Establish Evidence Chain Model on Chinese Criminal Judgment Documents Using Text Similarity Measure

Yixuan Dong[1,2], Yemao Zhou[1,2], Chuanyi Li[1,2], Jidong Ge[1,2(✉)],
Yali Han[1,2], Mengting He[1,2], Dekuan Liu[1,2], Xiaoyu Zhou[1,2],
and Bin Luo[1,2]

[1] State Key Laboratory for Novel Software Technology,
Nanjing University, Nanjing 210093, China
gjdnju@163.com
[2] Software Institute, Nanjing University, Nanjing 210093, China

Abstract. One of the most prominent issues in criminal judgment documents is the insufficient evidence. In order to raise the level of judgment documents reasoning, we need to evaluate the quality of evidence in judgment documents. In the recent informatization of Chinese courts, the huge amount of law cases made it necessary to automate the evaluation of evidence. Constructing the model of evidence chain is the basis for assessing the quality of the judgment documents as evidence chain model can describe the relationship between evidence and fact as well as the relationship between evidence more intuitively. In trying to achieve all above mentioned, we propose a model of evidence chain based on Chinese criminal judgment documents. Automated text preprocessing of Chinese criminal documents creates semi-structured XML documents and in XML file, we can get evidence set and fact set. Key element extraction based on syntactic parsing is used to get keywords of each evidence and fact. Text similarity measure based on Word2vec and keyword overlap ratio calculation is used to get the connection point of evidence chain. Predefined weight of different kinds of evidence can be used to measure the importance of each evidence chain. Table format and graphical format make it possible for us to see the structure of evidence chain.

Keywords: Criminal judgment documents · Judgment documents reasoning
Big data · Evidence chain · Text similarity measure · Word2vec
Weight of evidence chain

1 Introduction

In the process of hearing a case, the Chinese People's Court writes judgment document according to law and details of the case, which records the process and result of each case. The judgment document is one of the most important instrument for resolving procedural issues and substantive issues. The reasoning of judgment documents refers to the explanation of the reasons in legal domain.

At present, many courts appear to pay little attention to the reasoning of judgment documents. In recent years, most of the injustices that have strong social influence are due to the error of fact instead of applying wrong law. The error of fact is most likely to happen when we do not have enough evidence or the quality of evidence is low [1]. At present, it is generally believed by all walks of life in the community that the reasoning of judgment documents is an important mechanism for promoting judicial fairness, implementing judicial openness and enhancing the credibility of the judiciary [2]. In order to achieve the goal of raising the level of reasoning of judgment documents, it is necessary to study the automated assessment techniques of judgment documents.

In order to extract evidence and construct evidence chain model of each judgment document, we have some challenges as follows:

(1) Although Chinese judgment document has a general fixed format, it is still written in natural language. So, how can we get fact and evidence information from the judgment document?
(2) There is so much information in an evidence record and a fact record, which kinds of key elements should be selected? How to extract these key elements?
(3) As every fact is inferred by a set of evidence, how can we build the relationship between fact and evidence?
(4) In criminal case, different evidence may combine together to prove one fact, how can we find the link point of different evidence?
(5) A criminal judgment document usually has much evidence, so we may generate many evidence chain units for one document, in this case, how can we rank these evidence chain units?
(6) How to store the results of evidence chain model and how to display the relationship clearly?

In this paper, we propose an approach to establish evidence chain model based on Chinese Criminal judgment documents. In order to extract the content from judgment document, we use Chinese natural language processing including Chinese words segmentation, syntactic parsing and regex match based on keywords. Also, in order to calculate the link point in evidence chain model, we use text similarity measure based on Word2vec.

2   Related Work

It is generally believed that the reasoning of judgment documents is an important criterion for judging the quality of judgment documents. Therefore, improving the level of reasoning in judgment documents is an important breakthrough to enhance the level of judicial services. For a long time, Chinese judges have a bad reputation for being unreasonable in judgment documents. This situation has been highly criticized by the academic circles [3]. In order to improve the level of reasoning in judgment documents, we can start from the following three aspects: First, the reasoning about evidence, second, the reasoning about facts and the third is about the reasoning of law [4]. There is an internal connection between evidence and fact, that is, evidence is the basis for proving the facts of a case [5]. One of the essential attributes of evidence is relevance.

Evidence unrelated to fact has no value in judgment documents. The vast majority of evidence cannot be directly related to the facts, and need to be arranged with each other. The combination of the two evidences means that the two have reached a steady and solid link and formed the most basic evidence chain model [6].

The Study on Criminal Evidence Chain is the representative work of the research on evidence chain in China. This thesis defines the basic model of chain unit, main part of unit (chain unit body), key of unit (chain unit head) and connection point in evidence chain, and proposes different kinds of link, for example simple link, multiple link, net link and so on [7]. This thesis points out that each evidence chain has a weight according to its evidence content and the number of evidence in evidence chain.

In order to establish evidence chain model from criminal judgment documents, we need to extract evidence and fact content from criminal judgment document using Chinese Information Extraction. The main function of Information Extraction is to extract specific factual information from the text, which can be structured, semi-structured or unstructured text. Overall, the method of information extraction is divided into two categories: one is based on KDD (Knowledge Discovery in Databases) and data mining methods, mainly from structured and semi-structured data extraction information and the other one is using NLP and text mining methods. The goal is to discover new knowledge from unstructured, open texts and turn them into understandable and useful information [8]. News reports are based on events, in order to clearly illustrate the environment of an event, news reports usually contain the organic combination of six elements: Who, Where, What, When, When, Why, How. If the content of a news article contains these elements of the news, then the news content is considered complete [9]. Although the application scenarios of these papers are not as the same as this one, they are all based on Chinese text.

In order to calculate the link point of different evidence, natural language processing including Chinese Word Segmentation, syntactic parsing, word vector model and text similarity measurement is applied. Chinese word segmentation theory can be attributed to: three main word segmentation algorithm, combinatorial algorithm research, Chinese word segmentation disambiguation, unrecognized word recognition and research on Word Segmentation and POS Marking Evaluation [10]. In general, syntactic parsing methods can be divided into rule-based analysis methods, methods based on statistical analysis and methods of combining statistics with rules [11].

3 Approach

In this section, we describe our approach to establish evidence chain on Chinese criminal judgment documents in detail as follows. Section 3.1 presents the model of evidence chain and introduces an overview of the workflow in our approach. Section 3.2 introduces preprocessing of Chinese criminal judgment documents and how to get fact set and corresponding evidence set of the documents. Section 3.3 introduces the means of extracting keywords of each evidence item using natural language processing. Section 3.4 introduces how to get link points of each evidence chain using text similarity measure. Section 3.5 introduces the approach to calculate the weight of the evidence chain. Section 3.6 introduces different ways to display evidence model.

3.1 Overview

Figure 1 presents the form of evidence chain model. An evidence chain model is a three-layer hierarchical structure. First layer: Each evidence node is composed of evidence unit body and evidence head. Second layer: Link point is composed of evidence heads of two or more different evidence nodes if the heads can match each other. Third layer: evidence node set is connected to fact by link point.
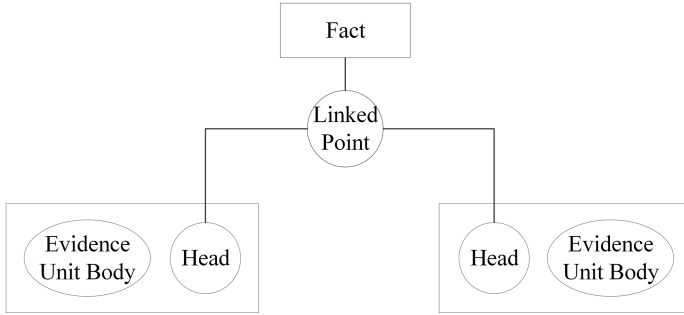
Fig. 1. Form of evidence chain model

Figure 2 presents an overview of workflow of our approach. Because of the characteristics of Chinese natural language, the process is based on Chinese word segmentation. The steps of our workflow are as follows:

(1) Preprocess the text of Chinese criminal judgment documents and extract fact content and corresponding evidence set of each fact.
(2) Extract five different kinds of key elements of each fact and evidence using natural language processing. In this method, we use 4W1H (What, Where, When, Who and How much) as five different kinds of key elements.
(3) Calculate the link points of each evidence head using Word2vec as the text similarity measure methods.
(4) Calculate the weight of each evidence chain.
(5) Display evidence chain model in both table format and graphical format.

3.2 Text Preprocessing

Text preprocessing step can be divided into three sub steps as follows: (1) Extract useful paragraph; (2) Get fact and corresponding evidence set from the paragraph; (3) Label the attributes of evidence.

3.2.1 Extract Useful Paragraph

Chinese judgment documents are usually written in a fixed format. According to the position of a paragraph, the keywords from the first sentence in each paragraph and the writing regulation of each part, we can split an article into seven different parts: headline, litigant participant, and litigant record, basic information of case, trail
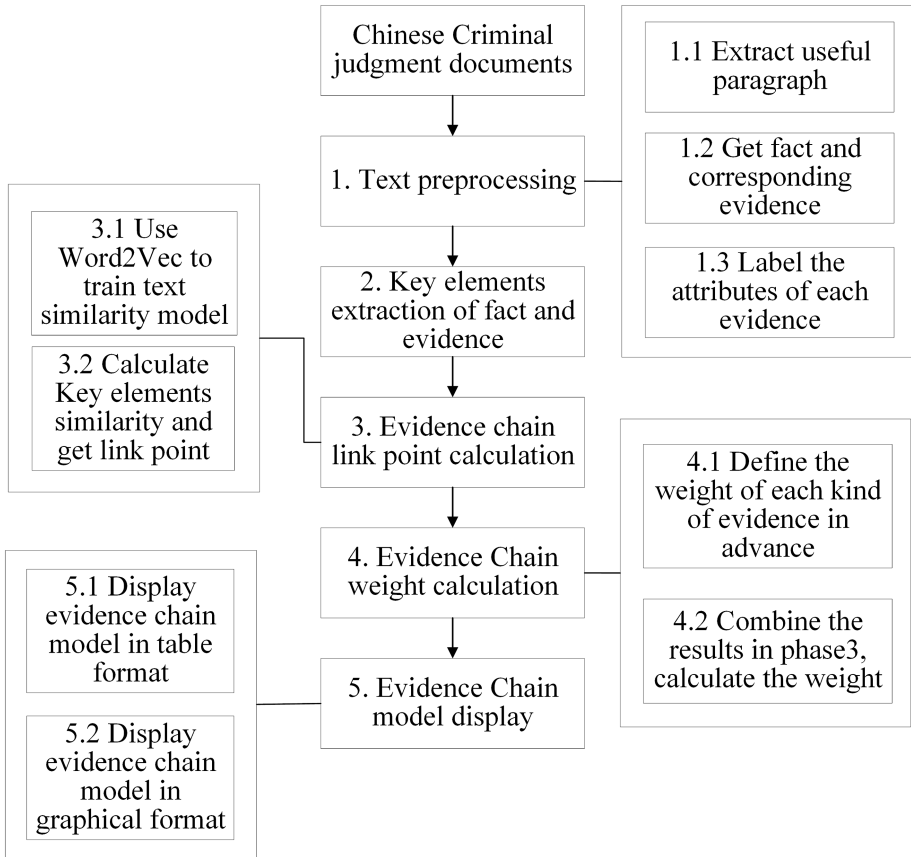
	Chinese Criminal judgment documents	1.1 Extract useful paragraph
	↓	1.2 Get fact and corresponding evidence
3.1 Use Word2Vec to train text similarity model	1. Text preprocessing	1.3 Label the attributes of each evidence
3.2 Calculate Key elements similarity and get link point	2. Key elements extraction of fact and evidence	
	3. Evidence chain link point calculation	4.1 Define the weight of each kind of evidence in advance
5.1 Display evidence chain model in table format	4. Evidence Chain weight calculation	4.2 Combine the results in phase3, calculate the weight
5.2 Display evidence chain model in graphical format	5. Evidence Chain model display	

Fig. 2. Overview of the workflow of establish evidence chain model

process, judgment result and end. If the first sentence in the paragraph has keywords like "法院", "书" and "号", then the paragraph belongs to headline. If the first sentence has words like "纠纷" or "起诉书", then it may not be the litigant participant part. The litigant record appears after the litigant participant part and the basic information of case comes after the litigant record. If the first sentence has keywords like "本院认为" or "审查认为", then it belongs to the trial process part. If the first sentence has keywords like "裁定如下" or "达成如下结论", then it belongs to the judgment result part. If the first sentence has keywords like "速录" and "书记", then it belongs to the end part.

3.2.2 Extract Fact and Corresponding Evidence

In order to get fact and evidence information, basic information is the most important parent paragraph. For example, if a paragraph of basic information contains words like "证据如下", "书证", "物证", "证人证言" and similar keywords, it may be an

evidence section. In terms of fact paragraph, we use the same method, if a paragraph contains words like "经审理查明", "认定XX事实" and similar keywords, then it may be a fact section.

Furthermore, when we get fact paragraph and evidence paragraph from basic information part of a judgment documents, we can extract the relationship between fact and evidence. For instance, if a paragraph contains keywords like "上述事实有经下列证据予以证实", "认定上述事实的证据有" and keywords like these words, the evidence set in this paragraph will be assigned to the fact. After this processing, fact and evidence will be related and the remaining evidence will be put in a specific evidence set called unrelated evidence set.

3.2.3 Label the Attributes of Each Evidence

Faced with more evidence, the three features should be considered as filtering criteria. They are authenticity, legality and relevance. Evidence without these characteristics cannot be incorporated into evidence model [12].

Due to the perfection and development of criminal justice in China, the reasoning of judgment mainly focuses on the evidence part of criminal procedure [13]. So, it is important for us to extract the attributes of criminal evidence as extracting correct attributes is the basic of establishing evidence chain model and evidence chain model can help a lot in the reasoning of judgment.

For each evidence item, we define three attributes for it as follows: the submitter of evidence, the type of evidence and the reasoning result of evidence. The methods to extract each attribute are as follows: (1) to extract submitter of evidence, we use regex match combined with natural language processing. For example, "证人陈某某证实" match the regex of "(.*)证实", so "证人陈某某" is extracted as the submitter of the evidence. Then, we use natural language processing to split "证人陈某某" to "证人" and "陈某某", and then "证人" will be filtered as stop words. (2) Chinese Criminal Procedure Law defines that there are eight different kinds of evidence type of evidence, so we use keywords matching to extract evidence type. (3) Reasoning result of evidence have two different types: accepted and not accepted, so we use keywords matching to extract reasoning result. For example, if the evidence contains words like "不采信", "不采纳", "不予采信" and "不予采纳", then the reasoning result is labeled as not accepted.

3.3 Keywords Extraction

As is mentioned in Sect. 3.1, we use 4W1H as the type of keywords. Because 4W1H have different characteristics, we need to apply different methods to extract each kind of keywords. In this section, we will introduce the methods used to extract keywords.

3.3.1 Extract What

What means the object in evidence and fact. For instance, What usually refers to the murder weapon in a murder case and refers to the stolen things in a theft case. What elements are all nouns and most of them belong to subject or object in a sentence.

In order to extract What, we combine regex match with syntactic parsing as follows:

(1) Regex match: for things that have fixed structure, we can use regex match to extract this element. For example, the title of a book between " 《" and "》 ".

(2) Syntactic parsing: we use HanLP as Chinese syntactic parsing tool. HanLP is based on Chinese dependent syntax analysis based on CRF (Condition Random Field) sequence annotation algorithm and maximum entropy dependency syntax analysis algorithm. This step can be divided into four sub steps as follows: Firstly, we use HanLP to analyze the structure of a sentence. Secondly, we get the description of each word using the sentence structure. For example, "黄色" is used to describe "上衣" and "塑料" is used to describe "热水瓶" in the sentence structure. Thirdly, we can get all subjects and objects of a sentence using syntactic parsing. Fourthly, we combine descriptions and subjects or objects extracted from the sentence to get What element. Finally, we filter stop words from the result.

3.3.2   Extract Where

Where refers to the site information in evidence and fact, for example, the crime spot or witness spot. Where elements are all nouns or location words and most of them appear after prepositions like "在", "于", "至" or similar prepositions.

Two methods are used to extract Where element as follows:

(1) Participle the sentence and extract the words that are labeled as "S" which means site.

(2) Regex match: define a list of prepositions and the words after the prepositions will be extracted as Where elements.

3.3.3   Extract Who

Who refers to the participants mentioned in evidence and fact. POS (Part of Speech) is used to extract Who elements from a sentence, we participle the sentence and get the words that are labeled as "nr" which means name of people or organizations. These words will be extracted as Who element.

3.3.4   Extract When

When refers to time information mentioned in evidence and fact. Time information in evidence or fact is usually in a fixed format so we use regex match to get time information. Here are two examples of the regex:"\d年\d月\d日" and "\d月\d日".

3.3.5   Extract How Much

How much refers to quantitative phrases in evidence and fact and amount information. They are usually in a fixed format: numbers always appear before quantifiers. So we use regex match to get amount information as well.

3.3.6   Generate Stop-Word List Based on Judgment Documents

Stop-word means word that have high frequency of occurrence but do not have important meaning in text such as "is", "of" or "the". The processing of stop words can greatly accelerate the speed of word segmentation and subsequent parsing [14]. In terms of

judgment documents, we have more words that need to be filtered as stop-word. For example, "原告", "被告", "证言" or similar words, these words have high frequency in judgment documents but have little meaning in law related information extraction. To construct a stop-word list that meets requirements talked above, we choose 100000 criminal judgment documents as our corpus and calculate IDF of each unique word. IDF (Inverse Document Frequency) refers to the inverse fraction of documents that contain a specific word. Words with low IDF can be regarded as potential stop words. After scanning all judgment documents and filtering common stop words, we calculate IDF of rest words and sort words in ascending order. Top N words have been chosen as stop words. Stop words we choose are added into the set of common stop-word list to construct the stop words list we use to filter candidate set of chain unit heads.

3.4   Link Point Calculation

As we have introduced in Sect. 3.1, link point is the fundamental component in evidence chain, so link point calculation is of great importance in establishing evidence chain. Link point calculation can be divided into two sub steps: (1) Calculate the Link point between evidence. (2) Build relationship between evidence and fact.

3.4.1   Calculate the Link Point Between Different Evidence

In terms of calculating the head of evidence we mainly use two methods:

(1) Calculate link point using equality of keywords extracted in Sect. 3.3.

In this method, we compare keywords of an evidence with keywords of another evidence in a specific keyword type such as "What". If these two keywords are equal to each other, then these two keywords will be extracted as the head of each evidence and the head will be the link point of these two evidence.

After the loop, we can get the link points between evidence.

(2) Calculate link point using text similarity measure of keywords extracted in Sect. 3.3 based on Word2vec.

In this method, we compare keywords of an evidence with keywords of another evidence in a specific keyword type such as "What". If these two keywords are similar to each other semantically, then these two keywords will be extracted as head of each evidence.

In calculating the similarity between words, we convert words into word vectors and measure the similarity between words by calculating the cosine distance of word vectors. We use Word2vec model to convert words into word vectors. The specific steps are as follows:

Firstly, train Word2vec model on all Chinese criminal judgment documents in the corpus. In this step, we choose all criminal documents and get the evidence part and fact part of each article. Then we participle the evidence part and fact part and get all the words except stop-word and use these words as the training corpus and train the Word2vec model on this training corpus. After this step, we get a Word2vec model that contains the word vector of each word.

Secondly, determine the threshold for similarity calculation. As there is no uniform standard to determine in which case two words vectors are similar. In this step, we choose 100 judgment documents from the training corpus randomly and participle evidence and fact part of each document. Then we calculate the cosine distance of every two words that lies in different evidence and respectively output the word pairs that have the cosine distance larger than 0.95, 0.9, 0.85 and 0.8. Combine the precision and recall of the similar words extraction, we choose 0.9 as the threshold value in similar words calculation.

Thirdly, participate the keywords that need to be compared and get the word list of it and then filter the stop-word from it.

Fourthly, compare the word similarity of each word that we get from the two keywords respectively by using Word2vec model we trained in step1 and get highest word similarity as the similarity between the two keywords. If the similarity of two keywords are higher than the threshold that we set in step two, these two keywords will be extracted as the head of each evidence.

As we can see from the Table 1, the first method is more efficient but may leave out some information and the second method is somehow less efficient but can extract more useful information. So we apply the first method on the comparison of keywords that belong to "When", "Who" and "How much" as these three types have a fixed format. We apply the second method on the comparison of keywords that belong to "What" and "Where" as these two types have the characteristic that the same thing may have different expression.

Table 1. Comparison of two methods in link point calculation

Method type	Advantage	Disadvantage
Equality of keywords	Simple calculation and efficient	Leave out keywords that have different expression but have same meaning
Text similarity of keywords	Extract more keywords that have different expression but have similar meaning	More complex calculation

3.4.2 Build Relationship Between Evidence and Fact

Section 3.2.2 shows that we extract fact and corresponding evidence from every judgment documents, and the remaining evidence will be put in unrelated evidence set.

So for the evidence which has not be related to a fact, we compare its similarity with each fact by calculating its overlap of keywords and fact's keywords and the evidence will be related to the fact which has the highest overlap with this evidence.

3.5 Weight of Evidence Chain Model

In this section, we mainly focus on how to calculate the weight of evidence chain model. As we may generate a lot of evidence chain models in a judgment document and some of them are of great importance while some are less important in helping

judges analyze the evidence information. So it is essential for us to calculate the weight of each evidence chain model.

In calculating weight of evidence chain model, we ask some experts in legal domain to decide how to calculate the weight of each evidence and get a conclusion that the weight should be decided by the evidence type and reasoning result. In terms of evidence type, authoritative evidence such as documentary evidence or physical evidence provided by police have higher weight than non-authoritative evidence such as testimony of witness. We define that the weight of authoritative evidence is 1.0 and the weight of non-authoritative evidence is 0.8. In terms of reasoning result of evidence, evidence that is confirmed by the judges has higher weight than unconfirmed evidence. So, each evidence item has a weight according to evidence type and reasoning result of evidence in Sect. 3.2.3. Then we add the weight of each evidence item in an evidence chain model and get the weight of an evidence chain model.

As for now, we can get the weight of each evidence chain model in a criminal judgment document automatically; in the future we will apply it by sorting the weight and give the top n evidence chain models of the judgment document as a reference for judges.

3.6 Display Evidence Chain Model

In this section, we will introduce two main ways that we apply to display evidence chain model: the excel format and the graphical format. Excel sheets can clearly present the content and attributes of each fact and evidence and graphical format can directly shows the relationship between fact and evidence. Evidence chain model information is now saved in JSON format. JSON means JavaScript Object Notation, it is a syntax for storing and exchanging data and it is in text format written with JavaScript object notation. Experimental data show that JSON is obviously superior to other data transmission formats in terms of data transmission efficiency, which provides an optimized reference scheme for the selection of data transmission formats in light-weight applications [15]. We can easily convert Json file to object and it is easy for person to read or write.

3.6.1   Generate Excel File of Evidence Chain Model

Evidence chain model of each Chinese criminal judgment documents is saved as an excel file with two sheets:

(1) Evidence sheet focuses on evidence part of a criminal judgment document. In this sheet, we display the evidence and its attributes including evidence content, evidence type, reasoning result and the head list of each evidence that we calculated in Sect. 3.4.1.
(2) Fact sheet focuses on the fact content and relationship between fact and evidence. Identifier, name and content of each fact are shown in the sheet in turn. At the same time, the corresponding evidence and their chain unit heads are displayed in the sheet.

3.6.2 Graphical Display of Evidence Chain Model

Graphical display of evidence chain model focuses on demonstrating the relationship of fact and evidence. As Fig. 3 shows, we can use the system with evidence chain model information in the excel file as input and extracts the evidence content from excel file. Then the evidence chain model we build can be displayed on the interface clearly.
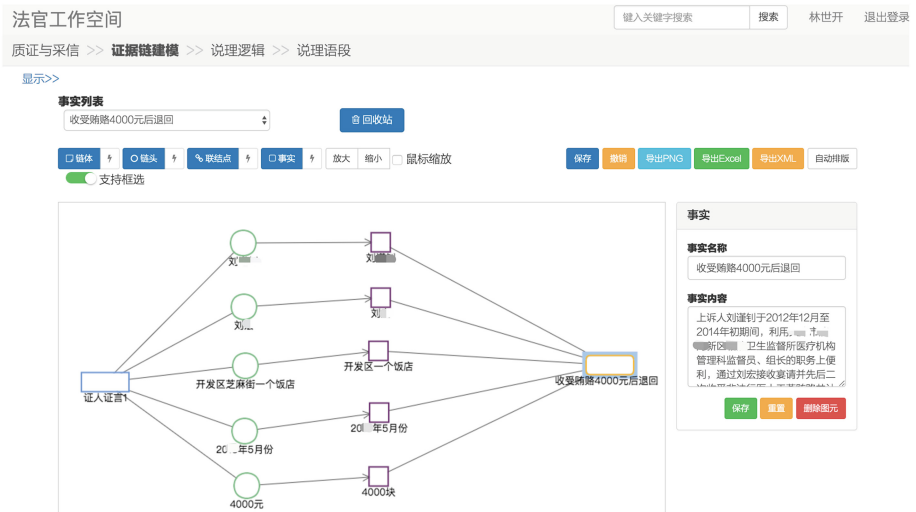

Fig. 3. Form of evidence chain model

4 Experiment and Result

In this chapter, we design some experiments to evaluate the effectiveness of each step in the approach of establishing evidence chain and give the result of each experiment.

4.1 Text Preprocessing

To evaluate the results of paragraph partition, we compare the results of our approach with the information of each case in the trial system database. As Table 2 shows, the accuracy of splitting basic information of a case if higher than 90% and for criminal case, the accuracy is even up to 97%.

As criminal case has more detailed evidence record, so we choose criminal judgment documents to extract fact and evidence. To evaluate the results of extracting fact and evidence set, we compare our results with the evidence and fact information in the trial system database. Table 3 shows the accuracy of extracting fact section and evidence section from basic information of case. It is approximately equal to 90% which means that our approach in text preprocessing is effective.

Table 2. Accuracy of paragraph partition

Type of case	Headline	Litigant participant	Litigant record	Basic information	Trial process	Result
First-instance of civil case	99%	93%	88%	94%	89%	91%
Second-instance of civil case	99%	92%	90%	90%	91%	97%
First-instance of criminal case	99%	97%	96%	97%	98%	98%
Second-instance of criminal case	99%	98%	96%	97%	98%	99%
First-instance of administrative case	99%	94%	88%	92%	91%	98%
First-instance of administrative case	99%	97%	96%	96%	95%	99%

Table 3. Accuracy of fact and evidence extraction

Type of case	Fact section	Evidence section
First-instance of criminal case	88%	88%
Second-instance of criminal case	90%	90%

4.2   Keywords Extraction

As there are no existing tools to extract keywords in terms of 4W1H, to evaluate the results of our approach, we randomly choose 100 criminal judgment documents and label the correct key elements by skilled legal practitioners. Then we compare our results with it and get the results shown in Table 4. As the table shows, recall and precision are almost more than 70%. The evaluation results prove that our methods to extract key elements can lay a solid foundation for finding out link points and establishing evidence chain model in later steps.

Table 4. Evaluation of key elements extraction

Key elements	Precision	Recall	F1
What	68.3%	79.1%	73.3%
Where	81.2%	85.4%	83.2%
Who	90.6%	87.9%	89.2%
When	100%	92.7%	96.2%
How much	98.7%	92.5%	96.9%

4.3   Link Point Calculation

In order to verify the effectiveness of link point calculation, we randomly choose 100 criminal judgment documents and ask some skilled experts in legal domain to label the

correct keywords of each evidence and label the head of each evidence chain model in each judgment document. Then we did two experiments, in the first experiment, we calculate the head simply using equality calculation based on the correct keywords that are labeled by skilled experts and in the second experiment, we calculate the head using text similarity measure combined with equality calculation and the accuracy of the two types are shown in Table 5. As we can see from Table 5, the F1 of link point calculation of What and Where elements rise from 81.6% to 90.5% when we use text similarity calculation instead of equality calculation.

Table 5. Accuracy of two methods in link point calculation

Method type	Precision	Recall	F1
Equality of keywords	97.7%	70.1%	81.6%
Text similarity of keywords	92.4%	89.8%	90.5%

5 Conclusion and Future Work

In this thesis, we propose an approach to establish evidence chain model from criminal judgment documents which is based on Chinese natural language processing. In the experiments, we choose 4W1H as key elements of each evidence and fact and combine equality and text similarity measure based on Word2vec to calculate the relationship of different evidence.

However, establishing evidence chain model from judgment documents is the very first step of assessing the reasoning of judgment documents. In the future, more work need to be done to evaluate the reliability of the fact according to the evidence content in judgment documents. Also, we need to tell which part of fact in judgment document do not have enough evidence to support it. This results will help judges write more reasonable judgment documents, thus promoting judicial fairness and implementing judicial openness.

References

1. Zhang, D.: The optimization of China criminal evidence system. Soc. Sci. Chin. **07**, 125–148 (2015)
2. Lei, X.: Predicament and way out of criminal judgment documents - taking death penalty cases as analysis sample. J. Grad. Sch. Chin. Acad. Soc. Sci. **15**(6), 85–89 (2015)
3. Ling, B.: How judges reason: Chinese experience and universal principles. Chin. Leg. Sci. **15**(5), 99–117 (2015)
4. Yang, Y.: A study of paragraphs of evidence reasoning in judgment documents. Appl. Linguist. **05**(3), 111–115 (2005)

5. Liu, Z., Xiong, Z.: Factual information of evidence. J. Southwest Chin. Norm. Univ. **31**(5), 87–91 (2005)
6. Li, Z.: Evidence chain and structuralism. Chin. Legal Sci. **2017**(02), 173–193 (2017)
7. Chen, W.: The study on the evidence chain. J. Natl. Prosec. Coll. **15**(4), 128–136 (2007)
8. Guo, X., He, T.: Survey about research on information extraction. Comput. Sci. **42**(02), 14–17 (2015)
9. Wang, W., Zhao, D.: Chinese news event ontology construction and auto-population. Comput. Eng. Sci. **34**(4), 171–176 (2012)
10. Feng, G., Zheng, W.: Review of Chinese automatic word segmentation. Libr. Inf. Serv. **55**(2), 41–45 (2011)
11. Shi, C.: Dependency parsing research. Intell. Comput. Appl. **3**(6), 47–49 (2013)
12. Cai, Z.: Criteria of evidence chain integrity and approach to examine. Lawyer World **3**, 11–13 (2003)
13. Chen, J.: On fact-finding in the court and evidence. Southwest University of Political Science and Law (2009)
14. Hua, B.: Stop-word processing technique in knowledge extraction. New Technol. Libr. Inf. Serv. **2**(8), 48–51 (2007)
15. Gao, J., Duan, H.: Research on JSON data transmission efficiency. Comput. Eng. Des. **32**(7), 2267–2270 (2011)