

# An Approach for the Chinese Question-answer System based on Document

Benyou Wang<sup>1</sup>, Jiabing Niu<sup>1</sup>, Liqun Ma<sup>1</sup>, Lipeng Zhang<sup>1</sup>, Yuhua Zhang<sup>1</sup>, and Peng Zhang<sup>1</sup>

Tianjin Key Laboratory of Cognitive Computing and Application, School of Computer Science and Technology, Tianjin University, Tianjin, China

**Abstract.** Question Answering system has gradually become a new trend within the field of information retrieval and NLP. It outperforms the conventional search engines, for the system is able to answer users questions automatically and accurately. Question Answering system based on English corpus has developed rapidly, whereas the Chinese corpus based Question Answering system still has some problems remains to be solved. Thus, developing a new Question Answering model, which is characterized by dealing with features of Chinese corpus is extremely essential. Different to the current deep learning model, our model uses the semantic and syntactic information in Chinese corpus and bases on the linearity of Chinese texts. Finally, our model turns out to perform better than other methods through experiments...

**Keywords:** Question Answer, graph theory, Hamilton cycles

## 1 Introduction

## 2

just test a cite [?]

## 3 Methods

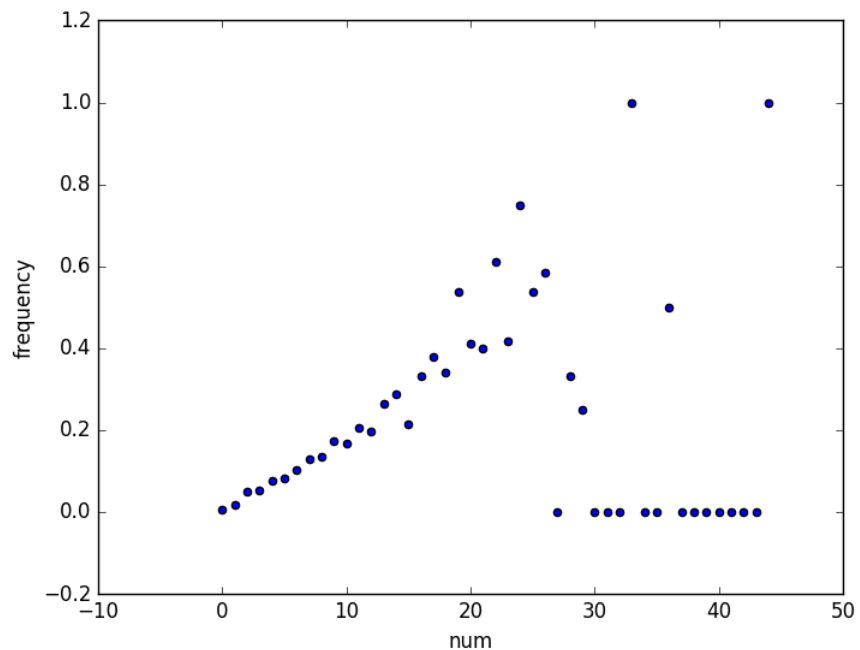
### 3.1 Data Exploration

### 3.2 Data Exploration

**Word Overlap and Character Overlap** It is considered that the more keywords in questions are matched with those in the answer sentences, the more likely that the answer is the correct one. We have found that there are some relationships between the overlapped characters or words and the occurrence frequency of the correct answer after analysing the given corpus.

Fig. 2

### Position Message in Overlapped Words



**Fig. 1.** This is the caption of the figure displaying a white eagle and a white horse on a snow field

## Word Overlap and Character Overlap

### 3.3 Data Preprocessing

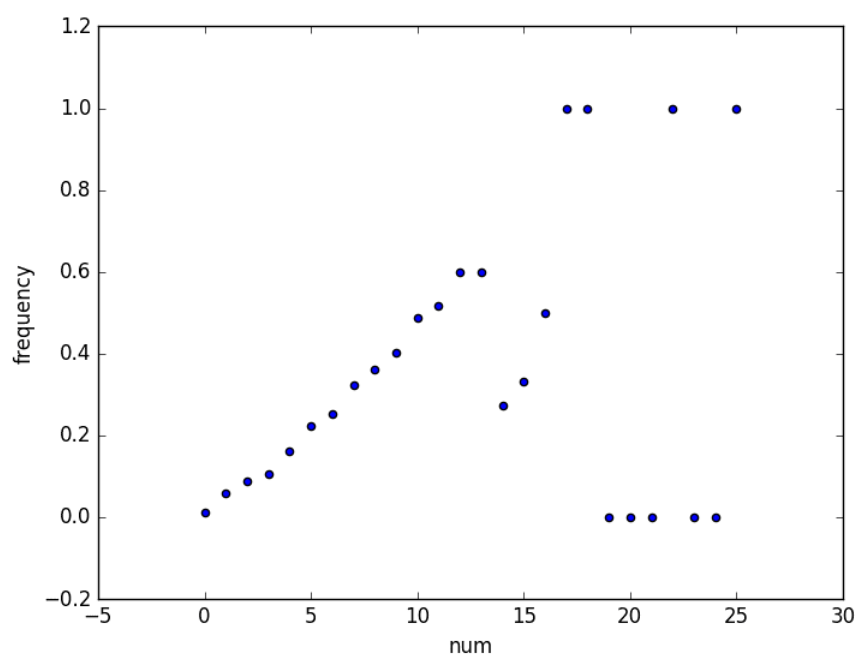
### 3.4 Feature Extraction

## Questions and Answers Type

## Overlap

## Other Conventional Methods

## Embedding



**Fig. 2.** This is the caption of the figure displaying a white eagle and a white horse on a snow field

### **3.5 Model Selection**

## **4 Experimental results**

## **5 Discussion and Conclusion**

## **References**