Intelligent Decision Support System Based on Natural Language Understanding

Yu Xingang
Auditing Department
Liaocheng University
Liaocheng, China 252000
yuxingang@lcu.edu.cn

Zhou Faguo, Zhang Fan
School of Mechanical Electronic and Information Engineering
China University of Mining and Technology
Beijing, China 100083
{zhoufaguo, zhangfan}@sina.com

Yang Bingru
School of Information Engineering
University of Science and Technology
Beijing, China 100083
bryang_kd@yahoo.com.cn

Abstract—In this paper, we simply introduced the traditional decision support system and its characteristics and limitations. And then some new types of intelligent decision support system are recommended, such as GDSS, GDSS, 3IDSS and IDSSKD. And at last IDSS based on natural language understanding are introduced detailedly. The prospects of IDSS based on natural language understanding is put foreword.

Keywords-intelligent decision support system, knowledge discovery, natural language understanding

I. INTRODUCTION

Decision Support System (DSS) [1], [2] is an information system that participates and supports the decision-making process of human beings. Through a series of human-machine conversations, it can provide a variety of dependable programs and check the decision-makers’ requirements and assumptions, thus achieving the aim of supporting decision-making. Until now, people have not reached an agreement upon the definition of DSS. Generally speaking, it is an interactive computer information system [3] that can help decision-makers make use of data and models and solve the problem of non-structures. It makes full use of suitable computer techniques and through the interactive human-machine model, helps and enhances the effectiveness of decision-making about semi-structures and non-structures [4].

The rest of this paper is organized as follows. Section II describes the traditional decision support system in briefly. Section III discusses some new types of intelligent decision support systems. Intelligent decision support system is introduced detailedly in section IV. Section V concludes the paper and prospects future work.

II. TRADITIONAL DSS

A. Basic Features of DSS

The basic features of DSS can be summarized in five parts as follows:

1. It is oriented toward the lack of structuralization and insufficiency of explanation that the managing staff often faced.

2. It combines the model or analytic techniques with the traditional techniques of data storage and information retrieval.

3. It can be easily used by non-professionals of computers in the mode of interactive conversations.

4. It emphasizes the flexibility and applicability to adapt to the changes of the environment and decision-making.

5. It can support (assist) rather than replace the decision-making of the high-ranking managing staff.

B. General Solving Process of DSS

Users input the description of the problem and requirements into DSS through the interactive language system. The interactive language system will recognize and explain this. The problem-solving system collects the data, information and technology about this problem through the knowledge base and database systems, and then according to this, it recognizes and judge the quality and solving process of the problem. It makes use of the model base system and builds the regular or mathematical model that is needed to solve the problem. In the method base, it recognizes the algorithm for the model solving, then it solves the model and evaluates the result. In the end, through the language system, it explains the result and gives out the output that has actual meaning and can be understood by the users. In the above solving process, if need be, users can interact with DSS and repeat the process for several times until a satisfactory result is achieved.

Intelligent decision support system (IDSS) is a result of combining decision support system (DSS) and artificial intelligence (AI). Its basic design thinking is to combine the knowledge reasoning techniques of AI and the basic function models of DSS.
III. SOME NEW TYPES OF IDSS

In order to adapt to different situations, people develop the IDSS from different aspects and in different ways. This can be summarized as follows:

A. Group Decision Support System (GDSS)

It supports group or collective decision-making; it combines communication techniques (net, telephone meeting, tele-transfer of information), computer techniques (multi-users system, 4GL, database, data analysis, OLAP, data storage, data warehouse, data mining) and decision support techniques (agenda setting), AI and reasoning techniques, decision-making model methods (decision tree, risk analysis, forecasting method etc.), structuralization group decision methods (Delphi method etc.)

GDSS [5] can provide three levels of decision:

At the first level, the GDSS aims at reducing the communications between decision-makers, connecting information and eliminating the obstacles of communication, such as the large screen showing various opinions, the voting and ballots collection equipments, the input of unanimous opinions and preferences, the telegraphic communication between members, etc. Its aim is to improve the exchange of information between members in order to improve the decision process. The so-called “telegraphic meeting system” falls into this category.

At the second level, GDSS provides the structural technique that is good at recognizing the processes and the dynamics of the system and the technique of choosing the analytic model and the judgment method for the decision process. In this system, the decision-makers usually work face to face, share the information and make the work plan collectively.

At the third level, GDSS mainly combines the techniques of the above two levels and employs the computers to inspire and instruct the communication method of the public, including the intelligent arrangement of expert consultation and meeting rules. This kind of high-level system is still at the level of experimental development.

B. Distributed Decision Support System (DDSS)

With the development of DSS, it is natural for people to expect the support of computers at a higher level of decision-making and at a more complex decision environment. This kind of support is oriented not only toward individual decision-makers or any decision-making group that stand for the same organization, but also toward decision-making organizations that have separate identity and are connected at the same time. It is impossible or in convenient to carry out some of the large-scale decision-making activity. These activities involve decision-makers that have different responsibilities. The information or important decision factors needed for the decision-making process are distributed among a larger area and this decision-making is a kind of organization decision-making or distributive decision-making. DDSS is an information system that is built to solve this kind of decision-making problem.

DDSS [6] is a computer net and is made up of many physically separate information processing nodes. Every node of the system includes at least one DSS or several functions that aids the decision-making. In contrast to the ordinary DSS, DDSS has the following characteristics:

As a specially designed system, DDSS supports the multi-level decision-making at different nodes and provides support or individuals, groups and organizations. It can not only provide support from one node to another, but also provide instructions and explanations for the results. The information can be shared in a more efficient way.

It can provide the mechanism and tools of communication between nodes and support human-machine, machine-machine and human-human communication. It has the ability of dealing with the possible conflicts between nodes and coordinating the operation between nodes. It has a severe internal contract and is open on the other hand. The expansion of the system or nodes is allowed, and meanwhile the nodes within a system are equal members and do not form a stratification relationship. Every node is autonomous.

C. Intelligent, Interactive and Integrated Decision Support System (3IDSS)

With the expansion of the applied field of DSS and the rising of the application level, DSS has entered into the decision-making activity of regional economic and social development strategy and the manufacturing and operating decision-making of large-scale enterprises. Not only do these decision-making activities involves all the aspects of economic activities and all the levels of management, but also all the factors are connected with each other and the decision-making environment is complicated. As for the application fields of the strategic planning of development such as the provinces, cities and counties, etc., the decision-making activity is also affected by political, social, cultural and psychological factors, etc. And the information available is incomplete and inaccurate. All these add to the difficulty of DSS. Under this environment, a new comprehensive decision-making DSS came into being, which is oriented toward decision-makers and the decision-making process. That is Intelligent, Interactive and Integrated DSS, abbreviated as 3IDSS [7] [8].

D. INTELLIGENT DECISION SUPPORT SYSTEM BASED ON KNOWLEDGE DISCOVERY (IDSSKD)

With the increasing speed of enterprise information, more and more resource including information flow, fund flow, human resource and so on is greatly restricted with existing information framework. Every kind of useful information has become a detached information island. All kinds of enterprise businesses (such as ERP, CRM, SRM and DRP, etc.) need urgently integrating in order to realize the sharing and interaction of information resources. However, how to utilize knowledge discovery to find a large amount of knowledge contained in countless information resources, has become the key point whether various kinds of integration information could play an important role. So, the organic integration of data mining systems and the other intelligent systems (such as DSS), especially of the enterprise business, which some users have already acquainted with, is very important for system to
exert full functions. And the overall structure model and its core components are described in the paper [9].

IV. INTELLIGENT DECISION SUPPORT SYSTEM BASED ON NATURAL LANGUAGE UNDERSTANDING

According to statistics, among the present digital information, 80% exists in the form of non-structures, and the majority of the non-structural information exists in the form of texts that are made up of natural language description. Therefore, DSS must analyze and deal with the natural language texts in a thorough way.

The IDSS based on natural language understanding combines the net searching technique and the technique that deals with information intelligently. It makes use of the technology management method to automatically collect, categorize, cluster, detect and focus topics from the large amount of information on the internet in order to detect and control the cyberspace public opinion and follow the news topics etc. IDSS then combines this with expert information in this field and gives out reviews, reports, diagrams, etc. Thus, it directs the opinions of the public in order for the users to understand the thinking trend of the public and provides the decision-makers with decisions and evidences for their analysis.

A. The Main Techniques

1) The automatic word segmentation technique

The automatic word segmentation technique is the basis of Chinese information processing and analysis. The wrong word segmentation will lead to completely wrong information semantics. The IDSS which is based on natural language processing holds the dictionary and the rules as the basis and makes a comprehensive use of the language model method which is based on probability analysis, thus making the rate of correctness 99% for the word parsing. And different parsing can be carried out to adapt to different applications.

2) Automatic keyword extraction and automatic summarization technique

After an analysis of the text meanings, the IDSS which is based on natural language processing combines the information of word frequency, word class and word position in order to obtain the keywords and summarization automatically. At the same time, the IDSS can make use of the technique of anaphoric resolution to make the abstract readable.

3) The full-text retrieval technique

The full-text searching engine of the IDSS which is based on natural language processing combines the traditional text retrieval technique and the latest WEB searching technique, thus increasing the function standards of the retrieval engine. At the same time, several related techniques are integrated. A rich variety of retrieval tools and retrieval methods such as synonyms can be provided.

4) Automatic classification and automatic clustering

Classification will study the training datasets and obtain the model for every class. Then the model will be used to classify the data that are unclassified. A typical application is that the system drills upon the classified documents and then classifies the document collection automatically. It differs from clustering in that it has a process of studying.

Clustering will put the data in the data collection into subgroups that have a certain meaning. The difference of the data in the subgroups will be made as large as possible. And the difference of the data in the same subgroup will be made as small as possible. A typical application is the data nodes in a spatial database will be clustered as different groups according to their spatial distribution.

5) Web Text Classification

There are thousands of new trading contracts in China, and how to categorize them efficiently for support decision is far beyond the ability of existing statistical software used by E-Commerce Center of Ministry of Commerce of the People’s Republic of China. Furthermore, with the rapid development of World Wide Web, more and more contract data are collected via Internet, so it is urgent to design efficient content-based retrieval, searching and filtering for the huge and semi-structured online repositories on the Internet.

Text classification, the assignment of free text documents to one or more predefined categories based on their content. Each document can be in multiple, exactly one, or no category at all. Text classification has many application areas, such as information management, real-time sorting of emails into folders, topic identification to support topic-specific processing operations, structured search and/or browsing, or finding documents that match long-term standing interests or more dynamic task-based interests, etc. In the research community the dominant approach to it is based on machine learning techniques: a general inductive process automatically builds a classifier by learning from a set of pre-classified documents, the characteristic of the categories. This system adopts the fusion of K-NN [10].

6) Web Text Clustering

The above-mentioned text classification technique requires a large number of labeled training examples. However, in certain cases, it is difficult to label some of them. Clustering can be applied to these collected texts. Hierarchical clusters of these unlabeled documents can be generated. Based on this, categorization can be operated indefinitely. Clustering via Self-Organizing Maps (SOM) is adopted.

The SOM is one of the major unsupervised artificial neural network models and often used to learn certain useful features found in their learning process. It basically provides a way for cluster analysis by producing a mapping of high dimensional input vectors onto a two-dimensional output space while preserving topological relations as faithfully as possible. After appropriate training iterations, the similar input items are grouped spatially close to each other. As such, the resulting map is capable of performing the clustering task in a completely unsupervised fashion. Furthermore, SOM approach is superior to other cluster analysis methods in data mining in terms of the power of data visualization. Thus, in this work SOM method was adopted to produce document cluster map for Web text mining.
7) **Similarity searching and duplicated webpages deletion**

Similarity searching is to find in the target collection the similar examples or target subgroups that are similar to what are being retrieved. For example, in the WEB, webpages, pictures or other information similar to the wanted content will be found.

In practical application, to the decision-makers, as for the same or largely similar documents, only one is enough. It is redundant if there is more than that. This has something to do with the duplicate reduction or filtering of the documents or webpages.

### B. The Main Characteristics

1) **Increase the work efficiency and enlarge the detecting and control area**

The IDSS based on natural language processing will collect and analyze the information on the internet. It makes use of the intelligent processing technique of the computer to collect and analyze the cyberspace public opinions. This not only increases the work efficiency and improves the guard and control of the opinions, but also amplifies the number of detecting and control of the media. The area of detect and control is enlarged and the high quality is achieved. The analysis and detecting and control of a large area have increased the credibility and accuracy of the detecting and control result. The cost of human resources is also reduced.

2) **Create intelligent opinion report**

The IDSS based on natural language processing will take advantage of the research results of the technology project institute of the University of Science and Technology Beijing. It combines the techniques of content management, technology management and related internet techniques. It can create an intelligent opinion report according to the hotspot social topics in a certain period. The opinion reports are vivid in pictures and texts and will show directly the hot topics in society and on the internet. A special opinion report will analyze and look into the important hot news. As for the urgent net opinions, the focus of the opinion and the development trend of the situation will be obtained.

In addition, by means of this system, the artificial mistakes will be decreased on a large scale. After the collection, organization, detection and analysis of the internet information sources, the artificial mistakes will be decreased. And a high-quality and high standard detection and control report will also be created. This will be beneficial for the related government departments to grasp the social thinking trends and provide dependable analytic evidences for them to make corresponding work plans.

3) **Help judge the positive and negative information on the internet news and commentaries**

The IDSS based on natural language processing support the artificial judging and annotating of the characteristics, such as emergency rate, secrecy rate, opinion trend, etc. Information can be arranged according to the characteristics of the information, such as time, opinion trend, etc. Through the systematic statistical and analytical function, it can carry out the statistical analysis of the positive and negative information about social emergencies and commentaries on the net and give out the straight diagram about the development trend which shows the characteristics of the net.

### C. The Prospects of the Application

The combination of the DSS with artificial intelligence and the internet has greatly increased the flexibility and expansionability of the traditional DSS. It has integrated the advanced technical tools of data mining and technology discovering, vague processing and natural language understanding, etc. The IDSS based on natural language understanding has become an interactive human-machine system that has a strong learning ability and adaptation. This is also the main development mode of the present DSS.

At the same time, the propaganda departments may effectively regulate the information of the net and advocate the healthy and beneficial consensus. This system has greatly put forward the development of the control and management of the net information, the organizing of the ordering and analysis of the information, dealing with the emergencies on the net and grasping in an all-round way the public opinion and situation.

### V. CONCLUSIONS

In this paper, we simply introduced some new types of intelligent decision support system. And then IDSS based on knowledge discovery and IDSS based on natural language understanding are described in detail. The prospects of IDSS based on natural language understanding is put foreword.

### ACKNOWLEDGMENT

This work is partially supported by National Nature Science Foundation of China (Grant No. 60675030).

### REFERENCES


