressed in several different relation types when they are in similar grammatical forms. For example, e_1-e_2 is presented in the form of 'IS_INCLUDED' in the sentence "*John used*$_{\mathbf{e1}}$ *the phone when driving*$_{\mathbf{e2}}$ *a car*." while e_3-e_4 is presented as 'SIMULTANEOUS' in the sentence "*John died*$_{\mathbf{e3}}$ *when bombing*$_{\mathbf{e4}}$ *the hotel*.". In contrast, other event types such as Time can be expressed with fewer language forms such as a limited number of prepositions or time nouns, e.g., time with preposition 'in, at, and on' or time of day in the week or year. Hence, they are easier to be detected.

Table 9 shows the comparison between our proposed method and baseline approaches (UTTime; TRelPro; Laokulrat et al., 2015; and CATENA). The numbers reported in Table 9 are the results of 5-fold cross-validation evaluation strategy for the baselines because the baselines are *supervised*. The evaluation shows that our proposed method, which is completely unsupervised, shows a competitive performance to the supervised baselines, i.e., better performance on precision and f-measure and competitive on recall on LS3RyIE system. We note that the proposed event network structure is dependent on the performance of the underlying Open IE system and hence in cases when the Open IE system cannot extract event mentions, the corresponding event nodes will not be created in the

Table 10
Details of the Causal-TimeBank and Causal-TimeBank Ext corpora.

Corpus	# docs	Causal-TimeBank		Causal-TimeBank Ext	
		#CLINKs	Mean CLINKs	#CLINKs	Mean CLINKs
ABC	3	9	3	16 (77%)	5.33
AP	2	20	10	48 (140%)	24
APW	11	25	2.27	39 (56%)	3.54
CNN	4	5	1.25	5 (0%)	1.25
EA	1	4	4	5 (25%)	5
NYT	6	22	3.66	30 (36%)	5
PRI	2	3	1.5	4 (33%)	2
SJMN	1	2	2	2 (0%)	2
VOA	2	6	3	6 (0%)	3
WSJ	78	236	3	347 (47%)	4.45
Total	110	332	3.02	502 (51.2%)	4.5

Table 11
Experimental result in Causal-TimeBank and Causal-TimeBank Ext.

Corpus	Open IE 5			LS3RyIE		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Causal-TimeBank						
ABC	87.50	77.77	82.35	77.78	77.78	77.78
AP	54.55	37.50	44.44	71.43	62.50	66.67
APW	88.23	60.00	71.43	77.27	70.83	73.91
CNN	100	33.33	50.00	66.67	66.67	66.67
EA	100	80.00	88.88	100	100	100
NYT	62.5	41.67	50.00	63.63	63.63	63.63
PRI	100	66.66	80.00	66.67	66.67	66.67
SJMN	100	50.00	66.66	100	50.00	66.67
VOA	83.33	83.33	83.33	83.33	83.33	83.33
WSJ	82.14	61.49	70.33	83.88	70.23	76.45
Causal-TimeBank Ext						
ABC	66.67	62.50	64.52	84.61	68.75	75.86
AP	43.33	27.08	33.33	73.91	45.95	56.67
APW	87.50	53.84	66.67	76.67	65.71	70.77
CNN	100	40.00	57.14	80.00	80.00	80.00
EA	100	80.00	88.89	100	100	100
NYT	72.72	42.10	53.33	66.67	63.16	64.87
PRI	75.00	75.00	75.00	75.00	75.00	75.00
SJMN	50.00	50.00	50.00	100	50.00	66.67
VOA	83.33	83.33	83.33	83.33	83.33	83.33
WSJ	79.20	54.57	64.62	79.71	65.26	71.76

event network and hence event relations will be missed. The lower recall of our method with LS3RyIE, noted in Table 7, can be explained on this basis. For example, in our experiments, the Open IE system extracted triples in which only one event is present as shown in Fig. 7. In such cases, it would not be possible to add any edges to the event network given the lack of extracted events. Consequently, our approach will not be able to recognize any additional relations based on such event instances.

Our approach benefits from the relation patterns extracted by Open IE to build the initial event network and bootstraps the temporal event extraction process by determining the type of temporal relation between two directly linked events. The advantage of our proposed work is two folds: 1) it is completely unsupervised and hence does not require any hand-annotated samples and infers indirect temporal relation types between events by systematically traversing the event network, and 2) it works at the document level and not sentence level and hence can identify temporal relations between events that have not been expressed in the same sentence. This is made possible due to the linking of different events in the network whose linking transcends individual sentences and forms a representation of events in the document.

4.2. Causal relation extraction

4.2.1. Experimental dataset

For benchmarking our approach in the second task, we use the Causal-TimeBank and its extended version described in Table 10. Causal-TimeBank has been developed by (Mirza & Tonelli, (2014a)), whose annotations for causality are taken from TempEval-3 corpus, containing 183 documents with 6811 golden events in total. In the Causal-TimeBank dataset, according to (Mirza & Tonelli, (2014a)), a total number of 332 event pairs for CLINK/CLINK-R can be identified from 2519 events available in the original TimeBank corpus. Besides, not all documents contain causality relations between events; from the total number of documents in TimeBank, only 110 (around 60%) of them contain explicit causal links. In Table 8, we report the statistics of causal relations found in the Causal-TimeBank dataset, along with the corresponding numbers of CLINKs associated with them. Regarding its extension (Causal-TimeBank Ext), different from Mirza et al. (2016a) that only uses cause-effect verbs around temporal events, we expand the

Table 12
Performance comparison in Causal-TimeBank and Causal-TimeBank Ext.

	Precision	Recall	F-measure
Causal-TimeBank			
CATENA (Mirza et al., 2016a)	73.70	53.80	62.20
Proposed method (Open IE 5)	79.72	59.45	68.11
Proposed method (LS3RyIE)	81.19	70.16	75.27
Causal-TimeBank Ext			
CATENA (Mirza et al., 2016a)	71.56	49.43	58.47
Proposed method (Open IE 5)	76.47	52.33	62.14
Proposed method (LS3RyIE)	78.54	64.31	70.72

S7: The U.S. maintains that under the U.N. charter, the Kuwaiti request triggers steps**e54** for the collective enforcement of international sanction**e55**.
…
S9: In a statement, the White House said it would do "Whatever is necessary" to ensure compliance**e66** with the sanctions**e67**.

Output:

e55 → e67	IDENTITY
e54 → e55	CLINK
e67 → e66	CLINK
e54 → e66	CLINK
e54 → e67	CLINK

Fig. 8. Identified samples based on transitive inference across multiple sentences.

dataset with new CLINKs based on combining CLINKs and TLINKs in the event network. We note that all expanded causal relations are based on original CLINKs and TLINKs in Causal-TimeBank, which were annotated using the CAT tool (Lenzi, Valentina, Moretti & Sprugnoli, 2012) by two annotators with Dice's coefficient of 0.73. As a result, based on the event network, we annotated an additional 170 causal relation events, which we refer to as Causal-TimeBank Ext. This increased the number of annotated CLINKs by 51.2%. The mean number of CLINKs are 3.02 and 4.5 in Causal-TimeBank and Causal-TimeBank Ext, respectively.

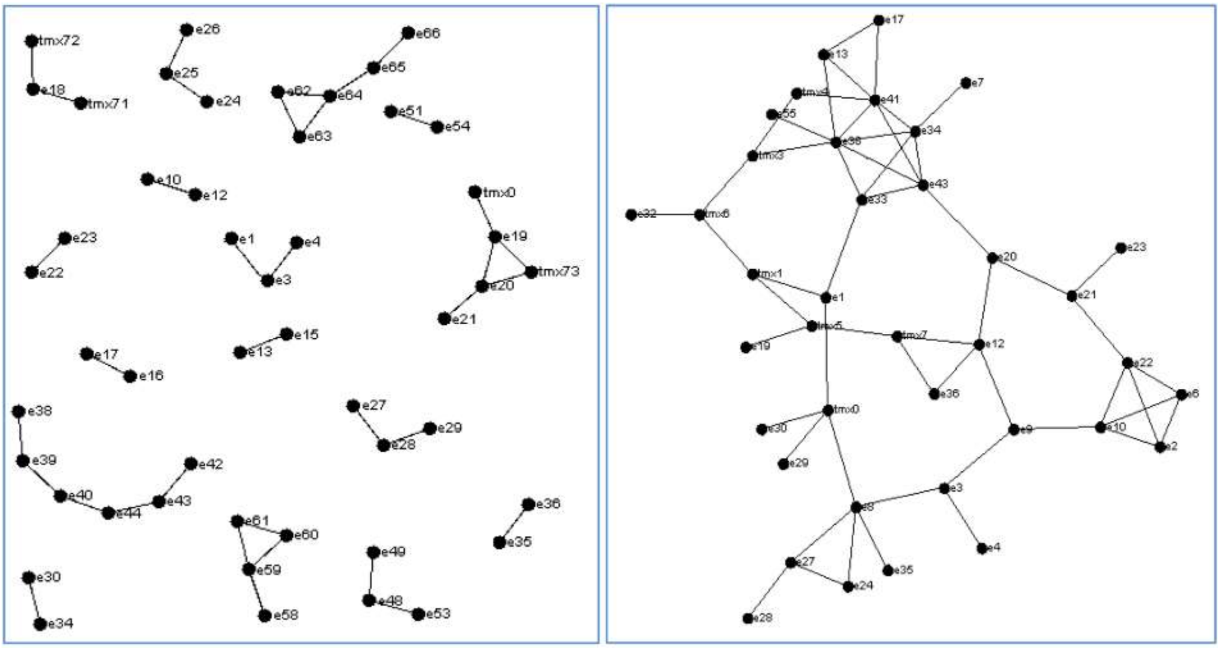
4.2.2. Evaluation results

We first extract patterns from Causal-TimeBank and its extension based on LS3RyIE and Open IE 5 systems to build the event network. Note that, we consider both causal links (CLINKs) and temporal links (TLINKs) on related causal events. As a result of this process, 638 and 930 triples were generated based on LS3RyIE and 592 and 884 triples were generated based on Open IE 5 that were then used to build the event network for both Causal-TimeBank and Causal-TimeBank Ext datasets, respectively. We select top ranking of 30 for casual-association verbs. Note that, we base on mean scale (λ) of score to remove noised verbs if they are larger than 2*λ (Do et al., 2011). Table 11 shows the detailed performance of our proposed method on both datasets. In Causal-TimeBank, our proposed method obtained 76.45% and 73.91% on F-measure for WSJ and APW compared to 66.67% on AP based on LS3RyIE while using Open IE 5, the system obtained 70.33%, 71.43% and 44.44% on F-measure for WSJ, APW and AP. Regarding WSJ, APW and AP in Causal-TimeBank Ext, our approach yielded 71.76%, 70.77% and 56.67% for F-measure based on LS3RyIE while the system obtained 64.62%, 66.67.77% and 33.33% for F-measure using Open IE 5, respectively. In both datasets, it should be noted that WSJ and APW account for a high proportion of the total documents while AP has a more complex event network with a higher number of CLINKs. In Causal-TimeBank, WSJ corpus contains 78 documents with mean CLINKs of 3 and APW has 11 documents with mean CLINKs of 2.27. While AP only has 2 documents with mean CLINKs of 25. Regarding Causal-TimeBank Ext, the mean CLINKs of WSJ is 4.45 and APW has a mean CLINKs of 3.5 while AP has a mean CLINKs of 24. In order to show how the performance of our proposed approach compares to the baseline system, known as CATENA (Mirza et al., 2016a); Table 12 presents a comparison on both Causal-TimeBank and Causal-TimeBank Ext corpora. CATENA exploits event linguistic features such as POS tags, chunking, and dependency paths to train a causal event relation classifier using SVM. Our method outperforms the baseline in both datasets. Our method yields 75.27% and 70.72% on F-measure based on LS3RyIE and 68.11% and 62.14% on F-measures based on Open IE 5 while CATENA obtained 62.20% and 58.47% for F-measure in Causal-TimeBank and Causal-TimeBank Ext, repetitively. Different from CATENA, our method is *unsupervised* and operates at the *document level* and so traverses the event network to identify the causal relations between indirectly connected events while CATENA is a supervised method that relies on sentence-level textual features to identify causal relations.

Fig. 8 shows several extracted relations from a sample document (document wsj900813-0157 from the WSJ corpus). In sentences S_7 and S_9 in this figure, a few temporal and causal event relations can be extracted by using rules based on syntactic and dependency features used in CATENA (Mirza et al., 2016a) such as "e_{55}-IDENTIFY-e_{67}" and "e_{54}-CAUSE-e_{55}", and "e_{67}-CAUSE-e_{66}". However, CATENA fails to correctly determine event relations such as e_{54}-e_{66} and e_{54}-e_{67}. This is primarily due to the fact that events in such relations have no DP connections with each other as they are not present in the same sentence. As such, CATENA cannot extract features pertinent to relations such as (e_{54}, e_{66}) and (e_{54}, e_{67}) and hence fails to identify such relations. In contrast, these relations have been detected in our approach because of the application of transitive inference on the event network, e.g., e_{54}-CAUSE-e_{66} and e_{54}-CAUSE-e_{67} could be determined via e_{54}-CAUSE-e_{55}-IDENTIFY-e_{67} and e_{54}-CAUSE-e_{67}-CAUSE-e_{66}, respectively. As a result, our approach shows up to 13% improvement on F-measure over the baseline.

4.3. Discussion

In our work, we proposed to build an event network based on triples from Open IE systems. We have used LS3RyIE because this system is able to generate high-precision extractions and can be flexibly customized to the underlying application domain compared to other Open IE systems (Vo & Bagheri, 2018). LS3RyIE exploits linguistic knowledge about the grammar of the English language to



(a) (b)

Fig. 9. Examples of event networks: a) A sparse network; b) A complex network.

first detect clauses in an input sentence and to subsequently identify the type of each clause according to the grammatical function of its constituents to extract relations with minimal domain-dependent background knowledge and the least amount of annotated training data. Regarding the overall performance, our system outperforms supervised methods such as CATENA (Laokulrat et al. (2015); Mirza et al., 2016a), TRelPro (Mirza & Tonelli, 2014b), and UTime (Laokulrat et al., 2013). These supervised methods exploit grammatical features such as syntactic and dependency features, but are limited to identifying intra sentence relations with explicit temporal and causal indicators. Our approach does not require any hand-annotated samples for inferring indirect temporal and causal relation types between events by systematically traversing the proposed event network. The network is built based on directly observable temporal and causal relations at the document level. It is possible to predict direct and indirect relations by linking different events in the event network whose linking transcends individual sentences and forms a representation of events across sentences. Hence our method can determine event relations both within the same sentence and across sentences in the whole document.

Now from a performance perspective, we note that the results produced by the event network can depend on the structure of the network including the number of event nodes, the number of relation edges and their connections in the network. Fig. 9 shows two sample event networks extracted from documents APW19980227.0487 and APW19980227.0425 in the APW corpus used in our experiments. As seen in the figure, one event network is quite sparse (Fig. 9a) while the other event network is quite dense (Fig. 9b). We find that when the event network is sparse, our approach is not effective in that it would not have sufficient number of edges in the network to perform inference. As result, our approach cannot identify a sufficient number of relations in such a case, e.g., relations R(e_1, e_{35}) and R(e_1, e_{27}), which should have been extracted, were not identified in Fig. 9a due to missing edges in the event network. On the other hand, the dense event network consists of a higher number of edges connecting the event nodes to each other and hence provides the opportunity for a higher number of inferred causal and temporal relations. For instance, the relation R(e_{33}, e_{41}), which is difficult to extract, was identified by traversing a long path consisting of $e_{33}\rightarrow e_{38}\rightarrow e_{41}$, $e_{33}\rightarrow e_{43}\rightarrow e_{41}$, $e_{33}\rightarrow e_{34}\rightarrow e_{41}$, $e_{33}\rightarrow e_{38}\rightarrow e_{13}\rightarrow e_{41}$. Based on this observation, we believe that our approach faces some limitations: 1) in light of the sparse event networks, our proposed approach is dependent on the performance of the underlying Open IE system and hence in cases when the Open IE system cannot extract event mentions, the corresponding event nodes will not be created in the event network and hence relations will be missed. 2) our method relies on reference mapping to identify similar event nodes in the event graph, which is currently performed through cosine similarity between the textual description of the events. However, more complex text matching methods for event types can be used in the future to improve the reference mapping process that can lead to better overall performance. The source code of our system is available at https://tinyurl.com/yd49et5m.

5. Concluding remarks and future work

In this paper, we have presented an unsupervised method for extracting temporal and causal relations between events by building an event network structure primarily based on information from Open IE systems. The event network is the basis for systematically exploring the possible temporal relations between events by considering how events can be reached from one another. We performed comparative benchmarking of our proposed method using the TempEval-3, Causal-TimeBank, Causal-TimeBank datasets and compared our work against several strong baselines. The results reveal that our method outperforms not only in temporal event extractions but also causal event extraction over other strong baselines.

Our future work will consist of addressing the two limitations of our work, namely quantifying the impact of the performance of the Open IE systems on our work and also exploring more systematic ways for performing reference mapping. We are also interested in extending our work in two other exciting directions in the future:

- The first direction we would like to explore is to see whether other forms of inference on the graph can result in the identification of better causal or temporal relations. For this purpose, we will explore graph distance and traversal methods such as random walks to establish a measure of relevance between events in the network.
- The second direction that we will explore is the application of neural embedding methods (Bagheri, Ensan & Al-Obeidat, 2018) for representing events, and using such representations for building the event network. We will explore how neural embedding-based event representations can be used to enrich the matching between vertices in the event network in order to optimize the weight of the shortest path in the network.

CRediT authorship contribution statement

Duc-Thuan Vo: Conceptualization, Software, Validation, Investigation, Formal analysis, Writing - original draft. **Feras Al-Obeidat:** Funding acquisition, Conceptualization, Validation. **Ebrahim Bagheri:** Funding acquisition, Conceptualization, Supervision, Methodology, Writing - review & editing.

References

Abacha, A. B., & Zweigenbaum, P. (2016). MEANS: A medical question-answering systems combining NLP techniques and semantic Web technologies. *Information Processing & Management, 51*, 570–594.

Abebe, A. M., Tekli, J., Getahun, T. F., Chbeir, R., & Tekli, G. (2019). *Generic metadata representation framework for social-based event detection, description, and linkage. Knowledge-based systems* In press, corrected proof, Available online 9 July 2019.

Ahn, D. (2006). The stages of event extraction. *Proceedings of the workshop on annotating and reasoning about time and events* July 23 - 23, 2006.

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM, 26*(11), 832–843 1983.

Amigo, E., Artiles, J., Li, Q., & Ji, H. (2011). An evaluation framework for aggregated temporal information extraction. *Proceedings of SIGIR workshop on entity-oriented search* July 28, 2011.

Bagheri, E., Ensan, F., & Al-Obeidat, F. (2018). Neural word and entity embeddings for ad hoc retrieval. *Information Processing & Management, 54*, 339–357.

Banko, M., Cafarella, M. J., Soderland, S., Broadhead, M., & Etzioni, O. (2007). Open Information Extraction from the Web. *Proceedings of the 20th international joint conference on artificial intelligence (IJCAI 2007)* (pp. 2670–2676). 06-12 January 2007.

Beamer, B., & Girju, R. (2009). Using a bigram event model to predict causal potential. *Proceedings of the international conference on intelligent text processing and computational linguistics* (pp. 430–441). Springer.

Bethard, S., & Martin, J. H. (2006). Identification of event mentions and their semantic class. *Proceedings of the conference on empirical methods in natural language processing (EMNLP 2006)* (pp. 146–154). July 2006.

Bethard, S., Martin, J. H., & Klingenstein, S. (2007). Timelines from text: Identification of syntactic temporal relations. *Proceedings of the international conference on semantic computing* (pp. 17–19). Sept. 2007.

Chambers, N., Cassidy, T., McDowell, B., & Bethard, S. (2014). Dense event ordering with a multi-pass architecture. *Transactions of the Association for Computational Linguistics, vol.2*, 273–284 2014.

Chambers, N., & Jurafsky, D. (2008a). Jointly combining implicit constraints improves temporal ordering. *Proceedings of the conference on empirical methods for natural language processing (EMNLP 2008)* (pp. 698–706). October 25 - 27, 2008.

Chambers, Nate (2013). NavyTime: event and time ordering from raw text. *Proceedings of the second joint conference on lexical and computational semantics (*SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013)* (pp. 73–77). Association for Computational Linguistics 2013.

Christensen, J., Mausam, Soderland, S., & Etzioni, O. (2011). An Analysis of Open Information Extraction based on Semantic Role Labeling. *International Conference on Knowledge Capture (KCAP)*. Banff, Alberta, Canada: KCAP.

Corro, L. D., & Gemulla, R. (2013). ClausIE: Clause-Based Open Information Extraction. *Proceedings of the 22nd international conference on World Wide Web (WWW 2013)* (pp. 355–366). 13-17 May 2013.

Denis, P., & Muller, P. (2011). Predicting globally coherent temporal structures from texts via endpoint inference and graph decomposition. *Proceedings of the 22nd international joint conference on artificial intelligence (IJCAI)* (pp. 1788–1793). July 16-22, 2011.

Do, X. Q., Chang, Y. S., & Roth, D. (2011). Minimally supervised event causality identification. *Proceedings of the 2011 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 294–303). Association for Computational Linguistics July 27-31, 2011.

D'Souza, J, & Ng, V. (2013). Classifying temporal relations with rich linguistic knowledge. *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: human language technologies (NAACL-HLT 2013)* (pp. 918–927). Association for Computational Linguistics June2013.

Etzioni, O., Fader, A., Christensen, J., Soderland, S., & Mausam (2012). Open Information Extraction. In Proceedings of the Second Generation. *Proceedings of the 22nd international joint conference on Artificial Intelligence (IJCAI 2012)* (pp. 3–10). 16-22 July 2011.

Fader, A., Soderland, S., & Etzioni, O. (2011). Identifying relations for open information extraction. *Proceedings of conference on empirical methods in natural language processing (EMNLP 2011)* (pp. 1035–1545). 27-31 July 2011.

Fan, B., Fan, W., Smith, C., & Garner, H. (2019). Adverse drug event detection and extraction from open data: A deep learning approach. *Information Processing & Management*. https://doi.org/10.1016/j.ipm.2019.102131 2019.

Fei, H., Ren, Y., & Ji, D. (2019). A tree-based neural network model for biomedical event trigger detection. *Information Sciences, 512*, 175–185 February 2020.

Fukui, K., Okada, Y., Satoh, K., & Numao, M. (2019). Cluster sequence mining from event sequence data and its application to damage correlation analysis. *Knowledge-Based Systems, 179*, 136–144 1 September 2019.

Gamallo, P., Garcia, M., & Fernandez-Lanza, S. (2012). Dependency-based Open Information Extraction. *Proceedings of the ROBUS-UNSUP 2012: joint workshop on unsupervised and semi-supervised learning in NLP at the 13th conference of the European chapter of the association for computational linguistics (EACL 2012)*EACL April 23-27, 2012.

Girju, R., Nakov, P., Nastase, V., Szpakowicz, S., Turney, D. P., & Yuret, D. (2009). Classification of semantic relations between nominals. *Language Resources and Evaluation, 43*(2), 105–121 2009.

Grishman, R., Huttunen, S., & Yangarber, R. (2002). Real-time event extraction for infectious disease outbreaks. *Proceedings of the second international conference on*

human language technology research (HLT) March 2002.
Han, R., Hsu, I., Yang, M., Galstyan, A., Weischedel, R., & Peng, N. (2019). Deep structured neural network for event temporal relation extraction. *Proceedings of the SIGNLL conference on computational natural language learning 2019 (CoNLL 2019)* November 3-4, 2019.
Hardy, H., Kanchakouskaya, V., & Strzalkowski, T. (2006). Automatic event classification using surface text features. *Proceedings of the 2006 AAAI workshop on event extraction and synthesis* 2006.
Hashimoto, C., Torisawa, K., Kloetzer, J., Sano, M., Varga, I., Oh, J. H., et al. (2014). Toward future scenario generation: Extracting event causality exploiting semantic relation, context, and association features. *Proceedings of the 52nd annual meeting of the association for computational linguistics (ACL-2014). 1. Proceedings of the 52nd annual meeting of the association for computational linguistics (ACL-2014)* (pp. 987–997). (Volume 1: Long Papers).
Hashimoto, Kazuma, Stenetorp, P., Miwa, M., & Tsuruoka, Y. (2015). Task-oriented learning of word embeddings for semantic relation classification. *Proceedings of the nineteenth conference on computational natural language learning* (pp. 268–278). Association for Computational Linguistics.
Hoang, T. B. N., & Mothe, J. (2018). Location extraction from tweets. *Information Processing & Management, 54*(2), 129–144 March 2018.
Hogenboom, F., Frasincar, F., Kaymak, U., Jong, F. D., & Caron, E. (2016). Survey of event extraction methods from text for decision support system. *Decision Support Systems, 85*, 12–22 May.
Ittoo, Ashwin, & Gosse Bouma, G. (2011). Extracting explicit and implicit causal relations from sparse, domain-specific texts. *Proceedings of the natural language processing and information systems - 16th international conference on applications of natural language to information systems (NLDB 2011)* (pp. 52–63). June 28-30, 2011. Proceedings.
Ji, H., Favre, B., Lin, W. P., Gillick, D., Hakkani-Tur, D., & Grishman, R. (2013). Open-domain multi-document summarization via information extraction: challenges and prospects. *Proceedings of the multi-source, multilingual information extraction and summarization. theory and applications of natural language processing*Springer.
Ji, H., & Grishman, R. (2008). Refining event extraction through cross document inference. *Proceedings of the 46th annual meeting of the association for computational linguistics (ACL 2008)* (pp. 254–262). June 2008.
Junuthula, R., Haghdan, M., Xu, K. S., & Devabhaktuni, V. (2019). The block point process model for continuous-time event based dynamic networks. *Proceedings of the World Wide Web conference (WWW 2019)* (pp. 829–839). May 13-17, 2019.
Kim, H. H., & Kim, Y. H. (2018). Video summarization using event-related potential responses to shot boundaries in real-time video watching. *Journal of the Association for Information Science and Technology, 70* Issue 2, November 2018.
Laokulrat, N., Miwa, M., & Tsuruoka, Y. (2015). Stacking approach to temporal relation classification with temporal inference. *Journal of Natural Language Processing, 22*(3), 171–196 2015.
Laokulrat, N., Miwa, M., Tsuruoka, Y., & Chikayama, Takashi (2013). Uttime: Temporal relation classification using deep syntactic features. *Proceedings of the second joint conference on lexical and computational semantics (*SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013)* (pp. 88–92). Association for Computational Linguistics June.
Laylavi, F., Rajabifard, A., & Kalantari, M. (2017). Event relatedness assessment of Twitter messages for emergency response. *Information Processing & Management, 53*(1), 266–280 January 2017.
Lenzi, B., Valentina, Moretti, G., & Sprugnoli, R. (2012). CAT: The CELCT annotation tool. *Proceedings of the eighth international conference on language resources and evaluation (LREC-2012)* 21-27 May 2012.
Li, H., Ji, H., Deng, H., & Han, J. (2011). Exploiting background information networks to enhance bilingual event extraction through topic modeling. *Proceedings of international conference on advances in information mining and management (IMMM 2011)* 2011.
Li, P., & Mao, K. (2019). Knowledge-oriented convolutional neural network for causal relation extraction from natural language texts. *Expert Systems with Applications, 115*, 512–523 January 2019.
Mani, Inderjeet, Wellner, B., Verhagen, M., & Pustejovsky, J. (2007). Three approaches to learning TLINKs in TimeML. Technical Report CS-07-268, Computer Science Department, Brandeis University. 2007.
Litkowski, K (2014). Pattern Dictionary of English Prepositions. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1274–1283). ACL.
Mausam, Schmitz, M., Bart, R., & Soderland, S. (2012). Open Language Learning for Information Extraction. *Proceedings of the conference on empirical methods in natural language processing (EMNLP 2012)* (pp. 523–534). 12-14 July 2012.
Mirrezaei, S. I., Martins, B., & Cruz, I. F. (2015). The triplex approach for recognizing semantic relations from noun phrases appositions, and adjectives. *Proceedings of the workshop on knowledge discovery and data mining meets linked open data (Know@LOD) co-located with extended semantic web conference (ESWC)* 2015.
Mirza, P.(2016b). Extracting temporal and causal relations between events. Ph.D Dissertation. 2016.
Mirza, P., & Tonelli, S. (2014a). An analysis of causality between events and its relation to temporal information. *Proceedings of COLING, the 25th international conference on computational linguistics: Technical papers* (pp. 2097–2106). August. Dublin City University and Association for Computational Linguistics 2014.
Mirza, P., & Tonelli, Sara (2014b). Classifying temporal relations with simple features. *Proceedings of the 14th conference of the european chapter of the association for computational linguistics (EACL-2014)* (pp. 308–317). Association for Computational Linguistics April2014.
Mirza, P., & Tonelli, S. (2016a). Catena: Causal and temporal relation extraction from natural language texts. *Proceedings of the 26th international conference on computational linguistics (COLING 2016)* (pp. 64–75). Association for Computational Linguistic 2016.
Navarro-Colorado, B., & Saquete, E. (2016). Cross-document event ordering through temporal, lexical and distributional knowledge. *Knowledge-Based Systems, 110*, 244–254 15 October 2016.
Negi, K., Pavuri, A., Patel, L., & Jain, C. (2019). A novel method for drug-adverse event extraction using machine learning. *Informatics in Medicine Unlocked, 2019*https://doi.org/10.1016/j.imu.2019.100190.
Olya, Hessam, M. (2014). Applying Dijkstra's algorithm for general shortest path problem with normal probability distribution arc length. *International Journal of Operational Research, 21*(2), 143–154. https://doi.org/10.1504/IJOR.2014.064541 2014.
Ozdikis, O., Oğuztüzün, H., & Karagoz, P. (2016). Evidential estimation of event locations in microblogs using the Dempster–Shafer theory. *Information Processing & Management, 52*(6), 1227–1246 November 2016.
Pal, H., & Mausam (2016). Demonyms and Compound Relational Nouns in Nominal Open IE. *Workshop on Automated Knowledge Base Construction (AKBC) at NAACL*. San Diego, CA, USA: NAACL.
Peng, H., Song, Y., & Roth, D. (2016). Event Detection and Co-reference with Minimal Supervision. *Proceedings of the 2016 conference on empirical methods in natural language processing (EMNLP-2016)* (pp. 392–402). November 1-5, 2016.
Pustejovsky, J., Castano, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A., et al. (2003a). Timeml: Robust specification of event and temporal expressions in text. *Fifth International Workshop on Computational Semantics (IWCS-5)* 2003.
Pustejovsky, James, Littman, J., Saurí, R., & Verhagen, M. (2006). TimeBank 1.2 Documentation. Technical report, Brandeis University. 2006.
Qian, Y., Deng, X., Ye, Q., Ma, B., & Yuan, H. (2019). On detecting business event from the headlines and leads of massive online news articles. *Information Processing & Management, 56*(6) November 2019.
Riaz, M., & Girju, R. (2013). Toward a better understanding of causality between verbal events: Extraction and analysis of the causal power of verb-verb associations. *Proceedings of the annual meeting of the special interest group on discourse and dialogue (SIGDIAL'13)* (pp. 21–30). 2013.
Rink, B., Bejan, C. A., & Harabagiu, M. S. (2010). Learning textual graph patterns to detect causal event relations. *Proceedings of the twenty-third international florida artificial intelligence research society conference 2010 (FLAIRS 2010)*.
Saha, S., & Mausam (2018). Open Information Extraction from Conjunctive Sentences. *International Conference on Computational Linguistics*. Santa Fe, NM, USA: COLING.
Schmidek, & Barbosa, D. (2014). Improving open relation extraction via sentence re-structuring. *Proceedings of the ninth international conference on language resources and evaluation (ELRA 2014)* May 26-31 2014.
Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland Publishing Company 1970.
Tanev, H., Piskorski, J., & Atkinson, M. (2008). Real-time news event extraction for global crisis monitoring. *Proceedings of the 13th international conference on natural*

*language and information systems: applications of natural language to information systems*LNAI: Springer-Verlag Berlin Heidelberg 2008.
Tatu, Marta, & Srikanth, M. (2008). Experiments with reasoning for temporal relations between events. *Proceedings of the 22nd international conference on computational linguistics (COLING 2008)* (pp. 857–864). COLING 2008 Organizing Committee.
UzZaman, N., Lorens, H., Derczynski, L., Allen, J., Verhagen, M., & Pustejovsky, J. (2013). Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. *Proceedings of the second joint conference on lexical and computational semantics (*SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013)* (pp. 1–9). Association for Computational Linguistics June2013.
UzZaman, Naushad, & Allen, J. (2010). TRIPS and TRIOS System for TempEval-2: Extracting temporal information from text. *Proceedings of the 5th international workshop on semantic evaluation* (pp. 276–283). Association for Computational Linguistics.
Verhagen, M., Sauri, S., Caselli, T., & Pustejovsky, J. (2010). SemEval-2010 task 13: TempEval-2. *Proceedings of the 5th international workshop on semantic evaluation, SemEval '10* (pp. 57–62). Association for Computational Linguistics 2010.
Vo, D. T., & Bagheri, E. (2017). Open information extraction. *Encyclopedia Semantic Computing Robotic Intelligence, 1*(1), https://doi.org/10.1142/S2425038416300032 2017.
Vo, D. T., & Bagheri, E. (2018). Self-training on refined clause patterns for relation ex-traction. *Information Processing and Management, 54*, 686–706 2018.
Wang, Z., Zhang, Y., & Chang, C.-. Y. (2017). Integrating order information and event relation for script event prediction. *Proceedings of the 2017 conference on empirical methods in natural language processing (EMNLP-2017)* (pp. 57–67). September 7–11, 2017.
Wu, F., & Weld, D. S. (2010). Open Information Extraction using Wikipedia. *Proceedings of the 48th annual meeting of the association for computational linguistics (ACL 2010)* (pp. 118–127). 11-16 July 2010.
Ye, H., & Luo, Z. (2019). Deep ranking based cost-sensitive multi-label learning for distant supervision relation extraction. *Information Processing & Management*. https://doi.org/10.1016/j.ipm.2019.102096 2019.
Zhang, H., Boons, F., & Batista-Navarro, R. (2019a). Whose story is it anyway? Automatic extraction of accounts from news articles. *Information Processing & Management, 56*(5), 1837–1848 September 2019.
Zhang, S., & Lv, Q. (2018). Hybrid EGU-based group event participation prediction in event-based social networks. *Knowledge-Based Systems, 143* 1 March 2018.
Zhang, Y., Xu, G., Wang, Y., Liang, X., & Huang, T. (2019b). Empower event detection with bi-directional neural language model. *Knowledge-Based Systems, 167*, 87–97 1 March 2019.
Zhao, Y., Jin, X., & Wang, Y. (2018). Document Embedding Enhanced Event Detection with Hierarchical and Supervised Attention. *Proceedings of the 56th annual meeting of the association for computational linguistics (ACL 2018)* (pp. 414–419). July 15-20, 2018.
Zheng, X., & Suri, A. (2019). Collecting event-related tweets from twitter stream. *Journal of the Association for Information Science and Technology, 70*(2), 176–186 February 2019.
Zhou, G., Qian, L., & Fan, J. (2010). Tree kernel based semantic relation extraction with rich syntactic and semantic information. *Information Sciences, 180*(8), 1313–1325 15 April 2010.
Hu, Z., Rahimtoroghi, E., and Walker, M. (2017). Inference of fine-grained event causality from blogs and films. In Proceedings of the Events and Stories in the News Workshop. Vancouver, Canada, August 2017. Association for Computational Linguistics.
Hu, Z., Walker, M. (2017). Inferring narrative causality between event pairs in films. In Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue. Saarbrücken, Germany. August 2017. SIGDIAL.