Postprint: Research on Strategic Intent of Corporate Mergers and Acquisitions Based on Entity Relationships

Authors: Liao Jianjun

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose / Significance] To address the challenge of acquiring corporate strategic intelligence in competitive intelligence, this paper proposes an automated and efficient strategic intent mining process to facilitate the collection of strategic intelligence information and enhance the effectiveness of strategic intent mining. [Method / Process] This study introduces entity relation extraction technology to explore a systematic solution for acquiring corporate strategic intelligence. Taking corporate merger and acquisition (M&A) events as an example, relation extraction technology is applied to collect corporate M&A relationships, entity attribute extraction technology is employed to reveal the strategic intent behind corporate M&A, and the effectiveness of the extraction system is validated through the M&A events of Alibaba Group. [Result / Conclusion] The proposed solution can effectively and automatically extract corporate M&A relationships from massive web information and identify business operation domains, revealing the strategic intent of corporate M&A through visual presentation, thereby achieving the goal of rapid and automated acquisition of corporate competitive intelligence.

Full Text

Strategic Intent Research of Mergers and Acquisitions Based on Entity Relations

Liao Jianjun1,2,3
1National Science Library, Chinese Academy of Sciences, Beijing 100190
2University of Chinese Academy of Sciences, Beijing 100049
3Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015

Abstract

[**Purpose/Significance**] To address the acquisition issues of enterprise strategic intelligence in competitive intelligence, this paper proposes an automatic and efficient strategic intent mining process to help enterprises better collect strategic intelligence information and improve the effectiveness of strategic intent mining. [**Method/Process**] This study introduces entity relation extraction technology to explore a systematic solution for acquiring enterprise strategic intelligence. Taking enterprise merger and acquisition events as examples, it applies relation extraction technology to collect M&A relations, employs entity attribute extraction technology to reveal the strategic intent of M&A, and verifies the effectiveness of the extraction system through Alibaba Group's M&A cases. [**Result/Conclusion**] The proposed solution can effectively and automatically extract enterprise M&A relations from massive online information and identify enterprises' business domains. By visually presenting the results, it reveals the strategic intent of enterprise M&A, achieving the goal of rapid and automatic acquisition of enterprise competitive intelligence.

Keywords: entity relation; relation extraction; mergers and acquisitions; strategic intent; competitive intelligence

1 Introduction

Mergers and acquisitions have become a major trend among Chinese Internet companies. On November 6, 2015, Alibaba Group and Youku Tudou Group jointly announced that they had signed a merger agreement for Alibaba to acquire Youku Tudou shares for approximately $4.5 billion in cash. Alibaba is an e-commerce company, while Youku Tudou is an entertainment video company. What was Alibaba's strategic intent in acquiring Youku Tudou, a company with a different business nature?

The Internet contains vast amounts of public information in public media, and acquiring business competitive intelligence from open information sources has become a current research hotspot. In enterprise competitive intelligence, users are most concerned about competitors' strategic planning intelligence, but this is also the most difficult to obtain. M&A events scattered across the Internet can reflect enterprises' strategic planning to a certain extent. Therefore, if M&A relations can be extracted from complex web pages, it will help enterprises more comprehensively analyze competitors' strategic intents and achieve the goal of knowing both themselves and their opponents. However, current research mostly focuses on the application of competitive intelligence in the M&A process [1-3], with few studies examining competitors' strategic intents through enterprise M&A. Thus, extracting M&A relations between enterprises from massive news web pages and revealing competitors' strategic intents is of great significance for acquiring enterprise competitive strategic intelligence.

2 Enterprise Mergers and Acquisitions

Enterprise M&A (merger & acquisition) generally refers to the behavior of an enterprise purchasing another enterprise's stocks or assets with cash or securities to enhance its competitive advantages and achieve business objectives. M&A can be divided into narrow and broad senses. Narrow M&A generally includes mergers and acquisitions, while broad M&A also includes various other forms of asset restructuring.

2.1 M&A and Competitive Intelligence

Dr. P. Dishman from Brigham Young University, a renowned scholar in competitive intelligence research on M&A, has publicly emphasized many times: "M&A is a competitive intelligence activity" [4]. Chinese intelligence expert Bao Changhuo pointed out that competitive intelligence is information and research about competitive environment, competitors, and competitive strategies [5]. Competitive intelligence is a process that collects business information, identifies and analyzes competitors' capabilities, plans, intentions, strengths, and weaknesses, and transforms them into intelligence to support business decisions. Enterprise M&A is one of the most important strategic decisions in enterprise development, and competitive intelligence plays a crucial role in decision-making success.

M&A strategy belongs to the category of strategic intelligence, usually formulated by senior management, and requires due diligence on the target company through competitive intelligence means to ensure M&A success. When conducting competitive intelligence investigations for M&A strategic planning, it is first necessary to clarify one's own M&A needs and also to master various relevant issues such as the target company's background and history, industry and enterprise marketing methods, manufacturing methods, financial information, and R&D plans. Therefore, M&A is also a competitive intelligence activity. Competitive intelligence runs through all processes before, during, and after M&A, collecting, analyzing, screening, and evaluating potential M&A targets through competitive intelligence analysis methods, thereby providing decision support for M&A strategy.

2.2 Strategic Intent of M&A

The purpose of M&A is to give full play to both parties' operational synergy effects, market share effects, and other advantages to enhance the enterprise's market competitiveness, which fully reflects the enterprise's development strategic intent. G. Hamel and C. K. Prahalad believe that strategic intent is "an ambitious and grand dream that provides motivation for the enterprise and enables it to embark on the road to future success" [6]. Simply put, strategic intent is the vision that an enterprise wants to achieve in its long-term development. The most common method for acquiring competitor intelligence is to collect competitors' strategic action plans, including product R&D, new business de-

velopment, and M&A. Obviously, for acquiring competitors' M&A strategies, the implementation of competitors' M&A is the best source of intelligence information.

Enterprise acquisition relation refers to the relationship formed when one enterprise acquires another. Usually, enterprise acquisition relations reflect certain strategic intents of the dominant enterprise and are therefore crucial for competitive intelligence analysis. For example, in May 2014, Google acquired the British big data company Rangespan, indicating Google's strategy to strengthen its e-commerce business through big data technology. The process and changing trends of M&A to some extent reflect information such as enterprise strategic dynamics and market operation tendencies. If an enterprise can obtain its competitors' M&A relation information, it will undoubtedly be beneficial for understanding competitors' business development strategies and enhancing its core competitiveness. This paper proposes a new method based on entity relation extraction to reveal enterprises' development strategic intents from the side, based on the operational synergy effects and market share effects of M&A.

3 Entity Relation Extraction

Manual methods of collecting competitive intelligence are no longer suitable for the surge in volume of open information sources, and intelligent information collection methods are urgently needed. Information extraction technology emerges to meet this need, aiming to extract important information from text and store it structurally for further organization, processing, and utilization. Entity relation extraction is a core task in information extraction, which extracts relations between named entities in text, such as employment relations, family relations, and cooperative relations. Entity relation extraction is widely used in natural language processing fields such as semantic retrieval, intelligent Q&A, and machine reading.

Named entities were proposed at the Sixth Message Understanding Conference (MUC6), with the goal of identifying specific entities in text, such as person names, locations, organizations, times, and numerical values [7]. The National Institute of Standards and Technology (NIST) defined 7 types of entity relations and 18 subtypes in the 2008 Automatic Content Extraction (ACE) evaluation plan [8]. O. Etzioni et al. [9] divided entity relation extraction methods into three categories: rule-based methods, supervised methods, and unsupervised methods.

Rule-based methods predefine rule structures for relations between entities and extract entity relations through pattern matching. C. Aone et al. [10] extracted relations between entities by analyzing content features of text and having domain experts write rule descriptions for entity relations. K. Humphreys et al. [11] first conducted extensive relational rule analysis on text, selecting rules that would not produce erroneous relations to extract entity relations. The advantage of rule-based methods is fast extraction speed and high accuracy; the

disadvantage is that they require manual participation in rule writing and are difficult to transplant to unrelated domains.

Supervised methods train models on manually annotated corpora through machine learning to identify entity relations in new text. Machine learning methods for training corpora include kernel function methods [12-13], logistic regression methods [14], enhanced parsing methods [15], and conditional random field methods [15-16]. These methods can solve relation extraction problems in a supervised manner. The disadvantage of this method is that it requires spending a lot of time and effort annotating corpora, but the advantage is that it can adapt well to other domains as long as there is sufficient training corpus.

Unsupervised methods automatically discover and annotate their own corpora through domain-independent extraction patterns. KnowItAll [17] was the first information extraction system to use unsupervised methods, capable of large-scale, domain-independent extraction of relations from web pages, using a set of general patterns to automatically instantiate specific relation extraction rules. M. Banko et al. [18] introduced the Open IE paradigm, proposed the concept of open relation extraction (ORE), developed the TextRunner system, used heuristic rules to automatically build rules in the Penn Treebank, and achieved large-scale relation extraction. O. Etzioni et al. [19] designed the ReVerb system based on rules for identifying relations from verb phrases, prioritizing relation identification and then extracting the left and right entities of the relation. L. D. Corro et al. [20] proposed the ClausIE system, which uses a dependency parser to parse sentences and achieve relation extraction. The advantage of unsupervised methods is that they do not require predefining entity relation types or rules, thus facilitating algorithm transplantation across domains; the disadvantage is that extraction results heavily depend on corpus quality and may produce meaningless results.

Since there are large amounts of M&A news on the Internet that need to be obtained and processed in a timely manner, and to avoid manual semantic annotation and training corpora while facilitating transplantation to research domains, this paper decided to adopt unsupervised methods to extract entity relations for M&A. That is, using verbs as control words, extracting verb relations from Chinese sentences through Stanford dependency parsing, and then finding the subject and object entities of verb relations to achieve entity relation extraction.

4 Acquiring M&A Strategy Based on Entity Relations

M&A relations in business information describe the connections formed between enterprise entities due to activities such as acquisitions and mergers. M&A relations are a type of shallow knowledge that can provide some factual information about enterprises from the side. When relevant M&A relations are organized to form a correlation network, M&A information transforms into useful competitive intelligence that supports enterprises' strategic decision-making. The

main research work of this paper is to design a specific entity relation extraction scheme based on entity relation extraction technology and taking domestic M&A relations on the Internet as examples to acquire enterprise strategic intelligence.

4.1 M&A Relation Extraction

The extraction of M&A relations essentially involves extracting acquisition relations between enterprise entities. M&A relations can be represented by relation triples <Company1, Verb, Company2>, where Company1 is the actor in the M&A relation (the acquiring enterprise), Verb is a synonymous verb of the M&A relation such as "merge," "acquire," or "purchase," and Company2 is the recipient in the M&A relation (the acquired enterprise). For example, from the sentence "Lenovo announced the acquisition of Motorola Mobility for $2.9 billion," we can extract the M&A relation between enterprises "Lenovo" and "Motorola Mobility" as <Lenovo, acquire, Motorola Mobility>. This paper uses StanfordParser to achieve dependency parsing of Chinese sentences. StanfordParser is a dependency parsing tool provided by the Stanford Natural Language Processing Group that can analyze sentence grammatical structures and support relation extraction in multiple languages including English, Chinese, German, French, and Arabic.

4.2 Enterprise Attribute Relation Extraction

After parsing the Chinese example sentence through StanfordParser, the following dependency relations are generated: nsubj(announce-2, Lenovo-1), root(ROOT-0, announce-2), prep(acquire-6, with-3), nummod(dollar-5, 2.9 billion-4), pobj(with-3, dollar-5), ccomp(announce-2, acquire-6), nn(mobile-8, Motorola-7), dobj(acquire-6, mobile-8), xsubj(acquire-6, Lenovo-1). To visually display the dependency relations between words in the sentence, a grammatical dependency graph is drawn, as shown in Figure 1 [Figure 1: see original paper].

As can be seen, there are two control verbs: "announce" and "acquire." "Announce" has only one subject relation "nsubj-Lenovo" and no object relation; "acquire" has both a subject relation "xsubj-Lenovo" and an object relation "dobj-mobile," yielding a relation <Lenovo, acquire, mobile>. However, this is not complete, as "mobile" does not equal "Motorola Mobility." Therefore, we must also consider the strong modification relation between nouns "nn-Motorola" to obtain the final relation extraction result <Lenovo, acquire, Motorola Mobility>.

From the M&A relation <Lenovo, acquire, Motorola Mobility>, we can obtain the fact that "Lenovo acquired Motorola Mobility," but what is the intent? This is still unclear. Through research on the enterprise's own attributes, the strategic intent of M&A can be revealed through the business scope of the acquired company. For example, Motorola Mobility's business scope is smartphones, so we can roughly understand that Lenovo's strategic intent in acquiring Motorola Mobility is to further strengthen its competitiveness in smartphones. To further

mine the strategic intent of M&A, we also need to identify the business scope or service domain of the acquired enterprise. Therefore, this paper links the acquired enterprise entity to the entry information in Baidu Baike. This has two benefits: first, it verifies whether the extracted acquired enterprise is a real enterprise entity; second, it obtains the business scope attribute information of that enterprise entity.

Baidu Baike [21] is an open and free online encyclopedia that aims to create a Chinese knowledge encyclopedia covering all fields and serving all Internet users. However, Baidu Baike has not yet opened a free API and does not support direct background queries of entry information. It can only extract the "business scope" from the enterprise entry information box on the front end through pattern matching methods, as shown in Figure 2 [Figure 2: see original paper].

First, the system simulates URL queries to search for entries in the encyclopedia, locates the detailed page of the enterprise, then uses the HTMLParser tool to crawl the webpage locally and parse it into a DOM tree, and finally obtains the value of the business scope "smartphone" by traversing the DOM tree of the page. In actual extraction, some entry information boxes do not have the "business scope" attribute, in which case synonymous information attributes such as "industry" or "service scope" can be considered.

4.3 M&A Relation Extraction Technical Architecture

Based on the above extraction strategies for M&A relations and attribute relations, this paper constructs an M&A relation strategic intent revelation system based on vertical search and entity relation extraction technology. The technical architecture is mainly divided into five layers: data preparation, information collection, information processing, information extraction, and intent revelation, as shown in Figure 3 [Figure 3: see original paper].

4.3.1 Data Preparation This layer uses Baidu's website vertical search function to retrieve news related to the research theme from Sina's technology news channel, and then pattern-matches the news URLs from the result list for later use. Since later stages require matching business domains of enterprise entities from Baidu Baike, this stage also needs to implement retrieval of relevant entries.

4.3.2 Information Collection This layer uses web crawling tools to load the list of relevant news URLs parsed in the data preparation stage and extracts the HTML text of relevant webpages.

4.3.3 Information Processing This layer mainly completes the task of cleaning and processing the crawled webpages, including: (1) main content extraction, which extracts the article's main content and filters out noise information

unrelated to the news report, such as hyperlinks and comments; (2) sentence splitting, which divides the entire article into individual sentences; (3) feature extraction, which extracts characteristic words according to the research theme, such as “merge,” “acquire,” and “purchase”; and (4) candidate sentence selection, which filters out the theme sentences to be extracted through characteristic words.

4.3.4 Information Extraction This layer mainly extracts M&A relations between enterprise entities, including: (1) Chinese word segmentation, which segments the theme sentences for natural language processing; (2) part-of-speech tagging, which tags the part of speech for each word in the sentence; (3) dependency parsing, which uses StanfordParser to perform dependency syntactic parsing of theme sentences to obtain dependency relations between words; and (4) relation extraction, which extracts M&A relations between enterprise entities based on verb relations, and then obtains business attribute relations from Baidu Baike through the information crawling function.

4.3.5 Intent Revelation Through M&A relations and business relations, this layer provides enterprises with summarized and targeted competitive intelligence analysis reports, presented through the Echarts visualization tool to achieve the purpose of supporting decision-making.

5 System Implementation

The framework of this enterprise M&A intelligence collection system adopts the MVC design pattern to implement a three-tier architecture: (1) the client user interface uses Jsp+Echarts technology to achieve visual presentation of strategic intent; (2) the application server side uses Tomcat+Spring Framework as the software support platform, using the Java programming language to implement system functions such as intelligence collection, relation extraction, and attribute mining; and (3) the database server side adopts MySQL 5.7, which not only has data warehouse functions but also supports JSON, enabling storage and integration of relation triple data.

5.1 Extracting M&A Relations

The experimental dataset was collected using Baidu search. Through Baidu’s website search function, news from Sina’s technology channel was retrieved with the search query “Alibaba + acquisition,” obtaining 1,200 relevant report URLs. The system loaded these URLs to crawl webpages, extracted the main content, and selected 1,066 candidate sentences with verb features such as “acquire,” “merge,” “annex,” and “purchase.” StanfordParser was then used to parse these sentences to extract enterprise M&A relations. Relations where the subject did not contain “Alibaba” and uncertain M&A relations containing words like “will,” “intends to,” or “suspected” were removed from the resulting relation set. The objects of these relations were linked to Baidu Baike for entity recognition,

and their corresponding "business scope or service scope" attributes were mined to form new enterprise attribute relations. Finally, the extracted M&A relations and attributes were output in JSON format as node and relation datasets, as shown in Figure 4 [Figure 4: see original paper].

As can be seen in Figure 4, there are null values because corresponding business scope attribute values could not be mined from Baidu Baike, which need to be manually added. Many duplicate relations also need to be processed. Finally, 24 enterprise M&A relations and 24 attribute relations remained.

According to Internet industry classification, the "business scope" was further standardized into corresponding industry domains, obtaining the industry domain distribution of acquired enterprises, as shown in Table 1 .

5.2 Strategic Intent Revelation

The system implements the generation of visual reports based on processed results from relation extraction and attribute mining for strategic intent analysis. Figure 5 [Figure 5: see original paper] can effectively reveal Alibaba's intent in acquiring enterprises. In the figure, pink balls represent acquired enterprises, green balls represent the business scope of acquired enterprises, and red balls represent the industry domains of acquired enterprises. It can be seen that software development, entertainment media, and information aggregation are the key industry domains of Alibaba's acquisitions, followed by cloud computing, information search, education services, finance, and other industries.

When Alibaba's acquired enterprises' industry domains are presented in a pie chart, the enterprise M&A domain analysis pie chart shown in Figure 6 [Figure 6: see original paper] is obtained.

It can be concluded that Alibaba's strategic intent can be roughly divided into three goals:

(1) **Building a diversified e-commerce business ecosystem.** From the perspective of the industry domains of acquired enterprises, to achieve the overall strategic vision of becoming the world's largest e-commerce service provider, Alibaba has acquired 10 related domains needed for its core business development, including software development, entertainment media, information aggregation, logistics, and finance, aiming to extend Alibaba's influence to every user. Through diversified M&A, it can not only conduct technology integration but also establish a diversified e-commerce business ecosystem by monopolizing market entry points.

(2) **Creating big data analysis cloud platform service capabilities.** In its big data strategy, Alibaba acquired 7 software development companies capable of big data analysis and 2 cloud computing companies, focusing on data as the center. By mining users' core business data and behavioral data, it identifies strategic investment directions, thereby forming an ecosystem covering data collection and aggregation in users' daily life,

business, social interaction, learning, and other domains. Finally, with Alibaba Cloud as the supporting platform, it establishes an e-commerce business ecosystem around users' multi-faceted needs.

(3) **Cross-border layout of O2O business model.** Alibaba has transformed its pure e-commerce thinking mode by cross-border acquiring O2O platforms such as Youku Tudou video websites, information aggregation, and entertainment media. This can not only absorb Internet user traffic but also collect rich user behavioral data resources. Meanwhile, using big data to mine user behavioral data can lay the foundation for commercial advertising. O2O, or Online to Offline, refers to making the Internet an offline transaction platform. Alibaba's development of the O2O business model can not only support the e-commerce platform but also expand new profit models to ensure its core competitiveness.

6 Conclusion

With intensifying enterprise competition, diversified operation has become a strategy for sustainable development. As a growing e-commerce company, Alibaba Group implements an M&A development strategy to achieve a diversified Internet O2O business model. This paper automatically acquires and mines Alibaba's acquisition intelligence information from massive information to reveal its strategic intent in M&A, providing a reference for competitive intelligence collection methods for Chinese enterprises. This method can solve the shortcomings of previous intelligence services that only focused on entity terms while ignoring relations between entities, advancing competitive intelligence analysis from traditional methods such as automatic classification and topic mining to the semantic mining level. It automatically acquires strategic intelligence information around relations between entities and provides visual presentation of strategic intent, facilitating rapid acquisition of competitive intelligence by enterprise intelligence departments.

Of course, this system solution inevitably has aspects that need further improvement. For example, the extraction accuracy of enterprise entities and their attributes needs to be further improved. Due to their diverse expressions and poor regularity, better solutions need to be proposed. The extraction of time and acquisition amount can be added to enrich the semantic information of M&A. A complete domain knowledge base can also be established, such as standardized enterprise institutions and their attribute information, to provide complete information for consistent processing of entities.

References

[1] Li Guoqiu, Lü Bin, Wu Kun. A competitive intelligence evaluation model for enterprise M&A target selection—Also on the lack of competitive intelligence in CITIC Securities' failed acquisition of GF Securities [J]. Information Studies: Theory & Application, 2008(1): 55-59.

[2] NORLING P M, HERRING J P, ROSENKRANS W A, et al. Putting competitive technology intelligence to work[J]. Research-technology management, 2000, 43(5): 23-28.

[3] WAMBUGU P W. An analysis of the effects of competitive intelligence practices on the performance of pharmaceutical companies in Nairobi[J]. International academic journal of human resource and business administration, 2015, 1(4): 1-13.

[4] BRUNER R F. Does M&A pay? A survey of evidence for the decision-maker[J]. Journal of applied finance, 2002, 12(1): 48-68.

[5] Bao Changhuo, Xie Xinzhou. Enterprise Competitive Intelligence System[M]. Beijing: Huaxia Publishing House, 2002.

[6] HAMEL G, PRAHALAD C K. Strategic intent[J]. Harvard business review, 2005, 83(7): 148-161.

[7] MANSOURI A, AFFENDEY L S, MAMAT A. Named entity recognition approaches[J]. International Journal of computer science and network security, 2008, 8(2): 339-344.

[8] Automatic Content Extraction 2008 Evaluation Plan (ACE08) [EB/OL].[2015-11-15].http://itl.nist.gov/iad/mig/tests/ace/2008/doc/ace08-evalplan.v1.2d.pdf.

[9] ETZIONI O, BANKO M, SODERLAND S, et al. Open information extraction from the web[J]. Communications of the ACM, 2008, 51(12): 68-74.

[10] AONE C, HALVERSON L, HAMPTON T, et al. SRA: description of the IE2 system used for MUC-7[C]//Proceedings of the seventh message understanding conference(MUC-7). Fairfax: VA, 1998.

[11] HUMPHREYS K, GAIZAUSKAS R, AZZAM S, et al. University of Sheffield: description of the LaSIE-II system as used for MUC-7[C]//Proceedings of the seventh message understanding conferences (MUC-7). San Francisco: Morgan Kaufmann Publishers Inc., 1998.

[12] ZHAO S, GRISHMAN R. Extracting relations with integrated information using kernel methods[C]//Proceedings of the 43rd annual meeting on association for computational linguistics. Morristown: Association for Computational Linguistics, 2005: 419-426.

[13] BUNESCU R, MOONEY R. Subsequence kernels for relation extraction[C]//Advances in neural information processing systems. Cambridge: MIT Press, 2005: 171-178.

[14] KAMBHATLA N. Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations[C]//Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Stroudsburg: Association for Computational Linguistics, 2004: 22.

[15] MILLER S, FOX H, RAMSHAW L, et al. A novel use of statistical parsing to extract information from text[C]//Proceedings of the 1st North American chapter of the association for computational linguistics conference. San Francisco: Morgan Kaufmann Publishers Inc., 2000: 226-233.

[16] CULOTTA A, MCCALLUM A, BETZ J. Integrating probabilistic extraction models and data mining to discover relations and patterns in text[C]//Proceedings of the main conference on Human Language technology conference of the North American chapter of the association of computational linguistics. New York: ACM, 2006: 296-303.

[17] ETZIONI O, CAFARELLA M, DOWNEY D, et al. Unsupervised named-entity extraction from the web: an experimental study[J]. Artificial intelligence, 2005, 165(1): 91-134.

[18] BANKO M, CAFARELLA M J, SODERLAND S, et al. Open information extraction from the Web[C]//Proceedings of the 20th international joint conference on artificial intelligence. San Francisco: Morgan Kaufmann Publishers Inc., 2007: 2670-2676.

[19] ETZIONI O, FADER A, CHRISTENSEN J, et al. Open information extraction: the second generation[C]//Proceedings of the twenty-second international joint conference on artificial intelligence. Menlo Park: AAAI Press, 2011: 3-10.

[20] DEL CORRO L, GEMULLA R. ClausIE: clause-based open information extraction[C]//Proceedings of the 22nd international conference on World Wide Web. New York: ACM, 2013: 355-366.

[21] Baidu Baike[EB/OL].[2015-11-15].http://baike.baidu.com/view/1.htm.

Note: Figure translations are in progress. See original paper for figures.