Linking Entities in Short Texts Based on a Chinese Semantic Knowledge Base

Yi Zeng, Dongsheng Wang, Tielin Zhang, Hao Wang, and Hongwei Hao

Institute of Automation, Chinese Academy of Sciences, Beijing, China
{yi.zeng,dongsheng.wang,hongwei.hao}@ia.ac.cn

Abstract. Populating existing knowledge base with new facts is important to keep the knowledge base fresh and most updated. Before importing new knowledge into the knowledge base, entity linking is required so that the entities in the new knowledge can be linked to the entities in the knowledge base. During this process, entity disambiguation is the most challenging task. There have been many studies on leveraging name ambiguity problem via a variety of algorithms. In this paper, we propose an entity linking method based on Chinese Semantic Knowledge where entity disambiguation can be addressed by retrieving a variety of semantic relations and analyzing the corresponding documents with similarity measurement. Based on the proposed method, we developed CASIA_EL, a system for linking entities with knowledge bases. We validate the proposed method by linking 1232 entities mined from Sina Weibo to a Chinese Semantic knowledge base, resulting in an accuracy of 88.5%. The results show that the CASIA_EL system and the proposed algorithm are potentially effective.

Keywords: Entity linking, Chinese Semantic Knowledge, Semantic similarity, Entity disambiguation.

1 Introduction

This paper introduces our results on the competition of entity linking task organized by the second conference on Natural Language Processing and Chinese Computing (NLP&CC 2013). The task of entity linking is to link the entity name from Sina Weibo microblog posts to Baidu Encyclopedia knowledge base. The organizer provides a part of Baidu Encyclopedia knowledge base and a number of Sina Weibo microblog posts within which entities are tagged. If the entities in the microblog posts can be linked to the entities in the knowledge base, we are supposed to output the identity of the entity in the knowledge base (denoted as "*KB_ID*"). If the knowledge base does not contain the given entity, we should output a "*NIL*". We build up an entity linking system called CASIA_EL according to the task requirement and produce one group of result. Finally, we ranked the second place among 10 teams with an overall linking accuracy of 88.5%.

This competition focuses on linking entities from short texts (e.g. microblog posts) to knowledge bases. The motivation of this work is to enable and evaluate possible

G. Zhou et al. (Eds.): NLPCC 2013, CCIS 400, pp. 266–276, 2013.

techniques to enrich existing knowledge base with dynamic and new knowledge from short texts based social media such as microblog. One of the most important step is linking possible entities from microblog posts with entities in the existing knowledge base, so that relevant entities within and outside the knowledge base can be linked together and organized as interconnected Web of knowledge. Some key issues need to be considered for the entity linking task. On the one hand, a specific entity may have various names while they actually refer to the same entity. In this case, we need processing techniques (such as creating and maintaining a synonym set) to link possible entity names together. On the other hand, different entities may share the same name. In this case, entity disambiguation algorithms are needed to link the entity from short texts with appropriate candidate entity in the knowledge base.

This competition is challenging in several perspectives. Firstly, microblog posts are very short (only 140 characters are allowed), and traditional entity linking algorithm may not perform well since they may be short of enough meaningful contents. Secondly, the representation of entity names and language style in microblog posts are extremely free, which brings many potential difficulties for linking the entity names to appropriate entities in the knowledge base.

In this paper, we propose an entity linking method and implement an entity linking system, named CASIA_EL, by utilizing various semantic relations among entities in a Chinese semantic knowledge base with knowledge from various sources. First of all, we construct the Chinese semantic knowledge base from scratch by using semantic techniques including RDF, N3, triple store, etc. In order to support the linking process, we automatically construct a synonym set from multiple Web-based wiki-encyclopedias. Secondly, in order to make a balance between the contents of a microblog post and the wiki page contents in the knowledge base, we attempt to make an understanding of microblog posts by adding relevant contents from the knowledge base. For example, the post "*My Company is bad, I like Apple*" is one single sentence with just two nouns while the description of "*Apple Inc.*" in the knowledge base may contain several hundred sentences. Therefore, when it is hard to find appropriate entities in the knowledge base, we extend the post by adding other documents of associated entities within the same context. For this example, we add the wiki page contents of "*Company*" from the knowledge base to this short post. This method is formalized as the stepwise bag-of-words entity disambiguation algorithm. With the experimental results on the competition, we conclude that the proposed entity linking method and the system is potentially effective.

The rest of the paper is organized as follows: Section 2 introduces related works of this paper. Section 3 introduces the constructed Chinese semantic knowledge base and the entity linking method. Section 4 provides evaluation results and detailed analysis on the performance of our system.

2 Related Works

Here we define the entity name from the microblog post as the target entity, and the possibly identical entities from the knowledge base as the candidate entity. The task of entity linking generally focuses on the problem of entity disambiguation that has been discussed extensively in different areas. There have been many studies addressing this issue, and one of the most widely used method is the Bag-of-Words (BOW) model [2, 3], which utilizes a bag of words to express the entity's document and calculate the similarity via vector cosine similarity. Some researchers extend the work by adding social network relations to obtain background knowledge. For instance, Bekkerman and McCallum utilized the inter-link and out-link structures of web pages as well as the similarity between two persons' page documents for disambiguation [4]. A graph-based framework was adopted by Jiang et al. to disambiguate person entity appeared on the Web by capturing personal information [5]. However, both the BOW method and the social-graph based method suffer the insufficiency of corpus or information.

Recently, because of the public availability of Wiki Encyclopedias and knowledge bases such as Wikipedia, DBpedia, and YAGO [1], some researchers take advantages of these sources, which contain large and abundant knowledge, to leverage entity disambiguation. For example, Han and Zhao constructed the semantic knowledge from Wikipedia, based on which the semantic similarity measurement is proposed [6]. Shen et al. utilized the taxonomy information in YAGO, a semantic knowledge base which is built based on Wikipedia and WordNet, to obtain agreement on the categories associated with candidate entities [7]. Compared to the efforts in English, Chinese entity linking has not been well studied and needs further investigations. Some recent progresses include Zhishi.me (a multi-source Chinese Web of Data platform based on infobox knowledge from Baidu Encyclopedia, Hudong Encyclopedia and Wikipedia Chinese sources) [8] and cross-lingual knowledge linking among different Chinese Knowledge Bases [9].

3 Chinese Semantic Knowledge Based Entity Linking

In this paper, we propose to investigate on the Chinese entity linking task through analysis and utilization of various semantic relations among entities from the Chinese Semantic Knowledge Base. In this section, we firstly introduce the fundamental work on how the Chinese Semantic Knowledge base is constructed from scratch in Section 3.1. Then we discuss techniques for linking unambiguous entities in Section 3.2. We detail our proposal on stepwise entity disambiguation based on the Chinese semantic knowledge base in Section 3.3.

3.1 Semantic Knowledge Base Construction

The competition organizer provides a portion of Baidu Encyclopedia contents in the form of XML, which are composed of entities, corresponding infobox knowledge, and unstructured wiki-page contents. We extract the infobox knowledge from the XML file and represent them in *N3* format. An example is provided in Table 1.

Table 1. An Example on Representing Infobox Knowledge in N3 Format

Infobox Knowledge	苹果公司, 外文名称, Apple Inc.
N3 format	http://www.ia.cas.cn/baidu_baike/resource/KBBD010956, http://www.ia.cas.cn/baidu_baike/resource/外文名称, "Apple Inc."^^xsd:string. http://www.ia.cas.cn/baidu_baike/resource/KBBD000001, http://www.w3.org/2000/01/rdf-schema#label "苹果公司"^^xsd:string.

For knowledge represented in N3 format, every entity is identified by an URL containing a unique knowledge base ID (denoted as "*KB_ID*"). The corresponding properties are expressed via a variety of relations such as the URI containing "外文名称" in the upper example.

Instead of traditional storage methodology which is generally based on relational databases, we employ semantic storage methodology based on Jena TDB[1] and the data are represented in RDF N3 format. Since the RDF data can be explained as a graph, consisting of interconnected nodes and relations, it is flexible to add new schema and new data (i.e. nodes and relations) without modification of the original database design.

In this Chinese knowledge base, every entity resource is with a variety of relations. The most significant relation is the "r*dfs*:*label*," by which a bag of possible input keywords is created for a given entity resource. The labels are generated from entity *name*, *English name*, *Chinese name* from the infobox knowledge. In addition, we extend the labels by synset, *nick names* and *redirect titles* crawled from Baidu Encyclopedia, Hudong Encyclopedia and Wikipedia Chinese pages (476,086 pair of synonyms are added). Also, we split the western people's name by "." into smaller keywords, for example, "*Michael·Jordan*" is split into "*Michael*" and "*Jordan*" and both of them are added as possible labels for "*Michael Jordan*". In this way, we enlarge the coverage of keywords for a given entity resource to support various entity names written by different user with different preference. Consequently, it is common phenomenon that several entity resources can share the same label. For example, as shown in Fig. 1, the Company "苹果公司(Apple Inc.)" and the fruit "苹果(Apple)" share the same label "苹果". In other words, given the entity name "苹果" in a microblog post, our system firstly returns two or more candidate entity resources from the knowledge base and they are delivered to entity disambiguation processing. The entity disambiguation process will be discussed in Section 3.3.

[1] `http://jena.apache.org/documentation/tdb/`

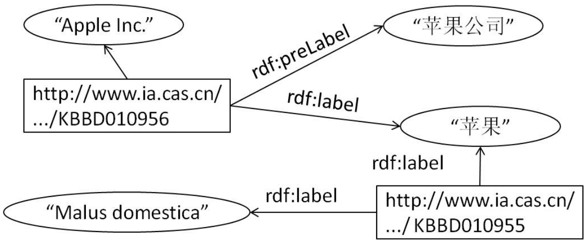

Fig. 1. An Example of Connected Semantic Data

Since the number of labels of an entity resource can be very large, we create a sub-relation "*CAS*:*preLabel*" (preferred label) of "*rdfs*:*label*" to express the representative keywords for an entity resource. Namely, we entitle higher priority to the relation "*CAS*:*preLabel*" than "*rdfs*:*label*". If the entity name is consistent with both "*preLabel*" and "*rdfs*:*label*" values, our system would performs linking to the corresponding "*preLabel*" entity resource first.

3.2 Linking Unambiguous Entities

As discussed in section 3.1, given an entity name appeared in a piece of short text, the retrieval of "*rdfs:label*" can result in one, several, or even no candidate entity resources returned from existing knowledge base.

For the case that only one single candidate entity resource was returned, we link the entity name with the corresponding entity resource in the knowledge base directly (In this case, the entity is unambiguous), and outputs the *KB_ID* of the entity.

For the case that there are no candidate entity returns, it is not clear whether there is no corresponding entity in the knowledge base or there might be typos for the entity name. Hence, we automatically deliver the entity name to Google's "*Did you mean * ?*" function to check whether it's a wrong spelling keyword (In the upper sentence, * is a possible correct alternative for the original entity name). If it returns possible new keywords, we change the entity name to the proposed new literal and process the new keyword again for the entity linking process. Otherwise, we assume there is no corresponding candidate entity in the knowledge base, and we produce the output result as "*NIL*".

In addition, considering the characteristics of Chinese entities, we made some extra processing if the original entity name cannot be directly found in the knowledge base. For example, in Chinese, many entities are with punctuation marks around them, such as book names, film, TV program names, etc. Hence, before a keyword is delivered to process, we check if it contains characters in the following set: {<, >, 《, 》, ", ","," , …}. If it contains, we process it as it is since the synonym set might be able to handle it, and if it returns "*NIL*", we delete them and re-process the keywords again. For example, " 《霸王别姬》 " is delivered to process as it is and if it outputs a "*NIL*," we split it into "霸王别姬" and process it again. For person names, when writing microblog posts, the character "·" in the middle of the given name and the family name is always replaced with a space or "-".

For example, “勒布朗·詹姆斯” is the entity name in the knowledge base, but in microblog posts, it is referred to as “勒布朗-詹姆斯” or “勒布朗 詹姆斯”. In our system, we firstly process the entity as it is, and if there are no relevant *KB_ID*, the space or “-” are replaced by “·” and the entity name is reprocessed again on the knowledge base.

3.3 Stepwise Entity Disambiguation Based on a Chinese Knowledge Base

If there are several candidate entity resources in the knowledge base that may be relevant to the target entity in the piece of short text, then we need the entity disambiguation process to decide which candidate should be linked to. Since every entity in Baidu encyclopedia corresponds to a wiki page (or a part of a wiki page) that is used to describe this entity, we can use these contents for the entity disambiguation process. Our general design is describing the target entity and the candidate entity with their contexts respectively, and then we compare the similarity of these two contexts to obtain the similarity between the target entity and the candidate entity. Based on these considerations, we propose the Stepwise Bag-of-Words (S-BOW) based Entity Disambiguation Algorithm, as shown in the following.

Algorithm 1: The Stepwise Bag-of-Words based Entity Disambiguation Algorithm

Input: target entity name (e_{st}) and the short text (d_{st}), semantic knowledge base (KB)

Output: *KB_ID* from the KB, which corresponds to e_{st}

1 Begin
2 Generate bags of word terms (denoted as $bag[d_{st}]$ and $bag[d_{kb}]$, only noun and entity type word terms are selected to put in the bags) which are used to describe d_{st} and several d_{kb}.
3 Get the intersection of $bag[d_{st}]$ and $bag[d_{kb}]$, which contains word terms that $bag[d_{st}]$ and $bag[d_{kb}]$ share, and obtain $sim(d_{st}, d_{kb})$ for each d_{kb}.
4 Get the biggest $sim(d_{st}, d_{kb})$ and output the corresponding *KB_ID*.
5 If more than one $sim(d_{st}, d_{kb})$ share the same value, extend $bag[d_{st}]$ to $bag'[d_{st}]$ and calculate $sim'(d_{st}, d_{kb})$, else go to Step 7.
6 Get the biggest $sim'(d_{st}, d_{kb})$ and output the corresponding *KB_ID*.
7 End

Here we briefly explain the general steps in this algorithm. Firstly, we get a piece of short text (denoted as d_{st}) which contains the target entity name (In this paper, a piece of short text refers to a microblog post from Sina Weibo.) and the contents that describe possible candidate entity resources from knowledge base (denoted as $d_{kb.}$ In this paper, the document corresponding to a candidate entity resource in the knowledge base is a (or part of) the wiki page from Baidu Encyclopedia). Then, we measure the similarity values between d_{st} and each d_{kb} based on the bag of words used to describe d_{st} and d_{kb}. The similarity value between d_{st} and each d_{kb}, is characterized as the number of shared word terms in both $bag[d_{st}]$ and $bag[d_{kb}]$, namely:

$$sim(d_{st}, d_{kb}) = |\, bag[d_{st}] \cap bag[d_{kb}] \,|. \quad (1)$$

The d_{kb} with the highest similarity value is selected and the target entity is linked to the corresponding entity resource in the knowledge base. However, in some cases, two or more d_{kb} still share the same similarity value with the d_{st}, and it is hard to select out only one of them. This often happens due to the short length of d_{st}, which leads to a relatively smaller number of word terms in $bag[d_{st}]$. On the contrary, the d_{kb} is generally very long and the amount of word terms in $bag[d_{kb}]$ is almost always big. In order to make a balance among the size of $bag[d_{st}]$ and $bag[d_{kb}]$, we enrich $bag[d_{st}]$ by adding bags of word terms that are used to describe each word terms in $bag[d_{st}]$.

$$bag'[d_{st}] = bag[d_{st}] \cup bag[t_1] \cup bag[t_2] \cup \ldots \cup bag[t_n]. \quad (2)$$

where $bag'[d_{st}]$ denotes the extended bag of word terms. $\{ t_1, t_2, \ldots, t_n\}$ is a set of terms obtained from the short text d_{st}. $bag[t_n]$ denotes the bag of word terms obtained from the wiki page which uses the term t_n as its title in the semantic knowledge base.

After extending the bag of word terms for d_{st} from $bag[d_{st}]$ to $bag'[d_{st}]$, the similarity among d_{st} and d_{kb} is recalculated according to the intersection of $bag'[d_{st}]$ and $bag[d_{kb}]$, as shown in the following formula:

$$\begin{aligned} sim'(d_{st}, d_{kb}) &= |\, bag'[d_{st}] \cap bag[d_{kb}] \,| \\ &= |\, (bag[d_{st}] \cup bag[t_1] \cup bag[t_2] \cup \ldots \cup bag[t_n]) \cap bag[d_{kb}] \,|. \end{aligned} \quad (3)$$

The selection criteria for candidate entity is the same, and the *KB_ID* with the largest $sim'(d_{st}, d_{kb})$ value is produced as the output.

This algorithm uses the idea of bag-of-words model in a stepwise manner. Nevertheless, the extension of the original short text only happens when there are more than one $sim(d_{st}, d_{kb})$ which share the same value. The next section will experimentally verify the proposed algorithm by using the competition data.

4 Experimental Results and Analysis

The NLP&CC 2013 Entity Linking contest provides 779 microblog posts within which 1232 entities are tagged, and these entities are required to be linked to the entities in the knowledge base. Therefore, it is not necessary to detect entity names automatically from the microblog posts, and the focus is on the linking process. According to the results produced by the organizer, the total amount of accurate output is 731, and the overall precision is 0.885. Detailed results are listed in Table 2.

Table 2. Evaluation Result Produced by the Competition Organizer

Accurate Output	precision	In-KB precision	In-KB recall	In-KB F1	NIL precision	NIL recall	NIL F1
731	0.885	0.8662	0.8456	0.8558	0.9036	0.9260	0.9146

The whole process of our entity linking system is based on the method discussed in Section 3. Here we extend our discussion based on the experimental results.

A key step for the S-BOW entity disambiguation algorithm is the generation of the bags of word terms that are used to describe the contexts of the target entity and the candidate entity. We use NLPIR [2], a refined version of ICTCLAS [3] for word segmentation and lexical category tagging. Table 3 provides practical reasons why noun and literal string are selected as basic elements to generate bag of word terms.

Table 3. The Effects of Lexical Categories on the Correctness of Entity Disambiguation

Lexical Category	Labels in ICTCLAS	Entity Disambiguation Correctness
noun	n, nr, nr1, nr2, nrj, nrf, ns, nt, nz, nl	0.75
verb	v, vd, vn, vf, vx, vi, vl, vg	0.685
adjective	a, ad, an, ag, al	0.48
literal string	x, xx, xu	0.714

We make additional test on the entity disambiguation correctness when only the specified lexical category is selected as the source to generate the bag of word terms. Here we can find that noun and literal string (usually entity type word terms, such as “GDP”, “CEO”, “KFC”, “DVD”, etc.) are better sources for generating the bag of word terms. Hence, these two lexical categories are selected for creating the contexts of target entity and candidate entity in the semantic knowledge base. In addition, we examined the effect of word term length on the correctness of entity disambiguation, and the result is listed in Table 4.

Table 4. The Effects of Word Term Length on the Correctness of Entity Disambiguation

Length	Entity Disambiguation Correctness
All lengths	0.655
Equal to or longer than 2 Chinese characters	0.776

According to Table 4, we can conclude that it can lead to negative effect when we involve the word terms that only contain one Chinese character. As a consequence, only the noun and literal string type of word terms whose length is greater than 1 are selected to represent the contexts of the target entity and the candidate entity.

During the design of the S-BOW entity disambiguation algorithm, we assume that extending the number of word terms can make a balance between the contexts of target entity and candidate entity. Fig. 2 provides an experimental evidence on the assumption, which reflects a general positive relevance between the number of entities in the context and the correctness of entity disambiguation.

[2] The NLPIR Chinese Word Segmentation System: `http://ictclas.nlpir.org/`
[3] The ICTCLAS Chinese Word Segmentation System: `http://ictclas.org`

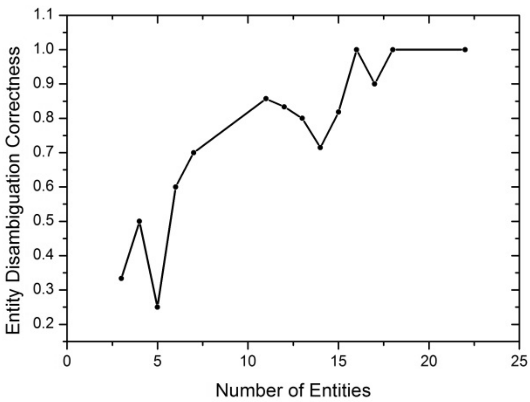

Fig. 2. The Relation Between the Number of Entities and the Entity Disambiguation Correctness

Although it seems that the number of candidate entities may not have any clear correlation with the entity disambiguation correctness, our system produce even better results when the number of candidate entities goes larger, as shown in Fig. 3.

Due to the reason that we add many synonyms to the original entity resource in the knowledge base (e.g. the family name and the given name of a person are added as synonyms for this person). There are many cases that the amount of candidate entities reaches more than 10, or even 50, as shown in Table 5 (with 56 candidate entities) and Fig. 4. Even in this case, the entity disambiguation process performs very well. It is noted that when the number is greater than 9, there are no incorrect disambiguations.

Table 5. An Example of Multiple Candidate Entities in the Knowledge Base

Target Entity and the Microblog Post	Candidate Entities	Name	Produced KB_ID
weibo id = aonierqiuyituiyi914 name id = 詹姆斯 content = “奥尼尔球衣退役了，突然联想到如果詹姆斯以后退役了，克里夫兰会退役他的球衣吗？？？？？？”	KBBD000035	詹姆斯·普雷斯科特·焦耳	KBBD000092
	KBBD000092	勒布朗·詹姆斯	
	KBBD000609	詹姆斯·西蒙斯	
	KBBD000707	詹姆斯·克拉克·麦克斯韦	
	KBBD000875	詹姆斯·弗兰克	
	KBBD000876	詹姆斯·弗兰克	
	…	…	
	KBBD018850	詹姆斯·瓦特	

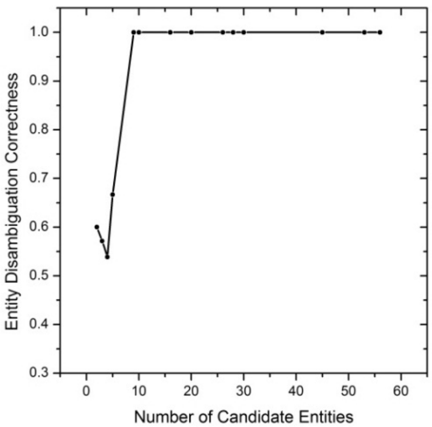

Fig. 3. The Relation Between the Number of Candidate Entities and the Entity Disambiguation Correctness in CASIA_EL

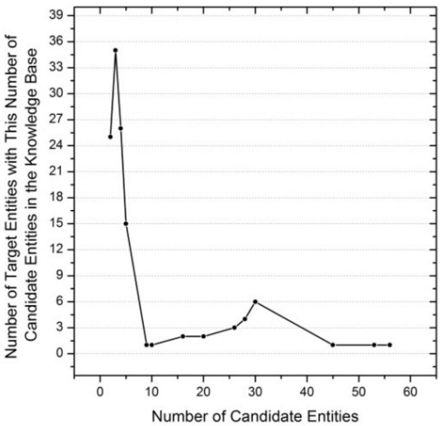

Fig. 4. Ambiguous Entity Number Distribution in the Test Data

Our system detected 161 entity disambiguation tasks in the test data. When we only try to disambiguate entities by using the original microblog posts, 123 entities were disambiguated, and the correctness is 82.1% (101 out of 123 entities were correctely disambiguated). When we extend the original microblog posts by adding the Baidu Encyclopedia contents (Following Step 5 and 6 in Algorithm 1), another 38 entities were disambiguated. For the second phase, the correctness is 63.2% (24 out of 38 entities were correctly disambiguated). Hence, the overall entity disambiguation correctness based on the proposed algorithm is 77.6% (125 out of 161 entities are correctly disambiguated).

5 Conclusion

Updating existing knowledge base with new facts is important to maintain the knowledge base and keep it fresh and updated, and the importing of new facts from the real-world requires entity linking in advance with existing knowledge base, challenging us with the named entity disambiguation.

There have been many studies on addressing entity disambiguation problem via a variety of algorithms. They are insufficient with information and some methods that adopt knowledge bases are less well-structured. In this paper, we proposed an entity linking method based on Chinese Semantic Knowledge where the named entity disambiguation can be addressed by retrieving a variety of semantic relations and analyzing the corresponding documents with similarity measurement. What's more, for the short texts (e.g. microblog posts) which are unfeasible for the similarity measurement, we extend the content by adding other documents of entities those that are within the same context. The proposed method was validated by linking 1232 entities, mined from Weibo, to Chinese Semantic knowledge base, resulting in an accuracy of 88.5%, demonstrating a satisfactory linking efficiency.

In this paper, we only considered extending short texts by adding relative contents from Baidu Encyclopedia pages, while the infobox knowledge of each candidate entity has not been taken into account. Intuitively, properties and property values are important for disambiguate one entity from the others. In the future, combining both unstructured contents and infobox knowledge in the wiki-pages seems promising for entity disambiguation and entity linking.

Acknowledgement. This study was supported by “the Young Scientists Fund” of the National Natural Science Foundation of China (61100128).

References

1. Suchanek, F.M., Kasneci, G., Weikum, G.: YAGO: A Large Ontology from Wikipedia and WordNet. Journal of Web Semantics 6(3), 203–217 (2008)
2. Bagga, A., Baldwin, B.: Entity-based Cross-document Coreferencing Using the Vector Space Model. In: Proceedings of the 17th International Conference on Computational Linguistics (COLING 1998), pp. 79–85. ACL, Montreal (1998)
3. Mann, G.S., Yarowsky, D.: Unsupervised personal name disambiguation. In: Proceedings of the 7th Conference on Natural Language Learning (CONLL 2003), pp. 33–40. ACL, Edmonton (2003)
4. Bekkerman, R., McCallum, A.: Disambiguating Web appearances of people in a social network. In: Proceedings of the 14th International Conference on the World Wide Web (WWW 2005), pp. 463–470. ACM Press, Chiba (2005)
5. Jiang, L., Wang, J., An, N., Wang, S., Zhan, J., Li, L.: GRAPE: A Graph-Based Framework for Disambiguating People Appearances in Web Search. In: Proceedings of the 9th IEEE International Conference on Data Mining (ICDM 2009), pp. 199–208. IEEE Press (2009)
6. Han, X., Zhao, J.: Named entity disambiguation by leveraging Wikipedia semantic knowledge. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM 2009), pp. 215–224. ACM Press, Hong Kong (2009)
7. Shen, W., Wang, J., Luo, P., Wang, M.: LINDEN: linking named entities with knowledge base via semantic knowledge. In: Proceedings of the 21st International Conference on the World Wide Web (WWW 2012), pp. 449–458. ACM Press, Lyon (2012)
8. Niu, X., Sun, X., Wang, H., Rong, S., Qi, G., Yu, Y.: Zhishi.me - Weaving Chinese Linking Open Data. In: Aroyo, L., Welty, C., Alani, H., Taylor, J., Bernstein, A., Kagal, L., Noy, N., Blomqvist, E. (eds.) ISWC 2011, Part II. LNCS, vol. 7032, pp. 205–220. Springer, Heidelberg (2011)
9. Wang, Z., Li, J., Wang, Z., Tang, J.: Cross-lingual Knowledge Linking across Wiki Knowledge Bases. In: Proceedings of the 21st World Wide Web Conference (WWW 2012), pp. 459–468. ACM Press, Lyon (2012)