Exploiting Entity-Level Morphology to Chinese Nested Named Entity Recognition

Chunyuan Fu, Yanqing Zhao, Guohong Fu
School of Computer Science and Technology, Heilongjiang University
Harbin 150080, China
Email: fuchunyuan@yahoo.cn, yanqing_zhao@live.cn, ghfu@hlju.edu.cn

Abstract

Named entity recognition plays an important role in many natural language processing applications. While considerable attention has been pain in the past to research issues related to named entity recognition, few studies have been reported on the recognition of nested named entities. This paper presents a morpheme-based due-layer labeling method to Chinese nested named entity recognition. To approach this task, we first employ the logistic regression model to extract multi-level entity morphemes from an entity-tagged corpus, and thus explore multiple features, particularly entity-level morphological cues for Chinese nested named entity recognition under the framework of conditional random fields. Our experimental results on different datasets show that our system is effective for most nested named entities under evaluation, illustrating the benefits of using entity-level morphology.

Keywords

Named entity recognition, Chinese nested named entity, Entity morpheme, Conditional random fields

1. Introduction

Named entity recognition (NER) seeks to identify and classify information units like names (viz. person, location and organization names) and numeric expressions (viz. date, time, money and percent expressions) in text. As an important subtask of information extraction, NER has received considerable attention from the natural language processing community over the past years. It has been a shared task of a number of conferences, including Message Understanding Conferences (Chinchor, 1999), Automatic Content Extraction (ACE), the Conferences on Natural Language Learning (CoNLL) and the Bakeoffs hosted

by the Special Interest Group of the Association for Computational Linguistics on Chinese Language Processing (SIGHAN Bakeoffs).

These shared tasks have greatly promoted the development of NER technology, especially English NER technology. However, to develop a practical NER system for Chinese is still a challenge. On the one hand, Chinese NER is usually formalized as a sequence labeling task with characters or words as the basic tokens. However, it is difficult to handle entity-internal structural features for Chinese NER based on either characters or words (Fu, 2009). On the other hand, while the recognition of simple named entities (NEs) is well studied over the past years, few works have been reported on the identification of nested NEs, which might be the major source of errors for existing systems for Chinese NER.

To address the above issues, this paper introduces the concept of entity morphemes into Chinese NER and presents a morpheme-based dual-layer chunking method to Chinese nested NER. To this end, we first employ the logistic regression model to extract multi-level entity prefixes and suffixes from an entity-tagged corpus, and thus explore a combination of multiple features, particularly entity-level morphological features for Chinese nested NER under the framework of conditional random fields (CRFs). Our experimental results show that the introduction of entity morphemes is beneficial to the improvement of Chinese NER performance.

This paper is structured as follows: Section 2 provides a brief overview of the related work on NER. In section 3, we analyze the structures of Chinese nested named entities. Section 4 describes a logistic transformation-based technique for entity morpheme extraction. Section 5 details the morpheme-based dual layer labeling approach to Chinese nested NER and. We report our experimental results in section 6 and finally conclude our work in section 7.

2. Related work

Over the past years, a number of machine learning approaches have been attempted for NER, such as hidden Markov models (HMMs) (Fu and Luke, 2005; Fu, 2009), maximum entropy models (MEMs) (Saha et al, 2008), support vector machines (SVMs) (Ekbal and Bandyopadhyay, 2008), conditional random fields (CRFs) (Liu et al, 2010). Zhang et al. (2008) explored multiple features, including both local and global constraint information, and thus combined them for Chinese NER under the ME framework. To reduce to reduce search space, they also introduced heuristic knowledge into their system. They showed that their system can achieve an F-score of 86.31% over the SIGHAN Bakeoff 2008 dataset. Tsai et al. (2005) employed the ME model to exploited multiple shallow linguistic information such as spelling, parts of speech, word forms, context and other features for biological NER. However, the performance for named entities with complex structures is not satisfactory. CRF proved to be more effective for sequence labeling tasks (Lafferty et al., 2001). Recently, a number of studies employ CRF to perform NER in different languages

or domains. Wang (2009) proposed a two-stage method that incorporates rules with CRFs to perform biological NER. More recently, She and Zhang (2010) incorporated CRFs with MNE rules for musical NER.

Nest NER is an important but difficult issue in the field of NER. While considerable attention has been pain in the past to research issues related to named entity recognition, there have been few attempts to investigate the problem of nested named entity recognition. Alex et al. (2007) applied a CRF-layer model to English NER, which first identifies the simple NEs embedded in nested NEs, and then recognizes other NEs. Finkel and Manning (2009) proposed a discriminatory selection algorithm to train a structural model for English nested NER. However, the method did not work well in news reports and biomedical field. Compared with other methods, it is also time-consuming.

This paper is primarily concerned with nested named entity recognition in Chinese news text. In particular, we attempt to exploit entity-level morphological features for Chinese nested NER. The purpose of this study is to determine if the use of entity-level morphological information would improve Chinese nested NER performance.

3. Nested named entities in Chinese

To investigate the structural characteristics of Chinese nested named entities, we use an entity-tagged version of the PKU corpus (Fu and Luke 2005). This corpus consists of one month of news text from the People's Daily, which has been annotated with 46 different part-of-speech tags and thirteen different named entity tags, respectively. It contains a total of 106430 named entities. In the present study, we focus on the three main named entities, namely person names, location names and organization names.

We analyze the form of nested named entities and educe three types as follows:

(1) A same type of NEs paratactically embedded in one named entity: In this case, a named entity contains a same type of entities, and there is no hierarchical relationship between these embedded NEs.

(2) Different types of NEs paratactically embedded in one named entity: In this case, a named entity involves different types of entities, and there is no hierarchical relationship between these embedded NEs.

(3) Various types of named entities nested in one named entity: In this case, a named entity consists of multiple entities and there is a hierarchical relationship between these embedded NEs.

Table 1 presents the structures of some typical nested Chinese named entities. We can see from this table that a nested Chinese organization name usually begin with a location name or an organization name as its prefix indicating its respective location or parent institute and ends with a suffix such as 大学 'university', 医院 'hospital' and 大使馆 'embassy' indicating the type of organization. With respect to its middle parts, they are made up of one or more person names, location names, organization names or other words. As for nested location names, they are observed to have similar characteristics. In short, most nested NEs begin

with an entity prefix and end with an entity suffix. Entity prefixes can be a person name, location name or organization name while entity suffixes such as 部 'department' usually indicate some attributions of NEs like types and administrative level.

NO.	Named entities	Nested structures
1	黑龙江大学 (Heilongjiang University)	[黑龙江/ns]LOC 大学]ORG
2	湛江市惠珍联合医院 (United Hospital of Huizhen in Zhanjiang City)	[[湛江市/ns]LOC [惠珍/nr]PER 联合/v 医院/n]ORG
3	中国驻南非大使馆 (Chinese Embassy in South Africa)	[[中国/ns]LOC 驻/v [南非/ns]LOC 大使馆/n]ORG
4	纽约联合国总部 (United Nations Headquarters in New York)	[[纽约/ns]LOC [[联合国/E-nt]ORG 总部/n]LOC]LOC
5	索非亚大教堂 (*Sofia Cathedral*)	[[索非亚/nr]PER 大/a 教堂/n]LOC

Table 1: Structures of some typical nested Chinese named entities

In addition, nested NEs usually have a complex hierarchy structure. Thereby, we can further distinguish different nested NEs in terms of their number of nested levels. Table 2 presents the distribution of different NEs with the number of nested levels.

NE type	Number of nesting level	Number	Percentage
Simple NEs	One-level	35124	81.5%
Nested NEs	Two-level	6864	16.0%
	Three-level	1046	2.4%
	Four-level	83	0.1%

Table 2: Distribution of different named entities in terms of nested levels

As illustrated in Table 2, we can see that more than 18% NEs in the corpus have nested structures, showing again that nested NEs are very common in Chinese. Furthermore, as can be seen from Table 2 that be most nested NEs have two-level nested structures.

4. Entity morpheme extraction

As we have discussed above, most nested named entities in Chinese begin with a prefix and end with a suffix. Such morphological information is obviously an important source of indicators for nested NER. However, there are some words or morphemes that usually occur before or after named entities but hardly appear as part of nested named entities. In

other words, we need to distinguish useless entity morphemes from informative entity morphemes while exploring morphological features for nested NER.

To this end, we first employ the logistic transform method in the logistic regression model to calculate the importance of some potential entity morphemes, and thus extract a set of useful entity prefixes and suffixes for nested NER. The details of logistic transformation can be seen in (Ashton, 1972).

Entity morpheme	Frequency		Probability
	Total number of entity morphemes in nested NEs	Total number of morphemes	
w_1 ⋮ w_k	m_1 ⋮ m_k	n_1 ⋮ n_k	p_1 ⋮ p_k
$\sum_i$	$m_\bullet$	$n_\bullet$	$P_\bullet$

Table 3: The contingency table for entity morpheme extraction

First of all, we construct a two-dimension contingency table (as shown in Table 3) from the training data. Where, $W = w_1\ w_2\ \cdots\ w_k$ denotes a set of entity morphemes extracted from the corpus, m_i ($1 \le i \le k$) is the number of the entity morpheme w_i that constitutes nested named entities, n_i ($1 \le i \le k$) denotes the total number of morphemes in the training data, and p_i is the relevant probability of the entity morpheme w_i constituting nested named entities.

In practice, the goal of logistic transformation to convert a linear scale into a probability measurement between 0 and 1. From the mathematical point of view, the value of the probability p ranges from 0 to 1. So the relationship between the independent variable and p is hard to describe by linear model. Furthermore, it is also hard to find and deal with the tiny change when the value of p is close to 0 or 1. Under this situation, we do not deal with p directly. Instead, we count a strictly monotonic function of p, namely $Q = Q(p)$, where $Q(p)$ is sensitive enough to small changes when p trends to 0 or 1. That is, dQ/dp should be directly proportional to $p/(1-p)$. The function Q can be formally defined as:

$$Q = Logit(P) = \ln(\frac{P}{1-P}) \tag{1}$$

When p changes from 0 to 1, the range of Q is $(-\infty, +\infty)$. For each entity morpheme w_i, we use P^* to represent Q. Thus, Formula (1) can be further rewritten as

$$P_i^* = Logit(P) = \ln(\frac{P_i}{1-P_i}) \quad (2)$$

The formula requires $P_i^* \neq 0$ or $P_i^* \neq 1$. In other words, $m_i \neq 0$ and $m_i \neq n_i$. So, if $m_i = 0$ or $m_i = n_i$, we must amend P_i^* using formula (3).

$$P_i = \frac{m_i + 0.5}{n_i + 1} \quad (3)$$

In this way, we can figure out the weight of each entity prefix and suffix and thus calculate their importance in forming nested NEs. Table 4 and Table 5 present the respective weights of some typical entity prefixes and suffixes based on logistic transformation.

Useful prefixes	Weights	Useless prefixes	Weights
中东 'Middle East'	5.5255	世界 'world'	-2.9435
美国 'America'	5.5636	华夏 'China'	-2.5649
淮 'Abbreviation of the Huai River'	4.4427	国家 'nation'	-1.5063
空军 'air force'	3.9120	中 'Abbreviation of China'	-1.1570
南亚 'South Asia'	2.5649	首都 'Capital'	-2.3573

Table 4: The weights of some typical entity prefixes based on logistic transformation

Useful suffixes	Weights	Useless suffixes	Weights
议会 'parliament'	3.7613	学校 'school'	-1.6831
委 'committee'	3.9195	所 'institute'	-1.5077
公司 'company'	0.3809	部门 'department'	-4.2022
院 'institute'	2.8140	大会 'conference'	-0.6497
支队 'detachment'	2.1972	企业 'enterprise'	-4.6914

Table 5: The weights of some typical entity suffixes based on logistic transformation

Meanwhile, we can further remove some useless morphemes in terms of their weights: If the weights of entity prefixes or suffixes are greater than 0, they can be identified as important prefixes or suffixes for nested NER. Based on this rule, we extracted 1878 informative entity prefixes and 881 entity suffixes for nested NER from 2804 candidate

entity prefixes and 1402 candidate entity suffixes, respectively. After a manual checking, we finally obtain 950 entity prefixes and 316 entity suffixes.

5. Dual-layer chunking for Chinese nested NER

Since most Chinese nested NEs have a two-level structure, we can reformulate Chinese nested named entity recognition as a morpheme-based dual-layer chunking task, which identifying all simple NEs in the first layer, and then resolve nested NEs in the second layer. This section details the proposed dual-layer method.

5.1 Task formulation

As mentioned above, the major purpose of our current study is to explore more informative cues, including entity-external contextual features and entity-internal morphological features for Chinese nested NER. To this end, we take morphemes as the basic tokens in entity formation and reformulate Chinese NER as a dual-layer chunking task on a sequence of lexical chunks. Here, a lexical chunk string is a sequence of morphemes associated with their corresponding lexical chunk tags (Fu et al., 2008). It should be noted that each word and its POS tag in the input sentence should be converted into a lexical chunks (as shown in Table 7) before entity chunking.

To explore more entity-internal informative clues for nested named entity recognition, and to consider the efficiency in training at the same time, we distinguish simple NEs from nested NEs and use a hybrid label scheme, which employs two separate tagsets to label simple NEs and nested NEs, respectively.

For simple NEs, we use the traditional BIO tagset, which consists of three tags, namely *B*, *I* and *O*, to denote the respective position patterns of a token within NEs. Where *B* indicates that the token is at the beginning of a multi-token entity, *I* denotes that the token is at the middle or the end of multi-token named entities, and *O* denotes that the token is an independent NE by itself.

Tag	Definition
B	The current token is at the beginning of a multi-token entity chunk.
I	The current token is at the second of a multi-token entity chunk.
M	The current token is at the middle of a multi-token entity chunk.
E	The current token is at the end of a multi-token entity chunk.
O	The current token is an independent entity chunk by itself.

Table 6: The tagset for named entity chunks

As shown in Table 6, we utilize an extended version of the BIO tagset for nested NEs. For convenience, we refer it to as BIO-E tagset. We believe that the BIO-E tags could

better represent the relatively complicated structures of nested NEs and thus provide a convenient way for exploiting more potential clues, in particular the entity-internal structural features for nested named entity recognition.

Table 7 illustrates the word-level and entity-level chunk representations of the sentence "广州/ns 标志/nz 公司/n 与/p 北京/ns 地质部/nt" (*Guangzhou Peugeot Company and China Ministry of Geology*) under the hybrid scheme. Where "n", "ns", "nt", "nz", and "p" are the PKU POS tags (Yu et al., 2003) for common nouns, toponyms, organization nouns, other proper nouns, and prepositions, respectively.

Word	POS tag	Morpheme	Word-level chunk tags	Simple NE chunk tags	Nested NE chunk tags
广州	ns	广州	O-ns	O-LOC	B-ORG
标志	nz	标	B-nz	O-nz	I-ORG
		志	I-nz	O-nz	M-ORG
公司	n	公司	O-n	O-n	E-ORG
与	c	与	O-p	O-p	O
中国	ns	中国	O-ns	O-LOC	B-ORG
地质部	nt	地质	B-nt	B-ORG	M-ORG
		部	I-nt	I-ORG	E-ORG

Table 7: An example: Chunk representation of words and named entities

5.2 CRFs for Chinese NER

Since CRFs have proven to be one of the most effective techniques for sequence labeling tasks (Lafferty et al., 2001), here we employ CRFs to perform the dual-layer chunking for Chinese nested NER. In comparison with other methods, CRFs allow us to exploit a number of observation features as well as state sequence based features or other features to NER.

Let $X = (x_1, x_2, \cdots, x_T)$ be an input sequence of Chinese morpheme or word tokens, $Y = (y_1, y_2, \cdots, y_T)$ be a sequences of entity-level chunk tags as shown in Table 7. From a statistical point of view, the goal of NER is to find the most likely sequence of entity chunk tags $\hat{Y}$ for a given sequence of morpheme or word tokens X that maximizes the conditional probability $p(Y \mid X)$. CRFs modeling uses Markov random fields to decompose the conditional probability $p(Y \mid X)$ of a sequence of entity chunk tags as a product of probabilities below.

$$p(y \mid x) = \frac{1}{Z(x)} \exp(\sum_{i=1}^{T} \sum_{j} \lambda_j f_j(y, x, i)) \quad (3)$$

Where $f_j(y, x, i)$ is the j^{th} feature function at position *i*, associated with a weight λ_j, and $Z(x)$ is a moralization factor that guarantees that the summation of the probability of all sequences of entity-level chunk tags is one, which can be further calculated by

$$Z(x) = \sum_{y} \exp(\sum_{i=1}^{T} \sum_{j} \lambda_j f_j(y, x, i)) \quad (3)$$

5.3 Features

The exploration of informative features is essential to nested NER. Based on the above morpheme-based representation of Chinese NEs, in this section we continue to exploit a variety of entity-internal and entity-external features for Chinese NER.

Word-level features. Lexical information plays an important role in NER. In the present study we consider multiple lexical features, such as morpheme forms, part of speech, and the position of entity morphemes within entities. Furthermore, we also take into account the contextual lexical information outside a lexical unit For example, given an n-morpheme sentence. The observation is not limited to the current morphme *i*, lexical features within a window of five morphemes, namely (*i*-2, *i*-1, *i*, *i*+1, *i*+2), are explored for NER.

Entity-level features. In addition to word-level and entity-external contextual features, we also take into account entity-internal prefixes and suffixes for Chinese nested NER. Considering the particularity and complexity of some special useful entity suffixes, we also re-rank the extracted entity suffixes shown in Table 5 according to their weights, and thus divide them into two levels based on a given threshold. In the present study, the threshold for ranking entity suffixes is set to 3. Table 8 illustrates the division of some typical entity suffixes.

Level	Suffix	Weight
Level 1	议会 ‘parliament’	3.761273
	委 ‘committee’	3.919495
Level 2	公司 ‘company’	0.380874
	支队 ‘detachment’	2.197225

Table 8: Re-ranking entity suffixes in terms of their weights

6. Experimental Results and Discussions

To test the validity of our method, we have conducted several experiments on different datasets. This section reports the experimental results.

6.1 Experimental setup

In our experiments, we employ the entity-tagged corpus (Fu and Luke, 2005). To achieve a morpheme-based system for Chinese NER, we transform the original word-based corpus to a morpheme-based format using the maximum forward matching method. Furthermore, we divide this corpus into two parts: 90% is used as the training data, and the rest 10% is for test. Table 9 shows the distribution of different named entities in the experimental corpora.

NE-type	Training data		Test data	
	Total	Number of nested NEs	Total	Number of nested NEs
PER	27913	—	1796	—
LOC	23770	935	2261	196
ORG	15483	5897	1603	965

Table 9: Basic statistics of the experimental data

In addition, three metrics, namely *recall* (*R*), *precision* (*P*) and *F-score* (*F*), are computed to evaluate our system. Where, recall is defined as the number of correctly recognized entities divided by the total number of entities in the test data, while precision can be interpreted as the number of correctly recognized entities is divided by the total number of entities yielded automatically by the system. *F-score* is the balanced value of recall and precision, namely $F = 2 * P * R / (P + R)$.

6.2 Experimental results

Our first experiment is designed to test the effects of different chunking tokens, namely morpheme-based chunking tokens and word-based chunking tokens on NER. This experiment is carried out by applying the CRF-based chunker to a morpheme-based dataset (referred to as CRF_M) and a word-based dataset (referred to as CRF_W), respectively, and comparing the relevant outputs. The experimental results are presented in Table 10.

As illustrated in Table 10, the morpheme-based system overall outperforms the word-based one. The reason may be due to the fact that it is more straightforward to explore morphological features for NER under a morpheme-based framework than under a word-based framework.

System	Nested NEs			All NEs		
	P	R	F	P	R	F
CRF_W	0.905	0.558	0.680	0.943	0.607	0.731
CRF_M	0.819	0.613	0.701	0.919	0.633	0.750

Table 10: Results for NER with different basic tokens

Our second experiment is aiming at investigating the effects of lexical features and multi-level entity morphological features on nested NER. In this experiment, we take CRF_M in the first experiment as a baseline, and additionally introduce the lexical features (referred as CRF_ML) and the multi-level prefix and suffix morpheme features (referred to as CRF_MM), and then evaluate the related outputs, respectively. The experimental results are presented in Table 11 and Table 12.

System	NE-type	Nested NEs			All NEs		
		P	R	F	P	R	F
CRF_M	ORG	0.850	0.651	0.737	0.890	0.672	0.766
	LOC	0.646	0.429	0.515	0.917	0.655	0.764
CRF_ML	ORG	0.862	0.745	0.799	0.906	0.752	0.822
	LOC	0.811	0.595	0.686	0.921	0.931	0.927
CRF_MM	ORG	0.849	0.843	0.846	0.891	0.843	0.867
	LOC	0.865	0.684	0.764	0.947	0.945	0.946

Table 11: Results for ORG and LOC recognition with different morphological features

As can be observed from Table 11, the performance for nested NER increases with more features introduced. Take nested organization name recognition for example. the F-score is 73.7% for the baseline system. The number increase to 79.9% after using lexical features, and further to 84.4% after the introduction of multi-level entity morphological information. Similar trends can be observed with regards to nested location name recognition. This shows in a sense that both lexical features and entity morphological information are equally important cues for nested NER.

Methods	Nested NEs			All NEs		
	P	R	F	P	R	F
CRF_M	0.819	0.613	0.701	0.919	0.633	0.750
CRF_ML	0.855	0.720	0.781	0.941	0.898	0.919
CRF_MM	0.851	0.817	0.833	0.944	0.929	0.938

Table 12: Overall NER performance with different morphological features

Table 12 presents the overall NER performance for the second experiments. As can be seen in this table, the overall F-score for all types of NEs can be improved from 75.0% to 91.9 after using lexical information, and further to 93.8%, illustrating not only the important roles of lexical information and entity-internal morphological features in nested NER, but also the significance of nested named entities in NER

To further demonstrate the effectiveness of our approach, we also conducted an open test on the IEER-99 and MET2 test data (Chinchor, 1999). Table 13 shows the distribution of named entities in the two corpora, and Table 14 lists the relevant experimental results.

NE-type	IEER-99		MET2	
	Total	Number of nested NEs	Total	Number of nested NEs
PER	504	—	170	—
LOC	962	60	585	84
ORG	483	236	318	224

Table 13: Basic statistics of NEs in the IEER-99 and the MET2 test datasets

Test data	Nested NEs			All NEs		
	P	R	F	P	R	F
IEER-99	0.678	0.753	0.714	0.923	0.905	0.914
MET2	0.759	0.789	0.774	0.910	0.908	0.909

Table 14: Experimental results on the IEER-99 and the MET2 test datasets

As can be seen from Table 14, our system performs better than our system can achieve an F-score of 0.714 and 0.789 for nested NER over IEER-99 and MET2 datasets, respectively. The respective overall F-score for all NEs are 0.914 and 0.909.

Although the proposed approach is effective for most nested NEs, it fails to yield correct results for some complex nested NEs like 中国人民银行江西金溪县支行 '*Jinxi branch of China People's Bank in Jiangxi*', or for some NEs that contain multiple prefixes and suffixes, such as 山东省石油公司加油站 '*Gas Station of Oil Company in Shandong Province*'. The reason may be because that our current approach still relies on sequence labeling and thus does not works for some nested NEs with complicated structures.

7. Conclusions

In this paper, we have presented a morpheme-based dual-layer chunking method to Chinese nested NER. In comparison with previous studies, our approach offers a straightforward framework for exploring more features, including entity-internal features and contextual features for Chinese nested NER. Our experimental results on different test set show that

both lexical information and entity-internal morphological features are equally important for Chinese nested NER.

The results of the present study suggest two possibilities for future research. First, we have shown the benefit of using entity-internal morphological features in Chinese nested NER. However, our current study is still relatively primitive. Therefore, a more systematic research would be necessary to explore entity-level morphological cues for Chinese nested NER. Second, the present system still relies on the framework of sequence labeling, and does not works for some nested NEs with complicated structures. Thus, future research might usefully extend the present dual-level sequence labeling framework to a structural model that can handle complex structural information or entity formation patterns.

Acknowledgements

This study was supported by National Natural Science Foundation of China under Grant No.60973081 and No.61170148, Harbin Innovative Foundation for Returnees under Grant No.2009RFLXG007, and the Returned Scholar Foundation of Educational Department of Heilongjiang Province under Grant No.1154hz26, respectively.

References

Alex, Beatrice, Barry Haddow, and Claire Grover, 2007, Recognizing nested named entities in biomedical text, Proceedings of the Biological, Translational, and Clinical Language Processing, pp.65-72.

Ashton, Winifred D, 1972, The logit transformation: With special reference to its uses in bioassay, Hafner Publishing Company.

Ekbal, Asif, and Sivaji Bandyopadhyay, 2008, Bengali named entity recognition using support vector machine, Proceedings of ACL-IJCNLP'08 Workshop on NER for South and South East Asian Languages, pp.51-58.

Finkel, Jenny Rose, and Christopher D. Manning, 2009, Nested named entity recognition, Proceedings of EMNLP'09, pp.141-150.

Fu, Guohong, 2009, Improving Chinese named entity recognition with lexical information, Proceedings of ICMLC'09, pp.12-15.

Fu, Guohong, and Kang-Kwong Luke, 2005, Chinese named entity recognition using lexicalized HMMs, ACM SIGKDD Explorations, vol.7, no.1, pp.19-25.

Fu, Guohong, Chunyu Kit, and Jonathan J. Webster, 2008, Chinese word segmentation as morpheme-based lexical chunking. Information Sciences, vol.178, no. 9, pp.2282-2296.

Lafferty, John, Andrew McCallum, and Fernando Pereira, 2001, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, Proceedings of ICML'01, pp.282-289.

Li, Dingcheng, Karin Kipper-Schuler, and Guergana Savova, 2008, Conditional random fields and support vector machines for disorder named entity recognition in clinical texts,

Proceedings of ACL'08, pp.94-95.

Liu, Jingchen, Minlie Huang, and Xiaoyan Zhu, 2010, Recognizing biomedical named entities using skip-chain conditional random fields, Proceedings of ACL'10, pp.10-18.

Saha, Sujan Kumar, Sudeshna Sarkar, and Pabitra Mitra, 2008, A hybrid feature set based maximum entropy hindi named entity recognition, Proceedings of ACL-IJCNLP'08, pp.343-349.

She, Jun, and Xue-Qing Zhang, 2010, Musical named entity recognition method, Journal of Computer Applications, vol. 31, no.11, pp.2928-2931.

Tsai, TzongHan, Chia-Wei Wu, and Wen-Lian Hsu, 2005, Using maximum entropy to extract biomedical named entities without dictionaries, Proceedings of ACL-IJCNLP'05, pp.268-273.

Wang, Yefeng, 2009, Annotating and recognizing named entities in clinical notes, Proceedings of the ACL-IJCNLP'09 Student Research Workshop, pp.18-26.

Yu, Shiwen, Huimin Duan, Xuefeng Zhu, Bin Swen, and Baobao Chang, 2003, Specification for corpus processing at Peking University: Word segmentation, POS tagging and phonetic notation, Journal of Chinese Language and Computing, vol.13, no.2, pp.121-158.

Chinchor, Nancy, 1999, Overview of MUC-7/MET-2, Proceedings of MUC-7.

Zhang, Yuejie, Zhiting Xu, and Xiangyang Xue, 2008, Fusion of multiple features for Chinese named entity recognition based on maximum entropy model, Computer Research and Development, vol.45, no.6, pp.1004-1010.