様式7

<div align="center">

論 文 内 容 要 旨

</div>

報告番号	甲 先 第 **2 1 0** 号	氏 名	王 駿
学位論文題目	Research of Extracting Relations and Constructing International Network From News (ニュースから関係抽出および国際ネットワークの構築に関する研究)		

内容要旨

With the rapid development of communications and information technology, especially the emergence of Internet, information travels faster and easier than ever before. Internet extends the range and scope of communications around the world. Countries become closer and form a huge and complex international relation network. Based on the background, researchers hope to study international relations through the view of network, analyze the structure of international relation network, understand effect of different relations on international situation and reveal the evolution of international relations. It is called the research of international relation network which arouse more and more attention recently. In this thesis, we describe our research of extracting relations and constructing international network, an integral system that can construct international network from news is proposed as well.

One thing should be pointed out that network approach is a formalization method. It has some natural defect such as loss of information while facing with the non-formalization problem of studying international relations. The network approach cannot solve all the problems of international relations as well. Therefore, the main target of our research is not to offer a package that can solve everything. We try to innovate based on current work and provide a better formalization approach to deal with international relations. Our research can offer a tool to assist researchers of international relations. On the other hand, it provides an approach to help common users getting an intuitive impression about international phenomenon. Through surveys of current work, we find that there are some defects in the constructing of international relation network which is the basis of related work. Current networks are generally constructed on structured data through manual or semi-manual methods. It limits the source of data, consumes

too much manpower and time, and also cannot ensure timeliness and consistency. In order to solve these problems, we design a method that can construct international relation network from unstructured data of texts. The method achieves the goal to recognize countries and their relations automatically by analyzing texts. It expands the data source of international relation network greatly and also provides an efficient way of acquiring knowledge from information which can be used to solve the problem of information explosion. It is well known that Internet has became the richest source of information with the development of network especially Web 2.0. Thousands pieces of news are published on Internet every day. Through the news, people can learn what happened in the world easier and more rapid than even before. Meanwhile, too much information also immerses people in the ocean of information. People can hardly read all news and gain a comprehensive view about the world. It is a classic difficulty of the information that can be expressed as cannot see the wood for the trees. The international relation network transfers various events among countries into different relations and exhibits them in the form of visual graph. Graph has the feature of intuitive so that people can grasp the overview of international phenomena at a first glance. With the help of visual graph, people can also find hot spots and search what they are interested in easily. Overall, the labyrinth of international phenomena becomes clearer and easier to understand through international network.

The major innovations of this thesis can be summarized as follows.

(1) The method of constructing international relation network. As there has been no way which can construct international relation network from text automatically in current work, we originally propose our method of construct international relation network from large scale texts through text mining which is under the unified theory of information, knowledge and intelligence. It expands data source from structured data to unstructured texts and provide an efficient way to solve the problem of information explosion. We also build a comprehensive system including data acquisition, network construction and visual interface to realize our method. Through experiment, we verify the reliability of our system.

(2) Definition and extraction of international relations. Based on related research about social network and interpersonal relationship, we point that intensity, quality and status are three important features of international relations. They affect the structure and evolution of networks in different way.

We design several methods to extract these features from texts and construct the international relation network. Results of experiments show satisfied performance. Finally, we attempt to propose a new frame of defining international relations which contains five new types of relations based on the three features.

(3) Recognizing sentiment of relations between entities. Sentiment between countries reflects quality of relations which is an important feature of international relations. Although some work has been done on sentiment analysis, there is little work about recognizing sentiment between entities from texts especially in Chinese. A method composed of three steps is proposed. Through entities recognition and extraction, sentiment related region detection and sentiment determination; we can obtain sentiment between entities on sentence level. We compare different algorithm based on different principle (rule and machine learning) and different related region as well. The algorithm using CRF (conditional random fields) model based on syntactic dependency tree acquires best result.

(4) Design and realization of visual interface. We analyze defects in current visualization approaches of international relation network and propose a new method combined with GIS (Geographic Information System). The visual interface uses Google Maps as substrate, integrates Google Maps API and Mysql database and achieves international relation network visualization on electronic maps. The interface also provides some research functions that allow users to search interesting details and see different aspects of the international network more clearly.

(5) International relation network analysis. We analyze the constructed international relation network through visual interface and obtain some interesting discoveries about hot spots and characters of network structure.

様式9

<center>論文審査の結果の要旨</center>

報告番号	甲先 第 2 1 0 号	氏　名	王　　駿

審査委員	主査　寺田　賢治 副査　獅々堀　正幹 副査　任　福継

学位論文題目

Research of Extracting Relations and Constructing International Network From News

（ニュースから関係抽出および国際ネットワークの構築に関する研究）

審査結果の要旨

　本論文では，機械学習方法とテキストマイニング技術を活かしたニュースから関係抽出および国際ネットワーク構築に注目している。

　まず，国際関係ネットワークを構築するための関係獲得方法を提案した。メインアイディアはデータ・情報・知識・知能を融合したアドバンスド知能概念に基づいて，大規模テキストから関係を抽出するアルゴリズムを開発した。ここに，テキストマイニング技術として，従来の構造化データから非構造化データまで拡張された方法を研究した。実験結果から本研究で提案した関係抽出手法の有効性を確かめることができた、

　次に，国際関係の定義と国際関係の抽出を行った。特に，関係性質・関係強度など国際関係の特徴について，大規模なデータを分析した上で，様々な考察を行った。また，５つの関係の新しいタイプを提案し，これらのタイプに基づく国際関係構築手法を試みた。

　さらに，関係抽出に感情要素を導入し，感情認識アルゴリズムを開発した。国際関係は人と人との関係に似ていると仮定し，感情要素が大きな役割に立つと考えられる。よって，研究室で構築された中国語感情コーパスを活かし，感情認識アルゴリズムを開発した。特に，構文的従属性木に基づいたＣＲＦ（条件付きのランダム・フィールド）モデルを利用したアルゴリズムは最良の結果を得た。

　最後に，国際ネットワークを構築し，ビジュアルインターフェイスを開発し，検証を行った。ここに，従来の研究で実現できなかった国際関係の可視化の困難点を分析し，GIS（地図情報システム）と融合した手法を提案した。

　提案された手法に基づいて，実験システムを構築し，様々な評価実験を行ったが，提案された手法の有効性を確かめることができた。

　以上本研究は，当該分野の既存の問題を解決した貢献から価値のある研究であり，本論文は学位論文としての水準を満たし，博士（工学）の学位授与に値するものと判定する。