

Construction of Event Knowledge Graph based on Semantic Analysis

Yixin SONG

Abstract: At present, the research and application of enterprise credit event information mainly takes the data of enterprise credit events as a dimension of enterprise credit evaluation, and lacks in-depth analysis and mining of the content of special events. On the basis of sorting out the connotation of enterprise credit events, the article firstly proposes a model with evolutionary features, network structured features and unstructured features of text data for the knowledge graph of enterprise credit events; then, the events in enterprise credit events are extracted in the form of case study, the named entities and dependency relationships in the text statements are analyzed, and the events with subject-predicate object relationship as the main form are extracted; secondly, the statements are analyzed. Finally, the extracted events and relationships are matched to form a knowledge graph of corporate credit events. The study applies the mapping research method to the field of corporate credit event research, and realizes the process analysis of the evolution of corporate credit events using knowledge graph.

Keywords: corporate credit events; text data; knowledge graph; semantic analysis

1 INTRODUCTION

Corporate credit crisis can be considered as a social event which is closely related to the public interests. It appears in response to a special event and spreads rapidly within a certain period of time; being affected by the multiplier effect, it may cause stronger social repercussions on the network. If not resolved in time, credit crisis may quickly turn into a collective emergency with both views and conflicting views [1]. In reality, with the rapid development of Internet technology, especially the popularity of various social platforms, the new media environment is getting more open, convenient and diverse, which promotes and adulterates the wide dissemination of various voices of consensus. And a good consensus environment can provide sustainable and high-quality development of ideological guarantee for enterprises [2]. Although credit event crisis is a new form the enterprises face in network environment, if we can take effective measures, it is possible to turn crisis into opportunity, not only improve the traumatic corporate image, but also reshape a better corporate image. So, it can make the credit event crisis become a driving force of corporate image building and a catalyst of corporate visibility improvement.

By analyzing the evolution process of many corporate credit event cases and summarizing the evolution rule of crisis which occurred in the business process, it can provide decision support for the business management of enterprises; meanwhile, the monitoring system of corporate credit events can also help the society monitor the changes of enterprise operation in time, guarantee health and safety of the network environment, and prevent the deterioration of enterprise operation from causing bad social impact [3, 4]. In the background of the network era, it is necessary to establish a mechanism of rapid response to credit events and to improve corporate credit monitoring methods. Enterprises should learn to use data to process financial disclosure information and build a corresponding corporate credit risk response system. When encountering a network crisis, it is needed to establish friendly network relationships, when the enterprise is in a consumer trust crisis, it is needed to establish an effective strategy to repairing trust [5-7]. With events, relationships between events, and then representation in an organizational system, an event knowledge base can be constructed and a matter mapping can be formed. Therefore, by sorting out and

visualizing the corporate crisis events and events relationships behind specific corporate credit events on the Internet [8, 9], using natural language processing technology to semantically analyze the corporate credit event texts, structuring the cause-effect relationships among corporate credit events, and demonstrating the evolution process of corporate credit in the form of visualization, this article aims to summarize the evolution law of corporate credit.

The rest of the paper is structured as follows: constructing a research framework for corporate credit event knowledge graph based on basic theory; extracting the events of financial data based on semantic analysis; causally extracting events based on cause-effect relationships; forming a corporate credit event knowledge graph; and finally, a research summary.

2 ENTERPRISE CREDIT EVENT METHOD DESIGN

2.1 Knowledge Graph Analysis Framework

A corporate credit event describes the beginning, evolution and outcome of corporate crisis event, which consists of one or more corporate crisis events. If the corporate crisis event continues to develop, it will not only have a negative impact on social reputation and credibility of enterprise, but seriously endanger the survival of the enterprise. Knowledge graph is the description and display of a specific event of an enterprise. Knowledge graph of enterprise credit events refers to the news of enterprise credit events occurring within a period of time, extracting the events and relationships from them, mining the veins of the development of credit-related events and displaying the data.

The construction of corporate credit event knowledge graph mainly includes two parts: event extraction and extraction of event relationship. Event extraction mainly refers to extracting the behavior of enterprises in corporate credit events. The enterprise credit events mainly come from different channels and belong to text type data, which needs mining technology to turn the unstructured data into structured data. Therefore, this paper adopts dependency parsing analysis and named entity recognition technology to extract topics and events with subject-predicate relationship respectively. Relationship extraction mainly refers to extracting cause-effect relationships in corporate

credit events by using cue words and different cause-effect patterns.

Because causal relationships may consist of several events, the extracted events need to match with relationships to form a knowledge graph based on public opinion, which is clarified in terms of connotation and extension, as well as the interaction with other public opinions, online and offline public opinions, print media, airwave media (Fig. 1).

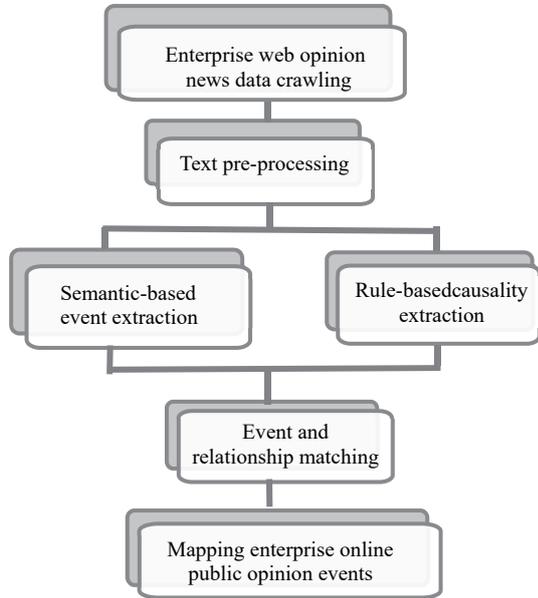


Figure 1 The construction process of enterprise network public opinion event graph

2.2 Data Preparation

This study collects negative credit event information from enterprises which have significant credit event events and have financial operation difficulties, default, break the contract and other major breach of trust in daily operations. The corporate credit event data are collected using python, and the sources of collected credit event information include government websites, news websites, industry websites and other formal news media websites, shielding data sources such as WeChat and Weibo, which are mainly personal and self-published.

[/w, ofo/nx, why/ryv, repeatedly/d, stuck into/vi, capital chain/nz, crisis/n, rumors that/n, ?/w, Dai Wei/nr, still/d, can/v, "/w, arbitrary/a, "/w, how long/ryt, ?/w,]/w, since/p, ofo/nx, refuse/v, DiDi/q, of/ude1, acquisition/v, offer/v, then/f, , /w, about/vn, ofo/nx, of/ude1, negative/b, news/n, has/d, time to time/d, emerged/v, o /w, since/p, Mo/b, bai/v, being/pbei, Mei/b, tuan/n, acquired/v, , /w, Haro/nz, bicycle/n, into/v, Ali/nt, arms/n, , /w, bent on/d, maintaining/v, /a, status/n, its/ude1, ofo/nx, , /w, how/ry, can/n, avoid/v, repeated/d, being caught in/v, capital chain/nz, crisis/n, of/ude1, rumors/n, ?/w

Most of the words obtained in the above-mentioned word separation stage are useless and have little impact on the analysis. Therefore, in this paper, these words are defined as deactivated words, including orientation words, prepositions, quantifiers, auxiliary words, punctuation, non-semantic words and tone words, such as "of, ah, this, if, that, so", relying on the existing deactivation thesaurus of Harbin

Institute of Technology, the deactivation thesaurus of Sichuan University Machine Learning Intelligence laboratory, and the deactivation word list of Baidu, and other deactivation thesaurus. This paper integrates the words into a more comprehensive word list, with 1598 deactivated words.

2.3 Text De-Duplication

Since these media platforms often reproduce the same articles, there are many duplicate contents in the obtained financial text information, which will affect the accuracy of the text analysis and cause large deviations in the analysis results, such as the frequency of "fake" to improve the occurrence of core words. Therefore, de-duplication is required before text analysis. The process of text de-duplication includes: generating a list of subwords; building a dictionary based on the text set to obtain the number of features; building a corpus based on the dictionary to generate a sparse vector of search words; training a TF-IDF model with corpus to calculate the similarity between texts (Fig. 2), and removing the texts with high similarity.

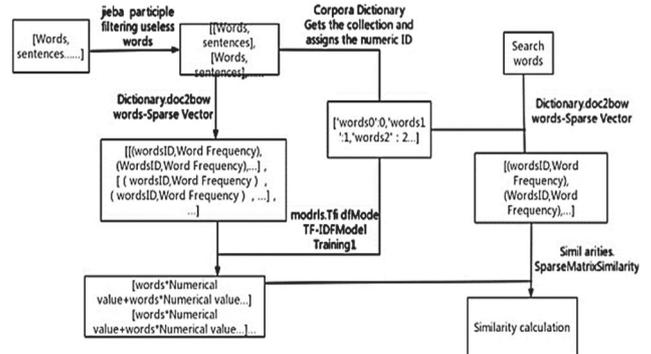


Figure 2 Text de-duplication algorithm flow

- Generating a list of subwords;
- Building a dictionary based on the text set and obtaining the number of features;
- Building a corpus based on the dictionary and generating a sparse vector of search terms;
- Training TF-IDF models with the corpus;
- Calculate the similarity between texts and remove the similarity between texts and remove the texts with high similarity.

Taking the financial text data of OFO as an example:

"Miss that old forest [original microblogging June 11 news, recently a media report said ofo currently owes 1.5 billion yuan, deposit."

The balance is only about 3.5 billion yuan, with less than 500 million yuan of cash available on the books. It is understood that ofo founder and CEO Dai Wei has publicly stated that ofo already has 200 million users. ofo started with a deposit of 99 yuan, and in June 2017 increased the deposit for new users to 199 yuan. If we follow ofo's official claim of 200 million users, each user deposit at 99 yuan before the price increase and the cumulative free deposit of nearly 3 billion."

"The little yellow car is really good, quite like riding it, hope to get through it! [Original Weibo] [Ofo's darkest

hour, how long can Dai Wei last? [eat melon]]The high maintenance costs, operating costs and fierce competition have put the shared bicycle in a profitable predicament, [meow] ofo, which insists on independent development, has recently been exposed to negative news of a tight funding chain and a large number of layoffs, [doge] although ofo has denied it, but the outside world believes that ofo has been under pressure for a long time, how long can Dai Wei hold out?"

The article can obtain the number of de-duplicated texts by generating subwords, building dictionaries, forming sparse vectors and calculating the similarity of texts (Tab. 1).

Table 1 Example of result

Example sentence	Subword results
sentence 1	miss, that, old forest, \, 【, original Weibo, 】 , , June, 11th, news, , , recently, , a media report, claims, ofo, currently, owes, 1.5, billion yuan, , , deposit, balance, only remains , about, 3.5, billion yuan, , , carrying amount, can, available, cash, less than, 500, million yuan, . , as, it is understood that, , , ofo, founder, and, CEO, DaiWei, has, , said that, , , ofo, has, had, 200, million, users, . , ofo, started with, deposit, of, 99, yuan, , , in , June, 2017, increased, the deposit, for, new users, to, 199, yuan, . , if, follow, ofo's, official, claim, of, 200, million, users, \ , each, user, deposit, at, 99, yuan, before, price, increase, and, cumulative, free, deposit, nearly , 30,
Sentence 2	little, yellowbicycle, isreally, good, , , quite, like, riding, it, , , hope to, getthrough, it, ! , , 【, Original Weibo, 】 , , 【, ofo's, darkest , hour, , , How, long, can, DaiWei, last, ? , [, eat melon,] , 】 , high, maintenance, costs , \ , operating , costs , and , fierce , competition , haveput , shared , bicycle , in , aprofitable, predicament, , , [, meow,meow,] , insists on, independent , development, , ofo, has, recently, been, exposedto, funding, chain, tight, , , large number of, layoffs, of, negative , news, , , [, doge,] , although, ofo, has, denied it, , , but, the outside world, believes, ofo, hasbeen, underpressure, for a long time, , , how, long, can, DaiWei, hold out, ? , http , : , / , / , t , . , cn, /
Sentence 3	claim, , , that, rumors of, DiDi, restarts , acquisition, ofo, is, not, true, , , ofo, will, maintain, long-term, independent, development, . , in, statement, , , ofo, small, yellow bicycle , also , mentioned , its , future , development, goals, , , that, by, 2025, , , ofo, will, provide, travel, service, to, 20, , billion, users, worldwide, . , , ofo's, the last stand, , is, price

secondly, pruning the dependency tree to remove the non-conforming dependencies and generate a dependency subtree; finally, selecting the subject-verb and verb-object relations with the named entity as the core, and obtain the subject-verb and verb-object pairs.

Based on this theory of syntactic parsing, this paper represents the dependency tree as $T = (V, A, R)$: Here V denotes the set of nodes, specifically the words in the utterance; A is the set of directed arcs, which refers to the inter-word dependency relations, where the starting point of the arc is the dominant word of the dependency relation and the ending point of the arc is the dominated word of the dependency relation; R is the root node of the dependency tree, which refers to the core verb of the word, and T meets the following three conditions: (1) The entry degree of the R node is 0. (2) the entry degree of nodes other than R is 1; (3) there is a directed path from R to any node. The dependency relations analysis results of the statement are obtained by performing dependency analysis on the example sentence (Fig. 3).

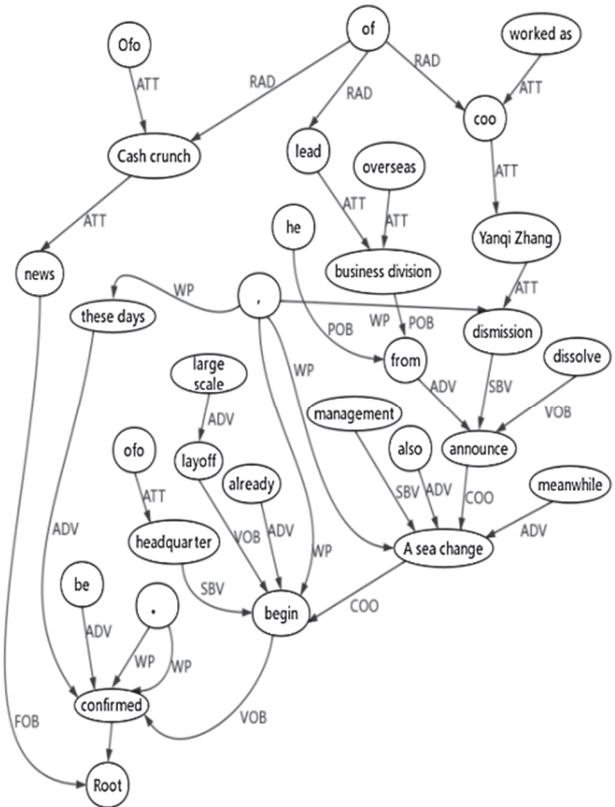


Figure 3 Text dependency tree

3.2 Triadic Event Extraction

This study focuses on identifying the names of people and organizations, and the named entities were identified for the example sentences in the previous section to obtain the results shown below.

Named entity recognition results

Recently /O, /O the news /O that /O Xiao /O Yellow Car /O is /O financially /O tight /O was /O confirmed /O. /O ofo /S-Ni headquarters /O has /O started /O massive /O layoffs /O, /O while /O management /O also /O underwent /O drastic changes /O, /O Zhang Yanqi /S-Nh, who /O was /O COO /O left /O, /O by /O him /O The /O overseas /O division /O led by /O announced /O the dissolution /O of /O. /O

From this sentence, the name of an organization "ofo" and the name of a person "Zhang Yanqi" can be identified. In this paper, the word segmentation is also modified in a certain way in syntactic annotation. The process is to use the established entity dictionary to syntactically annotate the whole. This approach not only allows the syntactic tree structure to be simplified, but also allows the next obtained triad of dependencies, with some complete semantic information. Suppose that the document, $D = \{s_1, s_2, \dots, s_n\}$, s represents a sentence in the document. For any one of these sentences s_i , after segmentation and deactivation there is $s_i = \{w_1, w_2, \dots, w_m\}$, here, w represents the words in the sentence. After the dependent syntactic analysis of the sentence yields the graph $G = (s_i, E_i)$, here, $G = (s_i, E_i)$ represents the set of all edges in the graph. This paper is a description of the features in the sentence using the relational structured triad of the two words connected by an edge in the graph corresponding to this edge. That is, if an edge in the graph is e_{ij} , the e_{ij} two ends of the word w_i, w_j , the edge e_{ij} indicates the dependency relationship as r_{ij} , here, w_i denoted by the dominant word, w_j denoted by the dependent word, the ternary relationship can be described as (w_i, w_j, r_{ij}) .

Table 2 The part of results

Institution: ofo	Person: Dai Wei	Address: Ideal International Building
ofo, repeatedly caught in, funding chain crisis rumours ofo, rejects, drip acquisition ofo, maintain, independent status ofo, split up, lacking ofo, suspected, layoffs ofo, start, layoffs ofo, refund, deposit ofo, make, massivelayoffs ofo, shortage, of, money ofo, conducts, business exploration ofo, insist, atough, independent, development path ofo, mortgage, 4.45 million bikes in first-tier cities ofo, responds, to journalists who say the company is experiencing organised smears ofo, started, launching the sale of body advertising ofo, founded in, 2014 ofo, offers, bike-sharing service ofo, hit, ofo's third anniversary birth poster ofo, average daily active users, reached, 5,748,400 ofo, carries out, body advertising ofo, has, 200 million users	Davy, refuses, merger Dai Wei, states, willing to fight Dai Wei, is, a lonely dream stickler Davy, Rejects, Merger Offer Dawei, loss, good cash opportunity Dai Wei, President of the Students Union in Beijing University Dai Wei, in debt, 6 million yuan Dai Wei, who joined the blockchain project Dai Wei, rather than,offended by Zhu Xiaohu, an early investor in ofo	ofo, located at, ideal international building

According to the definition of the ternary dependency group, forming the example sentence dependency display, a

total of 15 ternary relationship groups can be constructed. From these, the better dependency triads are selected, and the following principles are used in this paper: 1) select triads with strong dependency relations; 2) give priority to dependency triads containing entity words. In this paper, we focus only on the relationships between verbs and nouns in sentences, and use these dependency pairs to mine user intentions. A total of subject-verb-object relationship events exist in the financial text data, and some of the results formed are shown in Tab. 2.

4 RULE-BASED CAUSALITY EXTRACTION

The extraction of cause and effect relationships mainly consists of the extraction of cause and effect clauses as well as event tuples from the clauses. The cause and effect clauses are extracted according to syntactic relationships and matching rules, and the two types of financial text data are stored in a structured manner. The extraction of causal relations consists of two tasks: 1) the identification of causal clauses and 2) the extraction of causal relations. This paper focuses on summarizing the forms of causal relations corresponding to each type of cue word by using causal cue words, including conjunctions, verbs, prepositions and adverbs, based on lexical and grammatical features.

Let W be a collection of words, then the collection $S = \{w_1, w_2, \dots, w_n\}$ can be represented as a sentence S , and S after lexical annotation the collection $S' = \{w_1/d_1, w_2/d_2, \dots, w_i/d_i, w_n/d_n\}$ can be obtained, where w_i is the lexical nature of $d_i, w_i \in W, i = 1, 2, \dots, n$. According to the causal syntactic model, the rules for judging causal sentences are clarified (Tab. 3).

Table 3 Causal extraction rule

Cause and effect structure	Expressions
Rule 1: From cause to effect monofocal	if $w_i \in Cue_{c_1}$ and $d_i = c$, so $Cause = \{w_2 w_3 \dots w_i\}, Effect = \{w_{i+1} \dots w_n\}$
Rule 2: Cause to effect centering	if $w_i \in Cue_{c_2}$ and $d_i = c$, so $Cause = \{w_1 w_2 \dots w_{i-1}\}, Effect = \{w_{i+1} \dots w_n\}$
Rule 3: from effect to cause centered	if $w_i \in Cue_{c_3}$ and $d_i = c$ $Effect = \{w_2 w_2 \dots w_{i-1}\}, Cause = \{w_{i+1} \dots w_n\}$
Rule 4: from cause to effect matching	if $(w_i \in Cue_{c_2}$ and $d_i = c, w_i \in Cue_{d_2}$ and $d_i = d) \cdot (w_i \in Cue_{c_4}$ and $d_i = c, w_i \in Cue_{c_5}$ and $d_i = d)$ so $Cause = \{w_2 w_3 \dots w_{i-1}\}, Effect = \{w_{i+1} \dots w_n\}$
Rule 5: from effect to cause matching	if $w_i \in Cue_{c_6}$ and $d_i = c, w_i \in Cue_{d_7}$ and $d_i = c$ so $Effect = \{w_2 w_2 \dots w_{i-1}\}, Cause = \{w_{i+1} \dots w_n\}$

Transformation of causal extraction rules into regular expressions for causal extraction (Tab. 4).

In this chapter, four steps are designed to carry out the extraction of event tuples: the first step performs lexical filtering. The words closest to the position of the causal cue word and of verbal nature are taken as the trigger words; the second step performs dependent syntactic analysis. The main task is the subject and the object corresponding to the trigger word; the third step is component filtering. The related components of the trigger word, subject and object are extracted respectively; the fourth step determines the

event tuple representation. The format {subject and its related components, trigger word and its related components, object and its related components} is used.

Table 4 Causal extraction of regular expressions

Causal structure	Regular expressions
RULE 1: from cause to effect monofocal	word_pairs = r^\s?(for based according because caused by dependent byvirtue dueto since)/[p c]+\s(.*)'[. ,]+(.*)'
RULE 2: from cause to effect centering	word_pairs = r'(.*)'[. ,]+.*(still thus so caused leadto so that? therefore so that? so so how has only relied on on this)/[p c]+\s(.*)'
RULE 3: from effect to cause centered	word_pairs = r^\s?(since? because becauseof if dueto as long as)/[p c]+\s(.*)'[. ,]+(.*)'
RULE 4: from cause to effect matching	word_pairs = [['because', 'thus'], ['because', 'for this reason'], ['since?', 'so'], ['because', 'for this reason'], ['because', 'for this reason'], ['unless', 'only']] r^\s?(%s)/[p c]+\s(.*)(%s)/[p c]+\s(.*)'
RULE 5: from effect to cause matching	word_pairs=[['why?the reason', 'because'], ['why?the reason', 'due to']] r^\s?(%s)/[p c]+\s(.*)(%s)/[p c]+\s(.*)'

The causal extraction rules are transformed into regular expressions for causal extraction, which ultimately leads to Tab. 5 by pattern matching methods.

Table 5 The examples for cause and effect extraction

Cause and effect structure	Cause and effect example sentences
RULE 1	Cause: Strained capital chain Tag: Due to Effect: Mass layoffs have begun at the head office, while senior management changes have been dramatic
RULE 2	Cause: problems with ofo's operations Tag: Causes Effect: It needs to give up some of the things, ofo has chosen to give up overseas markets, in July has withdrawn from overseas small cities, Australia, Germany and other cities will be difficult to see the small yellow car shadow for a while
RULE 3	Cause: failed to get its investment in OFO as hoped for with limited funds Tag: Because Effect: Ali's stake was intended to gain more voting rights but failed to do so
RULE 4	Cause: There is a limited amount of bicycle sharing in each city. Tag: because - so Effect: After Ali harvests the remaining assets of ofo, then these shares will be Ali's as well.
RULE 5	Cause: Drip has not launched a formal takeover bid Tag: Why - Because Effect: Rejected Drip's offer

5 ENTERPRISE CREDIT EVENT KNOWLEDGE GRAPH GENERATION

Event knowledge mapping focuses on dynamic events and the cascading, temporal and causal relationships among them, and represents them in the form of structured graphs for more efficient management of massive data. In particular, the mining of dynamic event information and logical event relationships. Corporate credit event group $\{e_1, e_2, \dots, e_i, \dots, e_n\}$, e_i represents the i event extracted from the

financial text of this enterprise credit event, the causal event group of enterprise credit event $\{[C_1, E_1], [C_2, E_2], \dots, [C_i, E_i], \dots, [C_n, E_n]\}$, $[C_i, E_i]$ represents the causal event in the enterprise credit event, C_i represents the cause event and E_i represents the effect event. Determine if e_i is in C_i or E_i . If $e_i \in C_i$, then e_i is a causal event; if $e_i \in E_i$ then e_i is a fruitful event; if $e_i \notin C_i$ and $e_i \notin E_i$, e_i is a parallel event. Construct causal event pairs by judging the category to which the group of events belongs. e.g. if $e_i = a \in C_i$, $e_i = a$ is a causal event; $e_i = b \in E_i$, $e_i = b$ is a fruitful event, then consider $e_i = a$ and $e_i = b$ to be a pair of causal event pairs and record as $(e_i = a, e_i = b)$. The above example is a one-cause-one-effect event pair. There are also cases of one cause and many effects, multiple causes and effects, and multiple causes and effects, respectively, expressed as

- (1) One cause, many effects: $(e_i = a, (e_i = b, e_i = c))$.
- (2) Multiple causes and effects: $((e_i = a, e_i = b), e_i = c)$.
- (3) multiple causes and effects: $((e_i = a, e_i = b), (e_i = c, e_i = d))$.

The events in the case were first extracted to obtain the event (Tab. 6).

Table 6 The events list extracted from the cases

Sequence	Type of relationship	Event
1	Subject-predicate object relationship	(government, ban, put in new cars)
2	Subject-predicate object relationship	(vehicle, (un)get, replace)
3	Subject-predicate object relationship	(market, none, car)
4	Subject-predicate object relationship	(to ban, to be, to block the back roads of shared bicycles)
5	Subject-predicate object relationship	(rivals, are, the two main giants of Mobiles)
6	Subject-predicate object relationship	(they, fight, price war)
7	Subject-predicate relationship	(customer, churn)
8	Subject-predicate object relationship	(they, choice, new round of price wars)
9	Postpositional verb-object relationship	(war, burn, money)
10	Subject-predicate object relationship	(ofo, attract, user)
11	Subject-predicate object relationship	(operate, is, key to burn money)
12	Subject-predicate object relationship	(businessman, choice, withdrawal)
13	Subject-predicate object relationship	(small yellow car, experience, heyday sharing period)
14	Subject-predicate object relationship	(ofo, movable, deposit)
15	Subject-predicate relationship	(refund of deposit, failure)
16	Subject-predicate object relationship	(ofo, to stay up, this winter)
17	Subject-predicate object relationship	(heart, fear, is cool)

A total of 16 events were extracted, of which 13 were subject-verb-object events, 2 were subject-verb events, and 1 was a stative-verb-object event. The cause-effect relations of the cases were then extracted to obtain the cause-effect events of the cases (Tab. 7).

The event pairs were matched with the cause and effect in the causal relationship to obtain the causal event pairs for the disclosed financial data, as shown in Tab. 8.

Table 7 The causal relationship of case extraction

Sequence	Cause and effect relationship Event
1	A large number of previously damaged vehicles could not be replaced after the Cause government banned new bikes from the market. tag effect no bikes are available in the market.
2	cause ofo's competitors are mainly the two main giants, Mobay and Haro, and they are constantly fighting price wars for customers, which results in mutual targeting of customers tag penetration effect A portion of customers are lost.
3	cause To win back customers, they opt for a new round of price wars tag In order to effect The cycle repeats itself, creating a vicious circle.
4	cause To win back customers, they opt for a new round of price wars tag In order to effect The cycle repeats itself, creating a vicious circle. cause although it did drag down many other bikes, but also caused ofo and mobi's own internal conflict, although in the market to take the initiative, but the later operation is still the key to burn money tag therefore effect has led to many capital businessmen choosing to withdraw their shares.
5	cause All sorts of negative news follows, ofo touches bottom line with deposit tag The effect of the failure to refund the deposit has caused a lot of criticism from users.

Table 8 A causal event pair extracted from a causal relationship

Serial number	cause event	fruitful event
1	(government, banned, put in new vehicles) (vehicles, (un)available, replace)	(market, none, car)
2	(rivals, are, the two main giants of Mobay) (they, fight, price wars)	(customer, churn)
3	(operate, is, key to burn money)	(businessman, choice, withdrawal)
4	(of, movable, deposit)	(refund of deposit, failure)

By extracting events and causality matching for ofo credit events, a total of 22 valid causal event pairs were obtained, and the causal pairs were put in temporal order to generate the knowledge graph shown in Tab. 9.

Table 9 ofo extracts causal event pairs

Serial number	cause event	fruitful event
1	Expansion leads to burnout	The runaway funds
2	High cost of market expansion	Cost poses a quality hazard
3	of's sloppy management	Employee exposure to corruption
3	Deposit refunds	of is short of money
4	The opponents are the two main giants of Mobay They are in a price war	Loss of customers
5	Operations is the key to burning money	Businessmen opting out of shares
6	of is mired in constant negative news	Deposit refunds
7	of misappropriates user deposits Money expansion leads to	Deposit refunds fail Sources say of is diverting deposits for supply chain

Serial number	cause event	fruitful event
	money burning	debts
8	of rejects Drip's acquisition	of is mired in constant negative news
9	of has a crisis of broken capital chain	of starts laying off staff Zhang Yanqi leaves Huge change in management Company exits overseas operations of owes hundreds of millions of yuan to a number of logistics suppliers of Cloudbird of misappropriates user deposits
10	Dai Wei has a veto	of lost to a veto
11	Refund of deposit failed	of has a credibility crisis
12	Dai Wei refused to merge	Drip replaces ofo drip acquires ofo rival small blue bicycle
13	of defaulted on payment	Shanghai Phoenix sues ofo
14	of misappropriated users' deposits of defaulted on payment of goods	Dai Wei became a deadbeat
15	of awareness of quality issues leading to a decline in service	Word of mouth declining
16	Drip's flurry of moves is nothing less than a bid for de facto control of ofo Drip sent a number of senior executives	Dai Wei refused to merge
17	Veto blocking ofo financing	of has a crisis of broken capital chain
18	Drip denies Dai Wei's multiple decisions The drops refused to sign	Veto blocking ofo financing
19	Dai Wei rejected the proposal of the management of the drip system	The drops refused to sign
20	The result is ofo cost compression	Cost poses a quality hazard
21	Staff exposure to corruption	of releases internal anti-corruption document of investigates and handles several cases of corruption
22	The government banned the launch of new cars	Market, none, car

Corporate credit event evolution analysis is to build a corporate credit event knowledge map to show the whole process of corporate crisis event development according to the chronology of events and the cause-and-effect relationship between events. Taking ofo as an example, from the knowledge map of ofo's corporate credit events, it can be analyzed that ofo's crisis events evolved through three stages as follows.

(1) From March 2017 to June 2017, due to the fierce competition in the bike-sharing market, ofo and Mobay, which are the industry giants, began to engage in a price game in order to compete for users. In order to expand its market scope, ofo increased the number of bicycles by controlling the cost of bicycles, resulting in a poor riding experience for users and a decline in word-of-mouth. At the same time, limited management capacity resulted in the occurrence of corrupt practices among employees within the company.

(2) From July 2017 to April 2018, due to the

preliminary price game and market expansion, it caused a shortage of funds for ofo and a large-scale financing was conducted to replenish the funds. Drip entered into ofo as an investor, and a fight for control with Dai Wei, the founder of ofo, occurred. On the one hand, Dai Wei rejected the acquisition by Drip, and on the other hand, Drip also blocked other financing for ofo through its veto, further contributing to ofo's funding shortage.

(3) From May 2018 to March 2019, ofo's funding chain broke, leading to a series of problems, including the misappropriation of user deposits to repay suppliers, layoffs, senior management departures and business shrinkage, and the founder, Dai Wei, was also subject to complaints over unpaid debts and was eventually listed as a "rogue".

It can also be seen from Tab. 8 that some of the causal events are not in chronological order, as some of the crisis events are hidden and not immediately revealed, it appears first in the credit events and in the subsequent credit events through media digging.

6 CONCLUSIONS

There is a certain degree of unpredictability and uncertainty in corporate crises, and as the economic environment is constantly updated, enterprises will continue to face new risks and challenges, and at the same time, new samples of corporate crises are emerging. In the context of the Internet, negative credit events have become an essential source of information for identifying abnormal business conditions and play an important role in the early warning process of corporate crises. However, at present, China's corporate credit management is not perfect, and some enterprises, in order to evade policy restrictions in order to achieve such purposes as illegal operation, capital expansion and stock price manipulation, enterprises usually adopt a cover-up or misleading approach to relevant negative news in order to avoid the management and the public from understanding and being informed of the real business situation of the enterprise.

The article collects enterprise credit event news data through crawlers and performs text pre-processing such as word separation and de-duplication on the disclosed financial texts to form a corpus for text analysis; secondly, it performs text dependency analysis and named entity recognition through semantic analysis techniques to extract the event triad in the disclosed financial text data; then it constructs event causality rules to extract the causality relationships in the financial texts; finally, it matches the events matched with the causal relationships to generate a knowledge graph of corporate credit events, and the evolution process of corporate credit events is analyzed and displayed through the graph.

There is quite a lot of information that will not be spread on the network. In the future, we will collect more complete data, such as WeChat microblogs, to make the model and results more scientific.

7 REFERENCES

[1] Gang, L. & Dong, Q. (2011). Analysis and research on the communication process of enterprise internet public opinion based on Web2.0. *Information Science*, 24(24), 95003-95007.

- [2] Yang, C., Shi, X., & Wei, C. (2009). Discovering event evolution graphs from news corpora. *IEEE Transactions on Systems, Man, and Cybernetics*, 39(4), 850-863. <https://doi.org/10.1109/TSMCA.2009.2015885>
- [3] Zhang, P., Yang, J., & Shi, S. (2019). Public Opinion Situational Awareness in Cyberspace. *Journal of Information Security Research*, 5(11), 1013-1020.
- [4] Zhang, L. & Zhang, N. (2020). Effectiveness of Trust Repair Strategies in the Crisis of Corporate Internet Public Opinion. *American Journal of Management Science and Engineering*, 5(1), 10-18. <https://doi.org/10.11648/j.ajmse.20200501.12>
- [5] Rim, H., Park, Y., & Song, D. (2020). Watch out when expectancy is violated: An experiment of inconsistent CSR message cueing. *Journal of Marketing Communications*, 26(4), 343-361. <https://doi.org/10.1080/13527266.2018.1523216>
- [6] Liu, D., Su, J., Song, L. et al. (2019). Application of Internet segmentation research based on Natural Language Processing technology in enterprise public opinion risk monitoring. *Journal of Medical Internet Research*, 1187(4), 042007. <https://doi.org/10.1088/1742-6596/1187/4/042007>
- [7] Wang, J. & Wang, X. (2019). Research on the Enterprise Public Opinion Spreading Model and Control Strategy in Two-layer Social Networks. *Journal of Managementence*.
- [8] Li, Y., Xu, M., Wen, X., & Guo, D. (2020). The Role of Internet Search Index for Tourist Volume Prediction Based on GDFM Model. *Tehnicki vjesnik - Technical Gazette*, 27(2), 576-582. <https://doi.org/10.17559/TV-20191231071057>
- [9] Tosun, S., Öztürk, A., & Taşpınar, F. (2019). Short Term Load Forecasting for Turkey Energy Distribution System with Artificial Neural Networks. *Tehnicki vjesnik - Technical Gazette*, 26(6), 1545-1553. <https://doi.org/10.17559/TV-20180814115917>

Contact information:

Yixin SONG

Accounting department,
Zhongnan University of Economics and Law,
182# Nanhu Avenue, East Lake High-tech Development Zone,
Wuhan, China
E-mail: 2295195085@qq.com