

Tibetan Word Sense Disambiguation Based on a Semantic knowledge Network Diagram

Lirong Qiu^{1*}, Xinmin Jiang¹, Renqiang Ling²

¹ School of Information Engineering

Minzu University of China

Beijing, 100081 China

² University of Wollongong, Wollongong

NSW, Australia

*qiu_lirong@126.com



*Journal of Digital
Information Management*

ABSTRACT: *The method based on semantic knowledge is the most dynamic research direction in rule-based techniques. This method has been proven effective in studying English and Chinese word sense disambiguation. This study proposes two methods for selecting the correct Chinese meaning of Tibetan ambiguous words from Tibetan sentences in Tibetan-Chinese parallel corpora using semantic knowledge from HowNet and translation information from the aforementioned corpora. We can use these methods to build Tibetan-Chinese parallel corpora with word sense tagging. The two proposed methods are 1) the word sense disambiguation method based on HowNet and Tibetan-Chinese parallel corpora, and; 2) the semantic knowledge-based method of network diagram word sense disambiguation.*

Subject Categories and Descriptors

I.2.7 [Natural Language Processing]: Text Analysis; **I.2.6 [Artificial Intelligence]:** Learning-induction; knowledge acquisition

General Terms: Algorithm, Performance

Keywords: Tibetan Information Processing, Word Sense Disambiguation, HowNet, Semantic Knowledge Network Diagram

Received: 29 April 2015, Revised 19 June 2015, Accepted 9 July 2015

1. Introduction

The main research techniques in word sense

disambiguation (WSD) include the dictionary-based WSD method, as well as the supervised and unsupervised WSD methods. The supervised WSD method depends on large-scale, high-quality word sense tagging corpora for training data. Based on training data, the corresponding probabilities for the context of a polysemy and a specific word sense can be counted. In this manner, word sense recognition problems of polysemy translates into classification problems of context. An unsupervised word sense tagging method requires neither dictionary knowledge nor a word sense tagging corpus but directly depends on a large-scale untagged corpus to learn and deduce the meaning of words. Thus, the objective of WSD is achieved.

At present, acquiring large-scale untagged corpus is difficult in Tibetan areas because large-scale tagging and training corpus is lacking. Thus, realizing absolute supervised and unsupervised methods is difficult. Considering the aforementioned reasons, we propose the combined approaches of a Sino-Tibetan bilingual dictionary, the semantic repository HowNet, and Chinese translation information to study the automatic WSD method for Tibetan word sense in Sino-Tibetan parallel corpora and semantic dictionary in section 4. This method mainly makes a selection with lexical semantic similarity and correlation calculation. However, it requires part-of-speech tagging in the corpus preprocessing stage. At present, Tibetan part-of-speech tagging technology remains immature in terms of comprehensive practical applications. The accuracy rate affects the result of WSD. Meanwhile, applying semantic knowledge from HowNet

only relies on using prime information to calculate similarity and relevance levels. Other semantic information from HowNet is not utilized. Considering the aforementioned reasons, this study presents a semantic-based knowledge network diagram WSD method to improve existing WSD methods used in processing minority languages.

In recent years, the network diagram-based WSD method has exhibited favorable disambiguation performance in international WSD evaluation tasks. Under these conditions, this study proposes Tibetan WSD based on a semantic knowledge network diagram. First, a network diagram-based WSD method is used as an example. Then, WSD methods based on HowNet, a Sino-Tibetan

bilingual corpus, and network diagrams are combined to improve the performance of the Tibetan word automatic sense disambiguation method. The semantic knowledge-based network diagram WSD method proposed in this study aims to eliminate part-of-speech tagging in the preprocessing stage, use a variety of semantic relationships from HowNet to extend semantic meaning options, and build a semantic relationship diagram. Word sense can be chosen by calculating the semantic correlation between content and context in a semantic relationship diagram and then converting this correlation into the weight information of the side of a network diagram to find the path with the maximum weight in the diagram. The basic process of the method is illustrated in Figure 1.

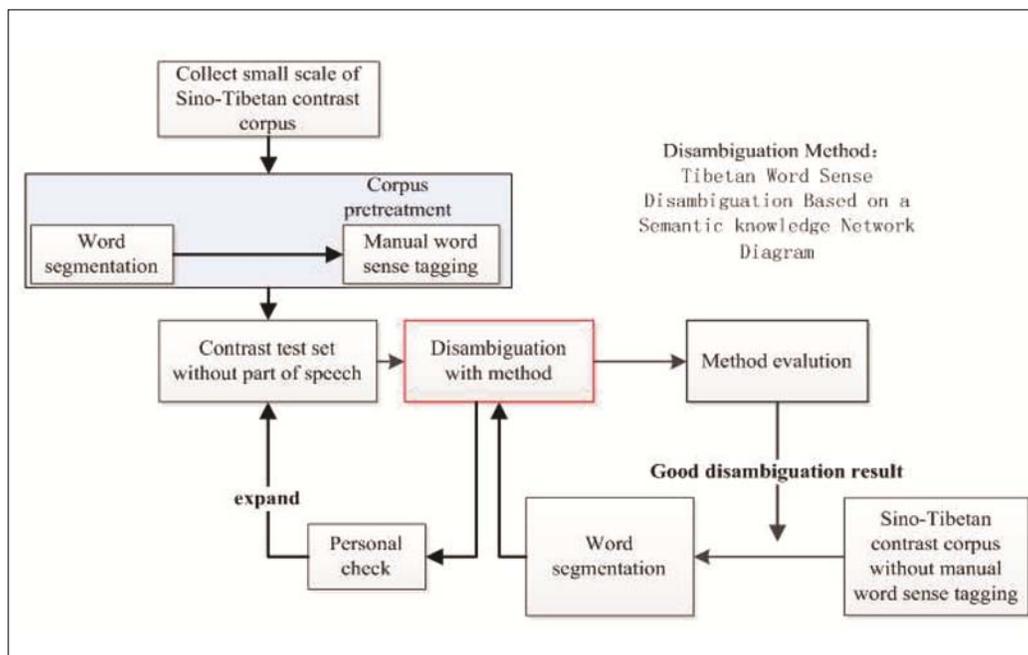


Figure 1. Flowchart of the semantic knowledge-based method of network diagram WSD

Based on the existing method, the Tibetan word sense is automatically disambiguated to build a large training corpus with word sense tagging. The application of the network diagram-based method in this study on WSD is briefly introduced.

The remainder of the paper is organized as follows. Section 2 provides an overview of related works on WSD as preliminaries to our work. Section 3 discusses the data pretreatment process. Sections 4 and 5 explain our work on the Tibetan WSD approach. The empirical analysis and its results are presented in Section 6. Lastly, Section 7 presents the conclusions, discussions, and future work.

2. Related Works

The WSD method, which is based on a network diagram, is both a knowledge-based and an unsupervised approach.

Navigli proposed a method based on a lexical chain, i.e., the structural semantic interconnection method [1]. This method creates a lexical chain of sentence pending disambiguation based initially on the semantic dictionary

WordNet. The vocabulary in the lexical chain has the smallest semantic distance in WordNet. A lexical chain forms a vocabulary chain with interconnections, and then, the “network diagram-based word sense of maximum connectivity” is selected as the sense of ultimate ambiguous words. This method has achieved the best results in an international evaluation of Senseval-3 unsupervised disambiguation for the whole word.

Navigli also explored unsupervised WSD through a repeated study of a connectivity graph. This method has few parameters and does not require a sense-annotated corpus to train an algorithm. He investigated several measures of graph connectivity to identify those that were best suited for WSD [2]. These measures included local and global measures. The local measure had four aspects, namely, degree centrality, eigenvector centrality and key player problem. The global measure also had four aspects, namely, compactness, graph entropy, edge density, and search strategies.

Agirre et al. presented a graph-based approach to WSD in the biomedical domain. This approach used knowledge

from the Unified Medical Language System Metathesaurus, which was presented as a graph [3]. The researchers used Personalized PageRank to judge the importance of a node in this graph. Experiments indicated that the best results were obtained using this method and this kind of knowledge source.

Agirre et al. also proposed a new WSD method based on random walks over large lexical knowledge bases that were built from WordNet and extended WordNet. This algorithm can deal with English and Spanish two kinds of datasets [4]. There are two types' methods that are tested in this random walks for WSD. They are static PageRank without context and personalized PageRank with context. This paper also gives a detailed analysis of the reasons which affect the efficiency of the algorithm.

Yang proposed a network diagram WSD based on word distance [5]. This method did not only consider the strength of semantic relations between words, but also their actual distance in ambiguous sentences. The method also considered the fact that the words that were close to an ambiguous word would have a greater effect than the words that were from an ambiguous word.

Hessami proposed a graph-based unsupervised WSD method. This method adds some ambiguity words' senses to a set G, and then builds a tree for all elements in the set G according to the lexical relation of WordNet [6]. Searching a tree to find the best path which can link to other senses that belongs to the set G. Building graph for all senses in the set G according to the search result of all lexical relation trees. At last, using the connectivity measure to find the best senses for ambiguity words.

Usbeck et al. uses linking data to fulfill a graph-based disambiguation of named entities [7]. This method makes use of some knowledge bases like WordNet, DBpedia and YAGO2 which can be called as underlying knowledge base to build disambiguation graph for all candidate resources of named entities that are extracted from the input text. And then, using the HITS algorithm to find authoritative candidates for discovered named entities. The whole method is called as AGDISTIS.

The aforementioned methods typically adopt WordNet as a knowledge source in studying English WSD. In Chinese studies, HowNet is generally chosen as a knowledge source because it can provide more semantic information than WordNet. The current study aims to choose the correct Chinese interpretation for ambiguous Tibetan words. The calculations of word semantic correlation are completed within the Chinese scope; thus, HowNet is selected as the knowledge source. In addition, ambiguous words that are farther from the context typically have less influence than closer words, and thus, this context considers the effect of semantic distance between words in context on semantic correlation calculation.

3. Corpus Preprocessing

For the method used in this study, the corpus preprocessing stage only has to undergo two stages, namely, corpus word separation and manual word sense tagging, as shown in Figure 1. Chinese and Tibetan word corpora must undergo word separation processing to enable further analysis. Manually tagging a Tibetan corpus provides contrasting data for the WSD method after the word sense tagging process is automatically completed using a computer. This process should be compared with the manually labeled word sense. The accuracy and recall rates of the automatic WSD method must be completed, and the WSD method must be evaluated.

In preprocessing a Tibetan-Chinese corpus, we use the Stanford Segmenter to split words. This software can be downloaded for free from <http://nlp.stanford.edu/>. The Stanford Segmenter can obtain F-scores from experiments on the Taiwan Academia Sinica corpus, the City University of Hong Kong Chinese corpus, the Peking University Institute of Computational Linguistics corpus, and the Microsoft Research Asia corpus, which are 0.947, 0.943, 0.950, and 0.964, respectively [8]. We also use the Tibetan segmenter developed by the China Ethnology and Anthropology Institute of the Academy of Social Sciences. The accuracy rate, recall rate, and F-score of this segmenter are 91.27%, 90.85%, and 0.9106, respectively, on the open test set [9].

4. WSD Method Based on Hownet and Tibetan-chinese Parallel Corpora

First, we introduce a Tibetan WSD method that is based on HowNet and Tibetan-Chinese parallel corpora. This method uses word similarity and relevance as evaluation criteria for selecting the definition of an ambiguous word. The calculation method for word similarity and relevance is used in the next section. This method is important; thus, we need to introduce it first. This method is also compared with the method based on a network diagram. After corpus preprocessing, a Sino-Tibetan comparison test set with word sense and split words can be obtained. In addition, the corpora that are used in this method must mark the parts of speech unlike the method shown in Figure 1. The following steps should be undertaken when selecting word sense from the Tibetan vocabulary in the test set.

- (1) Extract the Chinese interpretation of a notional word in a pending Tibetan sentence from a Sino-Tibetan bilingual dictionary (a structural word is not included in the study content of this paper). If a word has multiple meanings, then mark the Tibetan word as an ambiguous word and proceed to Step

- (2). Process the Tibetan words marked as ambiguous individually. Calculate the semantic similarity between semantic Tibetan word sense and Chinese words with the corresponding part of speech in the comparison corpus. Choose the eligible sense based on the result. The selection condition is that the results of the word

semantic similarity calculation must be greater than \acute{a} (the specific value of \acute{a} must be determined after conducting several experiments). If the number of eligible sense is greater than 1, then proceed to Step (3). If the aforementioned number is equal to 1, then set this sense as the specific interpretation of the ambiguous word; if the number is equal to 0, then select the calculation result with the maximum semantic similarity as the correct sense.

(3) Calculate the semantic relevancy between the

candidate sense of current ambiguous Tibetan vocabulary and other part-of-speech notional words. Then, select the calculation result with the maximum semantic relevancy in the candidate sense as the correct sense of the current ambiguous word.

During word sense collection, focus must be given on the special treatment of Tibetan verbs. These verbs typically occur during the change of tenses in a sentence. Table 1 provides examples of Tibetan verb tense changes.

No.	English Interpretation	Present Tense	Past Tense	Future Tense
1	dig	ཚོག་པ་	བཟོས་པ་	བཟོ་བ་
2	meet	འཇུག་པ་	བཅུག་པ་	གཞུག་པ་
3	read	ཟོག་པ་	བཟོགས་པ་	བཟོག་པ་
4	attach	ལྷུག	འལྷུགས	འལྷུག
5	declare	ལྟོག་པ་	འལྟོགས	འལྟོག
6	arrange	ལྟོག་པ་	འལྟོགས	འལྟོག

Table 1. Comparison table of three tenses of Tibetan verbs

In this study, a comparison table of three tenses of Tibetan verbs is established for the problem of multiple representations of Tibetan verb forms. For verbs expressed in past and future tenses, the verb tense comparison table must be consulted to identify the corresponding verb in the present tense. Subsequently, the corresponding Chinese interpretation must be collected from the dictionary.

The word sense similarity calculation in Step (2) adopts Equation (1). β_i ($1 < i < 3$) is the weighting parameter, and $\beta_1 + \beta_2 + \beta_3 = 1$, $\beta_1 > \beta_2 > \beta_3$.

$$sim(w_1, w_2) = \sum_{i=1}^3 \beta_i \prod_{j=1}^i sim_j(w_1, w_2) \quad (1)$$

$$sim_i(w_1, w_2) = \frac{\gamma \times h}{dis(s_1, s_2) + \gamma \times h + |high_{s_1} - high_{s_2}|} \quad (2)$$

$$sim_i(w_1, w_2) = \frac{\gamma \times h}{dis(s_1, s_2) + \gamma \times \max(h, (high_{s_1} - high_{s_2}))} \quad (3)$$

$$sim_i(w_1, w_2) = \frac{\gamma \times h}{dis(s_1, s_2) + \gamma \times \max(h, (high_{s_2} - high_{s_1}))} \quad (4)$$

This study uses Equations (2), (3), or (4) to calculate $sim_1(w_1, w_2)$. $dis(s_1, s_2)$ is the semantic distance between

two sememes, h is the lowest height of the public father sememe, and γ is the adjustable parameter. As shown in Equations (2), (3), and (4), the calculation of this similarity can be divided into three situations.

(1) s_1 and s_2 are at the same branch of the sememe tree. Apart from the relation among the semantic distance and the lowest height of public parent sememe, the similarity result has a negative relation with the height difference of the sememe level, as shown in Equation (2).

(2) For the common node of s_1 and s_2 , as well as initial $high_{s_2} - high_{s_1} > 0$, $sim_1(w_1, w_2)$ will have a negative relation with the maximum value between $high_{s_1} - high_{s_2} > 0$ and the lowest height of the public parent sememe, as shown in Equation (3).

(3) For the common node of s_1 and s_2 , as well as initial $high_{s_2} - high_{s_1} > 0$, $sim_1(w_1, w_2)$ will have a negative relation with the maximum value between $high_{s_2} - high_{s_1} > 0$ and the lowest height of the public parent sememe, as shown in Equation (4).

The sememe set calculation method from [10] is used to calculate $sim_2(w_1, w_2)$.

Using Equation (5) to calculate $sim_3(w_1, w_2)$, $sim_3^i(p_1, p_2)$

expresses the same number of relationships in the complete relation description between two concepts. Finally, the average of $sim_3^i(p_1, p_2)$ is calculated.

$$sim_3(w_1, w_2) = \frac{\sum sim_3^i(p_1, p_2)}{n} \quad (5)$$

After obtaining the similarity calculation result, we also need to calculate the relevance between word sense and other words with different parts of speech in the Chinese translation. According to HowNet theory, we select 16 dynamic semantic roles with high correlation with semantic relevance, namely: AccordingTo, MaterialOf, RelateTo, CoEvent, HostOf, OfPart, SourceWhole, belong, concerning, partner, scope, whole, ResultContent, ResultEvent, ResultWhole, and domain. Equation (6) is used to calculate the relevance between two concepts as follows:

$$rel(w_1, w_2) = \sum_{i=1}^{16} \varphi_i \times sim_i(p_1, p_2) \quad (6)$$

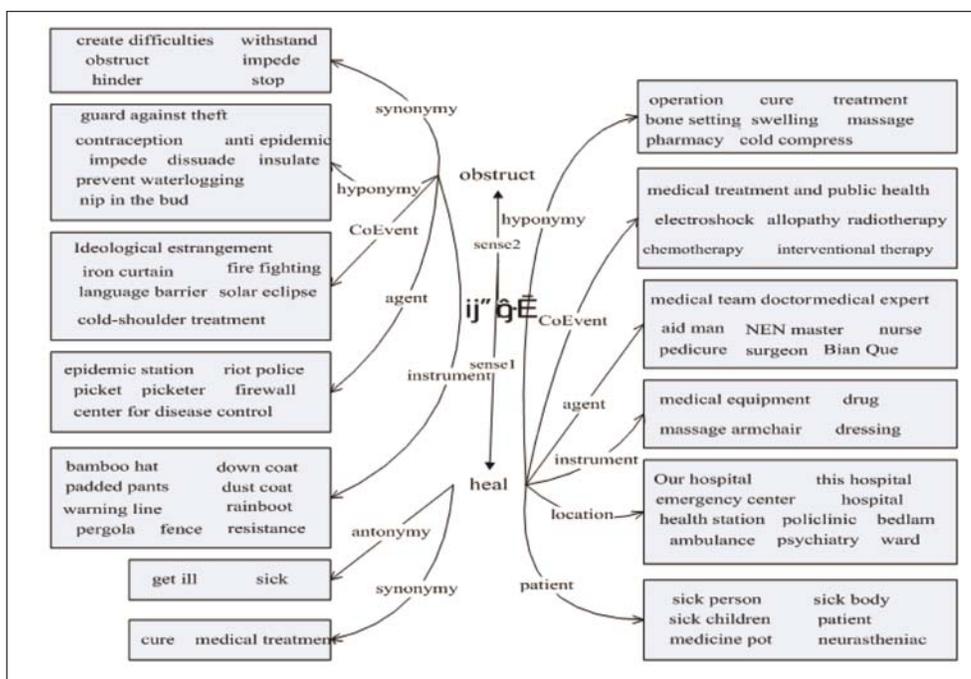


Figure 2. Semantic relation diagram of the Tibetan word

context to complete the sense disambiguation task. In this method, the restrictions of the context in selecting word sense embody the semantic correlation between ambiguous words and the context. A high-level semantic correlation presents that the senses agree with the current context. Performing only a semantic correlation calculation between various senses and context may lead to the incapability of such calculation because of data sparseness. Consequently, this method builds semantic networks on word senses, which can expand the correlation calculation objectives and enhance the accuracy of WSD.

The semantic networks of word senses are composed of the sense itself and other words with a semantic relationship or collocation with the sense. Semantic

φ_i is the weight of all the 16 semantic roles, which satisfies the condition shown in Equation (7):

$$\sum_{i=1}^{16} \varphi_i = 1 \quad (7)$$

$sim_i(p_1, p_2)$ is the similarity of the content in "AccordingTo" of the definition in HowNet. The rest can follow the order of the 16 dynamic semantic roles.

After calculating the similarity and relevance results, we must select the appropriate sense for the ambiguous word according to the summation of similarity and relevance results.

5. Semantic Knowledge-based Method of Network Diagram WSD

5.1 Basic Concept of this Algorithm

The core idea of the semantic knowledge-based method of network diagram WSD is that the restricted function must be used upon ambiguous word sense selection in

relations derived from the semantic repository HowNet, which utilizes multiple semantic relations, enrich the candidate semantic network and provide an objective vocabulary for the correlation calculation between context and word senses. The latter significantly enhances the accuracy of the correlation calculation. The highest correlation is selected as the final result.

5.2 Detailed Algorithm

In this study, the semantic diagram uses an ambiguous word as the core and multiple senses of ambiguous words as the first layer of extension nodes according to the defined contents of the semantic repository HowNet. Based on the extension nodes in the first layer, outerextension nodes with the semantic content derived

from the sense are built. The sides of the diagram include the side with the semantic relationships among the sense with other entities and these are undirected. Thus, the semantic diagram constructed in this study is an undirected graph. HowNet can obtain all kinds of semantic relations using semantic roles. A total of 90 kinds of semantic roles are defined in HowNet, and the relationships in HowNet itself are also defined, such as synonymous, antisense, and hyponymy relationships. These relationships can also be used as a source of semantic relations.

Based on the aforementioned theory, this study attempts to construct a semantic relation diagram on ambiguous words. For example, the Tibetan word “གཞོན་པ་” has two major senses: ‘obstruct’ and ‘heal’. The semantic relation diagram of the Tibetan word “གཞོན་པ་” is shown in Figure 2.

The semantic relation section in Figure 2 shows the semantic relations that are directly defined by HowNet, which are partly generated using semantic roles, the implications of which are described in Table 2.

Relation	Explanation	Relation	Explanation
synonymy	synonymous	agent	agent
antonymy	antonym	patient	patient
hyponymy	hyponyms	instrument	utensil
CoEvent	event with the same role framework	location	the location of an event

Table 2. Semantic relations and their explanations

Selecting the semantic relationship from the constructed semantic diagram in Figure 2 has a theoretical basis. The theory determines whether a semantic relationship can be added in the diagram, which depends on the interpretation of the current sense (the DEF item in HowNet) and whether this semantic relationship can generate links with other concepts. For example, the CoEvent semantic relationship of sense 2 “obstruct” can join the semantic diagram because “obstruct” {obstruct | 阻止 } from the semantic role of CoEvent, “Firefighting” (whose DEF is {affairs|事务:CoEvent={obstruct|阻止:patient={fire|火}} {remove|消除 md-patient = {fire|火}}}) shares a common role of semantic relation framework concepts with the preceding concepts. The joining of other semantic relations is also built on this foundation.

For the given Sino-Tibetan bilingual corpus, using a semantic relationship to disambiguate ambiguous vocabulary must undergo the following steps after the ambiguous words are determined.

(1) Establish a semantic relationship diagram for each sense of all ambiguous words in the current Tibetan sentence, as shown in Figure 2.

(2) Calculate the semantic relevancy between each relational term of every sense under each semantic relationship in the semantic relation diagram of the target ambiguous word and every word that has not been marked as unavailable in the current calculation window. Item correlation is calculated using the method discussed in Section 4. After obtaining the semantic relevancy of each item, Equation (8) is used to integrate and obtain the

disambiguation selection parameter of the current sense $Mark(mean_i)$.

(3) Sort each disambiguation for the sense of the current ambiguous word in descending order according to Equation (8). Select the sense with the maximum disambiguation selection parameter as the ultimate sense of the current ambiguous word.

$$Mark(mean_i) = \sum_j^{size(mean_i)} \beta_j \times \sum_k^{item_num(relation_j(mean_i))} \sum_m^{len_{sen}} rel_k(w_m, item_k) \quad (8)$$

Where $Mark(mean_i)$ expresses the disambiguation selection parameter of sense $mean_i$, $size(mean_i)$ expresses the semantic relation numbers of $mean_i$ in the current semantic diagram, and β_j expresses the weight parameter of each sense in calculating the disambiguation selection parameter. The detailed evaluation is provided in the experiment result section. $relation_j(mean_i)$ expresses the j^{th} semantic relationship of $mean_i$, which was acquired from the HowNet repository. $item_num(relation_j(mean_i))$ expresses the content item numbers of $relation_j(mean_i)$, which is the number of the relation items in each blue box in Figure 2. len_{sen} expresses the length of the corresponding Chinese translation of the current Tibetan sentence (calculate the number of words in a single sentence according to the result of word separation). $rel_k(w_m, item_k)$ expresses the semantic relevancy of K^{th} word and the current relation item in the corresponding Chinese translation.

A WSD network diagram is created using Figure 2 and the pending Sino-Tibetan bilingual sentence with the

assumption that a path in the network diagram must connect the unambiguous words in the Tibetan sentence and the correct sense that corresponds to the ambiguous word. The side then connects the Tibetan words, and the right sense comprises the sense with the maximum disambiguation selection parameter. Moreover, the side connects the Tibetan words in context.

6. Experiments and Results

We established the scale of 10000 sentences Tibetan-

Chinese parallel corpora in the research process. We have not experimented on all corpus data because of time limitation. We select 757 Tibetan-Chinese sentence pairs as our experimental materials. These sentences contain 10 Tibetan ambiguous words. All the Tibetan ambiguous words and their English interpretation are shown in Table 3.

Table 3 Tibetan ambiguous words used in the experiments and their English sense

Tibetan word	Sense 1	Sense 2	Sense 3	Sense 4
རབ་འཛིན་	shelter	dim		
སྒྲིག་པ་	signal	limbs	fringed pink	
དངོས་པོ་	material	oneself		
མཁས་པ་	master	scholar		
དཀའ་བ་	tire	difficult	trouble	
ལྷང་བ་མེད་པ་	cavalier	dim		
ཤིང་གཅིག་རྒྱ་རྩམ་	single-plank bridge	Indian yellow		
གཞོན་པ་	diligent	concentrate	descend	
གཤམ་དུ་འཛིན་པ་	admit	attach importance to		
འཛིན་པ་	grasp	remember	distinguish	be reluctant to leave

Table 3. Tibetan ambiguous words used in the experiments and their English sense “གཞོན་པ་”

The aforementioned parameter values are as follows.

0.025.

In Equation (1), $\beta_1 = 0.7$, $\beta_2 = 0.2$, $\beta_3 = 0.1$. In Equations (2), (3), and (4), $\gamma = 1.6$. In Equation (6), $\varphi_{domain} = 0.3$, $\varphi_{scope} = 0.2$, $\varphi_{RelateTo} = 0.1$, $\varphi_{CoEvent} = 0.1$. Each of the remaining 12 dynamic semantic roles has a weight of

The F-score is used to describe the final experimental results. To compare the experimental results, we also used the method proposed in Section 4 on the same experimental materials. The experimental results are presented in Figure 3.

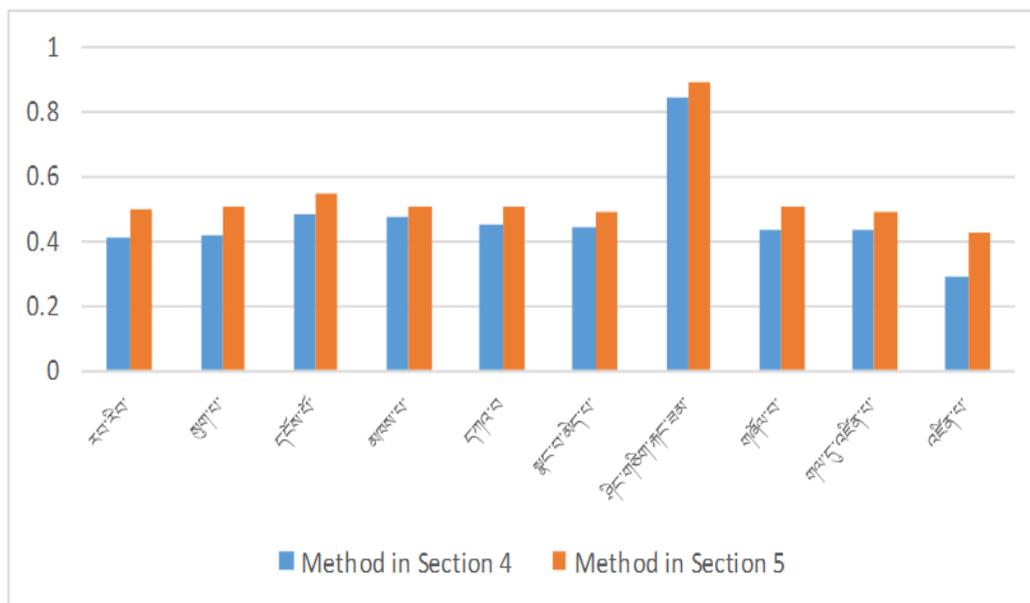


Figure 3. Experimental results

7. Conclusion and Future Work

The construction of basic resources is insufficient, particularly for the Tibetan language corpus, because research on minority language information processing lags in Chinese and English. In the case of lack of mass support for the Tibetan corpus, studying WSD using semantic knowledge is an ideal solution. This study utilizes HowNet as the knowledge source to study the Tibetan WSD method based on Tibetan-Chinese parallel corpora. The conclusions drawn from this study are as follows.

(1) This study proposes an improved sememe similarity calculation method and combines semantic distance, the lowest parent sememe height, and the height differences of sememe levels. A good calculation result is obtained.

(2) A total of 16 dynamic semantic roles are used to calculate relevance between concepts. Similarity and relevance are combined to select the appropriate sense for Tibetan ambiguous words according to Tibetan-Chinese parallel corpora.

(3) The rich semantic information from HowNet is utilized to construct semantic network diagram for every Chinese sense of the Tibetan ambiguous word. This operation can extend the calculation target of the context with sense and can alleviate the data sparseness problems to a certain extent. The experimental results also validate this conclusion.

Future work for our research includes the following aspects:

(1) Constructing large-scale Tibetan-Chinese parallel corpora to support the WSD method based on statistics. This method can solve inherent defects existing in the method based on semantic knowledge.

(2) Constructing a semantic network diagram with Tibetan words to apply our method and achieve purely Tibetan WSD. Such accomplishment can realize Tibetan WSD without the support of Chinese interpretation.

Acknowledgements

The work was supported by the National Natural Science

Foundation of China (No. 61331013) and the Program for New Century Excellent Talents in University (NCET-12-0579)

References

- [1] Navigli R., Velardi P. (2005). Structural semantic interconnections: a knowledge-based approach to word sense disambiguation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 27(7) 1075-1086.
- [2] Navigli R., Lapata M. (2010). An experimental study of graph connectivity for unsupervised word sense disambiguation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 32 (4) 678-692.
- [3] Agirre E., Soroa A., Stevenson M. (2010). Graph-based Word Sense Disambiguation of biomedical documents. *Bioinformatics*, 26 (22) 2889-2896.
- [4] Agirre E., Lacalle O., Soroa A. (2014). Random Walks for Knowledge-Based Word Sense Disambiguation. *Computational Linguistics*, 40 (1) 57-84.
- [5] Yang, Z. Z., Huang, H. Y. (2012). Graph Based Word Sense Disambiguation Method Using Distance Between Words. *Journal of Software*, 23 (4) 776-785.
- [6] Hessami E., Mahmoudi F., Jadidinejad A. H. (2011). Unsupervised Graph-based Word Sense Disambiguation Using lexical relation of WordNet. *International Journal of Computer Science Issues*, 8(6) 225-230.
- [7] Usbeck R., Ngomo A. N., Röder M., et al. (2014). AGDISTIS-Graph-Based Disambiguation of Named Entities Using Linked Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (4) 791-804.
- [8] Tseng H., Chang P., Andrew G., et al. (2005). A Conditional Random Field Word Segmenter for Sighan Bakeoff 2005. *In: Proceeding of the Fourth Sighan Workshop on Chinese Language Processing, Jeju Island, Korea: ACL*, p. 168-171.
- [9] Kang C. J., Jiang D., Long C. J. (2013). Tibetan Word Segmentation Based on Word-Position Tagging. *In: 2013 International Conference on Asian Language Processing (IALP), Urumqi, China: IEEE*, p. 243-246.
- [10] Xia T. (2007). Study on Chinese Words Semantic Similarity Computation. *Computer Engineering*, 33(6) 191-194.