As the rapid development of information industry, the amount of data on the Internet is increasing at an alarming rate. Current document level search engines for instance “baidu”,”google” are becoming increasingly inadequate for a wide range of query tasks. At the same time, entity search, which has been mentioned frequently in the Internet, has just started. However, with the growth of online content, more and more people are interested in finding entities or objects instead of just documents. We more and more often include entities in the key words of our search(e.g., person number, Agency name, Place name).We always try to structure more Meaningful search conditions around entities.

information fusion technology began in the early 1970s,and rapidly developed after 1980s.It is an information processing technology which comprehensively utilizates a variety of information resources in order to obtain a more objective and better understanding the nature of a thing.

Information fusion technology researches on how to process and integrate different information from multiple sources ,and makes these different forms of information complement each other then integrate together, so that information reaches maximize.

Information fusion technology is used in a retrieval system through the trade-offs and collection division information. That makes query results organized more reasonably, and information from different information sources can be connected into an organic entirety.This makes it easy for users query to a more complete, accurate, timely and effective and also concise, straightforward information.

In this paper, the main research works are as follows:

First of all, we collecte one of the industry news text from major sites.Through web extraction technology, we organize and build a corpus in the field of Chinese-language news oriented in a particular industry .

Secondly,we research the problem of the lack of search engines and users's search habits. Based on Baidu Encyclopedia,by extracting, sorting, classifing the entries, we get a dictionary based on certain sectors. Through machine learning methods we build the industry-oriented web information fusion prototype system.It is an entity-centric information fusion system. The purpose is using the entity concept to integrate information around entities,and makes it more convenient for ordinary Internet users to use the network resources more effectively.Through the weighting process of the information, Iintegrates text information to the network entity.It achieves network information entity-centric information fusion based on one of the industries.

Finally, we calculate the text with the entity in the the processing module. On the basis of semantic understanding, weighted the degree of correlation of the text entities,and get the degree of correlation between the entities in the text. Then calculate the correlation degree of entities according to the relationship of the text which the entities appear.