Chinese Hedge Scope Detection Based on Structure and Semantic Information

Huiwei Zhou, Junli Xu, Yunlong Yang, Huijie Deng, Long Chen, and Degen Huang

School of Computer Science and Technology,
Dalian University of Technology, Dalian 116024, Liaoning, China
{zhouhuiwei,huangdg}@dlut.edu.cn,{xjlhello,Sdyyl_1949,denghuijie,chenlong.415}@mail.dlut.edu.cn

Abstract. Hedge detection aims to distinguish factual and uncertain information, which is important in information extraction. The task of hedge detection contains two subtasks: identifying hedge cues and detecting their linguistic scopes. Hedge scope detection is dependent on syntactic and semantic information. Previous researches usually use lexical and syntactic information and ignore deep semantic information. This paper proposes a novel syntactic and semantic information exploitation method for scope detection. Composite kernel model is employed to capture lexical and syntactic information. Long short-term memory (LSTM) model is adopted to explore semantic information. Furthermore, we exploit a hybrid system to integrate composite kernel and LSTM model into a unified framework. Experiments on the Chinese Biomedical Hedge Information (CBHI) corpus show that composite kernel model could effectively capture lexical and syntactic information, LSTM model could capture deep semantic information and their combination could further improve the performance of hedge scope detection.

Keywords: hedge scope detection, structure information, semantic information.

1 Introduction

Hedges indicate uncertain or unreliable information, which are usually used in science texts. In English, 17.69% of the sentences in the abstract section and 22.29% of the sentences in the full paper section contain uncertain information on BioScope corpus [1]. In Chinese, 29.30% of the sentences contain speculative fragments on CBHI corpus [2]. In order to distinguish facts from uncertain information, hedge detection is becoming an important task for information extraction. The CoNLL-2010 Shared Task [3] was dedicated to detecting uncertainty cues and their linguistic scopes on English corpus. Chinese hedge information detection has also attracted considerable attention [4]. This paper focuses on Chinese hedge scope detection on the CBHI corpus. A hedged sentence taken from the CBHI corpus is shown as follows:

Sentence 1: *上述实验数据提示<scope>PCAF<ccue>可能</ccue>是一种HCC的抑癌因子</scope>，具有成为预测HCC术后愈后情况的生物标志物。*

(The above experimental data suggest that <scope>PCAF<ccue>may</ccue> be a tumor suppressor factors of HCC</scope>, and has become a predict postoperative HCC prognosis biomarkers.)

In sentence 1, the word "*可能 (may)*" is hedge cue and its scope is the statement that "*PCAF可能是一种HCC的抑癌因子 (PCAF may be a tumor suppressor factors of HCC)*".

Researches on hedge cue identification have been developed rapidly [5,6]. However, hedge scope detection remains a challenge, since hedge scope detection is dependent on syntactic and semantic information. This paper focuses on hedge scope detection from structure and semantic perspective.

Existing studies on hedge scope detection contain feature-based and tree kernel-based methods. Feature-based methods define a set of discrete features with "one-hot" representations based on lexical and flat syntactic information. Tree kernel-based methods could capture structured syntactic information by counting the number of common sub-trees [7]. However, both feature-based and tree kernel-based methods could not capture deep semantic information between cues and their linguistic scopes.

This problem motivates us to develop neural network models which could capture deep semantic information for scope detection. We propose a novel syntactic and semantic information exploitation method, which consists of a composite kernel and LSTM model. Composite kernel model is designed to capture lexical and structured syntactic information. LSTM model is adopted to explore deep semantic information. Furthermore, to fully utilize the nice properties of lexical, syntactic and semantic information, we explore a hybrid system to integrate composite kernel and LSTM model into a unified framework.

2 Related work

In this section, we review the literature related to this paper from two aspects: hedge scope detection and neural network approaches for Nature Language Processing (NLP) tasks.

2.1 Hedge scope detection

Existing researches for hedge scope detection mainly contain: rule-based and machine learning-based methods. Rule-based methods [8,9] compile heuristic rules by exploiting lexico-syntactic patterns for scope detection. Rule-based methods are simple and effective, but the extracted rules are hard to be developed to a new resource.

Machine learning-based methods formulate scope detection task as a classification issue, which classifies each token/sub-structure in a sentence as being the first element of the scope (F-scope), the last (L-scope), or neither (None). Machine learning-based methods mainly include feature-based and tree kernel-based methods. Feature-based methods design a set of discrete features with "one-hot" representations based on lexical and flat syntactic information. Morante and Daelemans [10] explore lexical features to predict F-scope, L-scope and None. Morante et al. [11] and Li et al. [12]

exploit flat syntactic features for scope detection. The above researches take tokens as classification units, which inevitably generate plenty of instances. To decrease instances, Zhu et al. [13] and Zou et al. [14] construct feature-based systems by taking phrase and dependency sub-structures as classification units, respectively. Tree kernel-based methods could capture structured syntactic information by counting the number of common sub-trees. Zhang et al. [15] use tree kernel-based methods to model structured syntactic information for relation extraction. Zhou et al. [16] and Zou et al. [17] investigate phrase sub-structures and dependency sub-structures respectively to capture structured syntactic information for scope detection.

Feature-based and tree kernel-based methods could effectively capture lexical and syntactic information. However, the extracted features with feature-based and tree kernel-based methods are discrete and could not capture deep semantic information.

2.2 Neural Network for NLP tasks

Neural networks could learn deep semantic representations without feature engineering. Especially, LSTM model [18] is superior in semantic representations of surface sequences. Xu et al. [19] use LSTM to pick up semantic information along the shortest dependency path between two entities for relation extraction. Zhou et al. [20] explore a series of semantic representations with LSTM model and further integrate diverse information for chemical-disease relation extraction.

Motivated by the success of LSTM model and Zhou et al. [20], we propose a hybrid system which consists of composite kernel and LSTM model to capture lexical, syntactic and semantic information for scope detection.

3 Methods

The corpus is preprocessed with Stanford Parser[1] to get lexical and syntactic information. To decrease candidate instances, we take phrase sub-structures as classification units adopting the way of Zhu et al. [13]. For the left (right) candidate phrase of a given cue, the leftmost (rightmost) word is F-scope (L-scope).

The hybrid system architecture consists of training and test phases as shown in Fig. 1. In training phase, lexical and syntactic features are captured by composite kernel model, and semantic representations are learned by LSTM model. In test phase, two models are applied to detect hedge scope. The predicted results of the two models are combined to optimize system performance finally.

[1] Available at http://nlp.stanford.edu/software/lex-parser.shtml

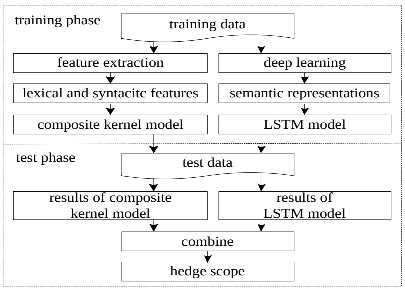

Fig. 1. Hybrid system architecture.

3.1 Composite kernel for hedge scope detection

The polynomial kernel.

The feature-based model is learned from lexical features with polynomial kernel $K_{poly}(x_i, x_j) = (x_i \cdot x_j + 1)^d$, where d is the dimension of polynomial kernel. We select widely-used features for scope detection as shown below. These features reflect lexical information of hedge and its candidate.

- *WordContext*: words of cue and its candidate in the window [-2, 2].
- *CandidateType*: the constituents of candidate phrase, such as NP, VP.
- *HedgePoS*: the part-of-speech (PoS) of hedge.

The convolution tree kernel.

The convolution tree kernel could effectively capture structured syntactic information. This paper focuses on the information combination, so we only adopt the extending phrase path tree (EPPT) [15] to explore the structured syntactic information for scope detection. EPPT includes the path from the hedge to its candidate, and the nearest neighbor tokens of both the hedge and its candidate in the phrase tree. The path from hedge to its candidate represents the most direct phrase syntactic information about the hedge and its candidate. Adding the neighbor structures could provide rich context syntactic information. For the phrase syntactic tree of sentence 1 as shown in Fig. 2(a), the EPPT about the hedge "*可能* (*may*)" and its L-scope candidate "*HCC*" is shown in Fig. 2(b).

The composite kernel.

To integrate the lexical and syntactic features, the composite kernel is defined by combining the polynomial kernel and the convolution tree kernel:

$$K_{com} = \gamma K_{tree} + (1-\gamma) K_{poly} \quad (1)$$

where $\gamma (0 < \gamma < 1)$ is the composite factor. The polynomial kernel K_{poly} and the convolution tree kernel K_{tree} are combined by the composite kernel K_{com}.

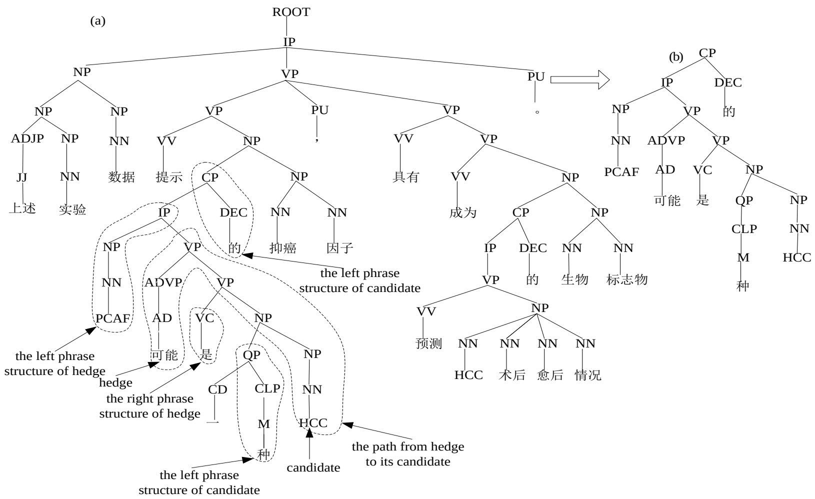

Fig. 2. Syntactic features extraction. (a) The phrase syntactic tree of sentence 1; (b) Extending phrase path tree (EPPT).

3.2 Long Short-Term Memory (LSTM) for hedge scope detection

LSTM model is a kind of recurrent neural network (RNN), which introduces a gating mechanism to avoid gradient vanishing and exploding. LSTM cell comprises four components: an input gate $i_t = \sigma(W_i \cdot [h_{t-1}; x_t] + b_i)$, a forget gate $f_t = \sigma(W_f \cdot [h_{t-1}; x_t] + b_f)$, an output gate $o_t = \sigma(W_o \cdot [h_{t-1}; x_t] + b_o)$, and a memory cell $c_t = i_t \odot \tanh(W_r \cdot [h_{t-1}; x_t] + b_r) + f_t \odot c_{t-1}$. These gates adaptively remember input vector, forget previous history and generate output vector, where $\odot$ denotes component-wise multiplication, σ represents the sigmoid function, W_i, b_i, W_f, b_f, and W_o, b_o are parameters of input, forget and output gates for the input x_t and the hidden state vector h_{t-1} respectively. LSTM processes the word sequence by recursively computing its internal hidden state $h_t = o_t \odot \tanh(c_t)$ at each time step. Our intuition is that context semantic information of hedge and its candidate is important for scope detection. We develop four LSTM models to explore deep semantic information related to hedge scope as following.

CanHedSeq-LSTM.
The context words of hedge and its candidate in the window [-2, 2] are jointed as a CanHedSeq sequence feeding to LSTM for recursively capturing context semantic representations of hedge scope. The dimension of word representations $x_w \in R^{d_1}$ is d_1. An illustration of the CanHedSeq-LSTM model is shown in Fig. 3.

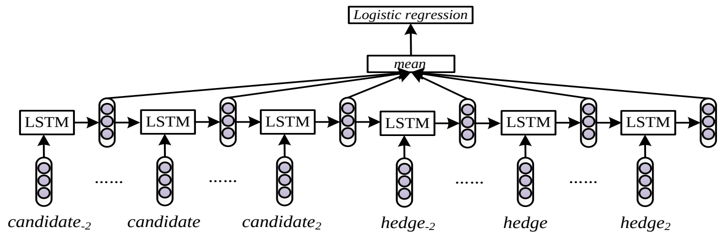

Fig. 3. Hedge scope detection based on CanHedSeq-LSTM

Bi_CanHed-LSTM.

We use two LSTM models: one LSTM model captures semantic information for context words of candidate in the window [-2, 2], and another LSTM model computes semantic information for context words of hedge in the window [-2, 2]. Afterwards, the last hidden vectors of two LSTM models are concatenated and fed to a logistic regression layer to detect scope. An illustration of the Bi_CanHed-LSTM model is shown in Fig. 4.

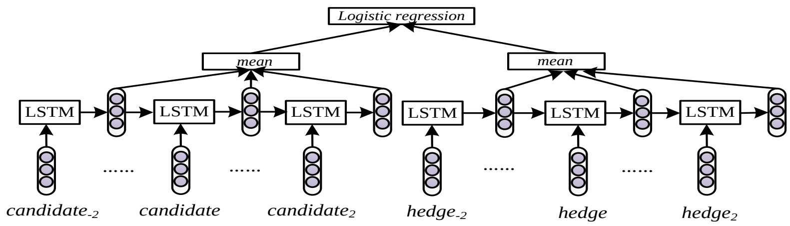

Fig. 4. Hedge scope detection based on Bi_CanHed-LSTM

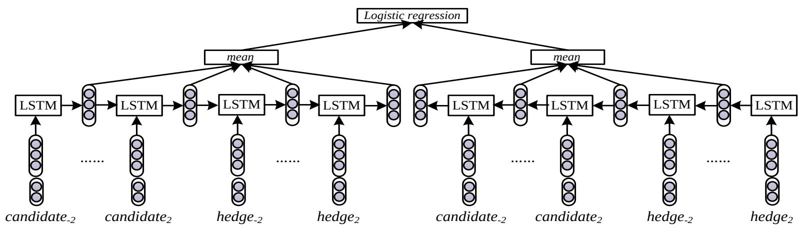

Fig. 5. Hedge scope detection based on Bi_CanHedSeq-LSTM

Bi_CanHedSeq-LSTM.

Both the history information and future information in a sequence are important for scope detection. In order to obtain the history and future information of CanHedSeq sequence, we use a forward LSTM and a backward LSTM to model the forward and backward CanHedSeq sequence, respectively. Afterwards, the last hidden vectors of two LSTM models are concatenated and fed to a logistic regression layer to detect scope. An illustration of the Bi_CanHedSeq-LSTM model is shown in Fig. 5.

Bi_CanHedSeq_Con-LSTM.
To further represent the information of *CandidateType* and *HedgePos*, we construct Bi_CanHedSeq_Con-LSTM model based on the Bi_CanHedSeq-LSTM. In the Bi_CanHedSeq_Con-LSTM model, the representation of *CandidateType* $x_c \in R^{d_2}$ is concatenated to the representations of the context words $x_w \in R^{d_1}$ of candidate to form a vector representation $x_w, x_c \in R^{d_1+d_2}$, and the representation of *HedgePos* $x_h \in R^{d_2}$ is concatenated to the representations of the context words $x_w \in R^{d_1}$ of hedge to form a vector representation $x_w, x_h \in R^{d_1+d_2}$. An illustration of the model is shown in Fig. 6.

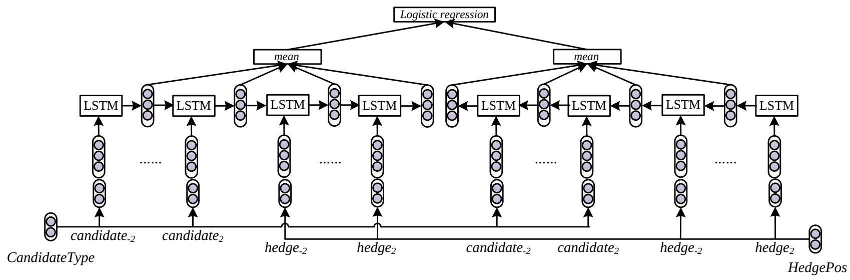

Fig. 6. Hedge scope detection based on Bi_CanHedSeq_Con-LSTM

3.3 Hybrid system for hedge scope detection

Both the composite kernel model and the LSTM model have their own advantages and could capture different information for scope detection. We propose a hybrid system integrating the composite kernel model $K(t_i)$ weighted by $\alpha \in [0,1]$ and the LSTM model $N(s_i)$ weighted by $1-\alpha \in [0,1]$.

The predicted results of the composite kernel model are the distances between the instances and the separating hyperplane, while those of the LSTM model are the probabilities of the instances. We adopt a uniform framework with sigmoid function σ to transform the distance into a probability as shown in equation (2).

$$P(H_i) = \alpha \cdot \sigma(K(t_i)) + (1-\alpha) \cdot N(s_i) \tag{2}$$

where t_i represents the lexical and syntactic features and s_i represents semantic representations of the hedge scope H_i in test data. The parameters $\alpha \in [0,1]$ could be controlled to investigate the impacts of composite kernel model vs. LSTM model. The sigmoid function σ is monotonic, and the point $P(y=1|f)=0.5$ occurs at the separating hyperplane $f=0$. Therefore, the boundary probability is set to 0.5 to separate boundaries from non-boundaries.

3.4 Postprocessing

To guarantee that all scopes are continuous sequences of tokens, we apply the following rules to hedge scope detection system.

(1) If one token is predicted as F-scope and one token as L-scope, the sequence will start at the token predicted as F-scope, and end at the token predicted as L-scope.
(2) If one token is predicted as F-scope, and none/more than one token is predicted as L-scope, the sequence will start at the token predicted as F-scope, and end at the token with the maximum L-scope predicted result.
(3) If one token is predicted as L-scope, and none/more than one token is predicted as F-scope, the sequence will start at the token with the maximum F-scope predicted result, and end at the token with the maximum L-scope predicted result.

4 Experiments and Discussion

Experiments are conducted on the CBHI corpus. The training and test data contain 7510, 1875 sentences respectively. We detect the linguistic scopes with golden standard cues. Stanford Word Segmenter toolkit[2] is employed to segment words and get PoS tag. SVM-LIGHT-TK toolkit[3] is used to construct the composite kernel model. LSTM model is developed based on Theano system[4] [21]. The evaluation of scope detection is reported by F1-score on tag-level and sentence-level. The tag-level takes the token as the evaluation unit, and evaluates the performance of the F-scope and L-scope classifiers respectively. The sentence-level corresponds to the exact match of scope boundaries for each cue.

4.1 Effects of composite kernel for hedge scope detection

The detailed performances of the lexical features with polynomial kernel under the condition $d = 2$ are summarized in Table 1. From the results, we can see that *Word-Context* features achieve poor results. With other features added one by one, the performance improves continuously and reaches 63.95% F1-score. All of the lexical features are effective for scope detection. Lexical features with polynomial kernel could obtain acceptable performance. However, the feature engineering is labor intensive and the extracted features with "one-hot" representations are discrete and only capturing shallow information for hedge scope detection.

We use composite kernel model to capture lexical features and structured syntactic information. Fig. 7 shows the performance of composite kernel with different composite factor γ. We vary γ from 0 to 1with an interval of 0.1. From Fig. 7, we can see that:

---

2 Available at http://nlp.stanford.edu/software/segmenter.shtml
3 Available at http://disi.unitn.it/moschitti/Tree-Kernel.htm
4 Available at http://deeplearning.net/software/theano/

(1) The sole tree kernel (γ=1) obtains 49.92% F1-score, which is worse than the sole polynomial kernel (γ=0). The composite kernel combining lexical and syntactic features with any composite factor γ could achieve higher performance than either one of them on both tag-level and sentence-level F1-score. The best performance of F-scope (L-scope) obtains 73.03% (85.07%) F1-score on tag-level. In sentence-level, we obtain 68.91% F1-score under the condition γ=0.3 . It indicates that tree kernel could capture useful structure information which hardly can be designed by feature engineering. The composite kernel could effectively realize the complementary of lexical and structured syntactic features.

(2) The performance of L-scope classifiers is usually better than that of F-scope classifiers. The main reason is that the distance of F-scope to its cue is longer than that of L-scope in a sentence on the CBHI corpus. The longer the distance from the scope boundary to its cue is, the harder the scope detection is.

Table 1. Performance of the Lexical features with polynomial kernel

Lexical	Bounary	P(%)	R(%)	F1-score(%)	Sentence-level F1-score(%)
WordContext	F-scope	82.84	48.91	61.67	54.51
	L-scope	67.32	64.37	65.81	
+CandidateType	F-scope	75.80	68.00	71.69	61.81
	L-scope	70.94	73.71	72.30	
+HedgePos	F-scope	75.94	67.52	71.48	63.95
	L-scope	75.16	85.87	80.16	

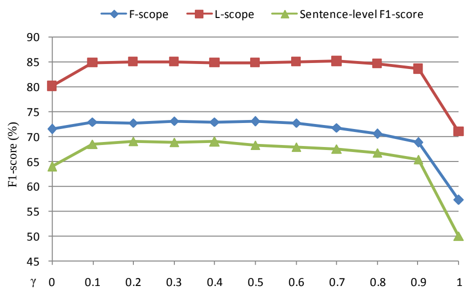

Fig. 7. The performance of composite kernel

4.2 Effects of LSTM for hedge scope detection

In our experiments, we use Word2Vec[5] toolkit to pre-train word representations on the SogouCS corpus[6]. The dimension d_1 of word representation is 100. The representations of *HedgePos* and *CandidateType* are initialized randomly with dimension 10. Table 2 shows the performance with four LSTM models.

[5] Available at https://code.google.com/p/word2vec/

[6] Available at http://www.datatang.com/data/list/s04-r020-t01-c03-la01-p3

Table 2. Performance of the semantic information with LSTM

LSTM	Bounary	P(%)	R(%)	F1-score(%)	Sentence-level F1-score(%)
CanHedSeq-LSTM	F-scope	65.26	60.00	62.51	55.73
	L-scope	63.01	74.24	68.17	
Bi_CanHedSeq-LSTM	F-scope	68.94	58.83	63.48	53.97
	L-scope	64.72	71.25	67.83	
Bi_CanHedSeq-LSTM	F-scope	60.08	65.81	62.81	55.15
	L-scope	71.12	67.89	69.47	
Bi_CanHedSeq_Con-LSTM	F-scope	65.24	68.96	67.05	59.09
	L-scope	74.93	80.64	77.68	

(1) Performance of scope detection obtains acceptable result under any LSTM models. This indicates that the context of hedge and its candidate could represent the hedge scope, and the four LSTM models could effectively capture semantic information of hedge scope.
(2) CanHedSeq-LSTM achieves 55.73% F1-score, which is 1.22% higher than the *WordContext* features with polynomial kernel. This indicates that CanHedSeq-LSTM model could capture the deep semantic information, while *WordContext* features only represent shallow semantic information.
(3) Bi_CanHed-LSTM obtains worse performance than CanHedSeq-LSTM. The main reason is that Bi_CanHed-LSTM with two LSTM models capturing the candidate and hedge contexts respectively, which may ignore the semantic relation between candidate and hedge contexts.
(4) Bi_CanHedSeq-LSTM achieves better performance, which obtains 55.15% F1-score. This indicates that Bi_CanHedSeq-LSTM can simultaneously capture the history and future semantic information, and both the history and future information are important for hedge scope detection.
(5) Bi_CanHedSeq_Con-LSTM obtains best F1-score 59.09%, which is 3.94% higher than Bi_CanHedSeq-LSTM. This indicates that the concatenation of *HedgePos* and *CandidateType* are effective for scope detection. However, the increasing amount is smaller than that *HedgePos* and *CandidateType* as lexical features with polynomial kernel. This may due to that the concatenations of *CandidateType* (*HedgePos*) and context words may bring noises at some time steps in LSTM.

4.3 Effects of weighting parameters

We investigate the impact of the parameters α that control the weighting of LSTM model vs. the composite kernel model. The composite kernel model under the condition γ=0.3 (68.91% F1-score) and each LSTM model are used in the hybrid system.

From Fig. 8, we can see that the trends of the four curves are similar. All starts from the initial F1-score of LSTM model, and then increases to the individual highest F1-score, finally falls below the initial F1-score (68.91%) of composite kernel model. The composite kernel is combined with CanHedSeq-LSTM (55.73% F1-score), Bi_CanHed-LSTM (53.97% F1-score), Bi_CanHedSeq-LSTM (55.15% F1-score) and Bi_CanHedSeq_Con-LSTM (59.09% F1-score) obtaining 69.92%, 69.76%, 69.49% and 69.33% F1-score, respectively. These indicate that both the composite kernel model and the four LSTM models have their own advantages and could capture different information for scope detection. Their combination could further improve the performance of scope detection.

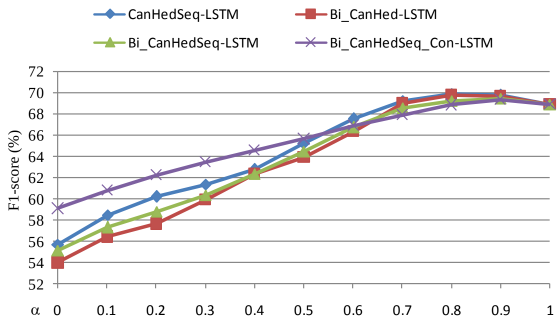

Fig. 8. The performance with different weightings

5 Conclusions and Future Work

Lexical features, syntactic structure features, and semantic representations are all particularly effective for hedge scope detection. We propose a hybrid system to integrate these information for Chinese hedge scope detection, which achieves 69.92% F1-score on the CBHI corpus. The hybrid system consists of the composite kernel model and the LSTM model. The composite kernel model could effectively capture lexical and syntactic information. The LSTM model could explore deep semantic information of hedge scope. In addition, four LSTM models are developed to explore deep semantic information related to hedge scope.

For the future work, we will explore other deep neural network models to capture more effective semantic information. Besides, we will explore other hybrid methods which integrate diverse information to further improve the performance of scope detection.

Acknowledgements. This research is supported by Natural Science Foundation of China (No. 61272375).

6 Reference

1. Szarvas, G., Vincze, V., Farkas, R., Csirik, J.: The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In: Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing, pp. 38-45. ACL, USA (2008).
2. Zhou, H.W., Yang, H., Zhang, J., Kang, S.Y., Huang, D.G.: The research and construction of Chinese hedge corpus. Journal of Chinese Information Processing. 83-89 (2015).

3. Farkas, R., Vincze, V., Móra, G., Csirik, J., Szarvas, G.: The CoNLL-2010 shared task: learning to detect hedges and their scope in natural language text. In: Proceedings of CoNLL, pp. 1-12. ACL, Sweden (2010).
4. Zou, B.W., Zhou, G.D., Zhu, Q.M.: Negation and Speculation extraction: An overview. Journal of Chinese Information Processing. 04, 16-24 (2015).
5. Wei, Z.Y., Chen, J.W., Gao, W., Li, B.Y., Zhou, L.J., He, Y.L., Wong, K.F.: An Empirical Study on Uncertainty Identification in Social Media Context. In: Proceedings of ACL, pp. 58-62. ACL, Bulgaria (2013).
6. Su, Q., Lou, H.Q., Liu, P.Y.: Hedge Detection with Latent Features. Chinese Lexical Semantics. Springer Berlin Heidelberg. 436-441 (2013).
7. Moschitti, A.: A study on convolution kernels for shallow semantic parsing. In: Proceedings of ACL, pp. 335. ACL, Spain (2004).
8. Özgür, A., Radev, D.R.: Detecting speculations and their scopes in scientific text. In: Proceedings of EMNLP, pp. 1398-1407. ACL, Singapore (2009).
9. Øvrelid, L., Velldal, E., Oepen, S.: Syntactic scope resolution in uncertainty analysis. In: Proceedings of CL, pp. 1379-1387. ACL, Beijing (2010).
10. Morante, R., Daelemans, W.: Learning the scope of hedge cues in biomedical texts. In: Proceedings of BioNLP, pp. 28-36. ACL, Colorado (2009).
11. Morante, R., Asch, V.V., Daelemans, W.: Memory-based resolution of in-sentence scopes of hedge cues. In: Proceedings of CoNLL, pp. 40-47. ACL, Sweden (2010).
12. Li, X.X., Shen, J.P., Gao, X., Wang, X.: Exploiting rich features for detecting hedges and their scope. In: Proceedings of CoNLL, pp. 78-83. ACL, Sweden (2010).
13. Zhu, Q.M., Li, J.H., Wang, H.L., Zhou, G.D.: A unified framework for scope learning via simplified shallow semantic parsing. In: Proceedings of EMNLP, pp. 714-724. ACL, USA (2010).
14. Zou, B.W., Zhu, Q.M., Zhou, G.D.: Negation and Speculation Identification in Chinese Language. In: Proceedings of ACL-IJCNLP, pp. 656-665. ACL, Beijing (2015).
15. Zhang, M., Zhang, J., Su, J.: Exploring syntactic features for relation extraction using a convolution tree kernel. In: Proceedings of ACL, pp. 288-295. ACL, New York (2006).
16. Zhou, H.W., Huang, D.G., Li, X.Y., Yang, Y.S.: Combining structured and flat features by a composite kernel to detect hedges scope in biological texts. Chinese Journal of Electronics, 20(3), 476-482 (2011).
17. Zou, B.W., Zhou, G.D., Zhu, Q.M.: Tree Kernel-based Negation and Speculation Scope Detection with Structured Syntactic Parse Features. In: Proceedings of EMNLP, pp. 968-976. ACL, USA (2013).
18. Tai, K.S., Socher, R., Manning, C.D.: Improved semantic representations from tree-structured long short-term memory networks. In: Proceedings of ACL-IJCNLP, pp. 1556–1566. ACL, Beijing (2015).
19. Xu, Y., Mou, L.L., Li, G., Chen, Y.C., Peng, H., Jin, Z.: Classifying relations via long short term memory networks along shortest dependency paths. arXiv preprint arXiv: 1508.03720 (2015).
20. Zhou, H.W., Deng, H.J., Chen, L., Yang, Y.L., Jia, C.: Exploiting syntactic and semantics information for chemical–disease relation extraction. Database, baw048 (2016).
21. Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Bengio, Y.: Theano: a CPU and GPU math expression compiler. In: Proceedings of SciPy, pp. 1-7. SciPy, Austin (2010).