Construction of Strategic Precision Planning Information System Based on Big Data Analysis

Ping Xu

School of Information Science and Technology Taishan University, Taian, China

Keywords: data mining; strategic precision planning; data integration; business integration

Abstract: In order to improve the efficiency of strategic precise planning information system, on the basis of analyzing the current strategic precision planning business needs, a component-based solution of strategic precise planning information system based on big data mining, computer and database technology is designed, and the overall framework of the system and an integrated database model used for volume data sharing are constructed. The strategic precision planning information system is realized by programming. The practice shows that the integrated system is of practical value for the planning work of relevant departments.

1. Introduction

The 21st century is a new digital era. The whole human society has entered the information digital society. The development of the information society must rely on the development of advanced network computer technology [1]. Internet technology enables enterprises to communicate with customers at zero distance, so that they can successfully complete transactions on the network. The explosive expansion of the Internet has made it a revolutionary technology in the new millennium, and has also brought blowout growth of data.

2. Method

Data mining technology originated in the 1980s. When the investment in artificial intelligence (AI) research project failed, AI was put forward in practical application. It is the result of people's long-term research on database technology [2]. Data mining is a new technology and it is the practical research of artificial intelligence for commercial applications [3]. At the beginning, business data was stored in computer databases, then developed to be able to access and query data, and developed to be able to effectively mine the correlation between data [4]. The concept of data mining is defined because it is related to statistics, actuarial, mathematical modelling and other fields. It enables database technology to enter a more advanced stage, to find out the potential relationship between data in massive data, and to promote the application of information at a higher level.

Data mining technology platform includes three main contents: algorithm and technology; data; and modelling ability. Figure 1 below is a schematic diagram of a logical process in which data is transferred on a mining platform and then transformed into usable information.
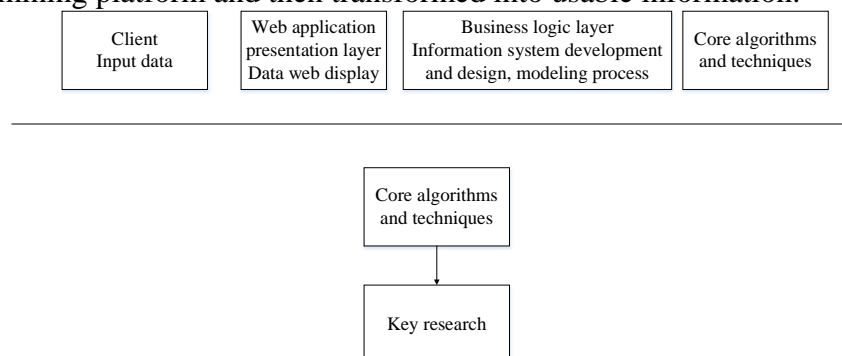

Figure. 1 Schematic diagram of information processing logic process

As can be seen from Figure 1, the core algorithm and technology used in data mining is the most central part of data mining platform, and the algorithms and technologies closely related to data mining mainly include:

First, machine learning. Machine learning includes two learning methods: neural network and decision tree algorithm. Neural network is a kind of self-organizing learning, while decision tree algorithm is to summarize rules from examples, which is a combination of computer and AI. Neural network algorithm is mainly used to solve complex problems and it consists of a large number of parallel and distributed processing units. Through learning experience in daily life, the connection strength between these experiences is adjusted, and then its empirical principle is applied to practical problems. The storage system of neural network is the physical connection between neurons to store knowledge and information, and to represent and store information and knowledge on neurons and their connections. Each neuron and its connections can only represent a part of information and knowledge. Only by connecting the whole neuron in a distributed way can a complete and specific concept and information be represented. Artificial neural network has many similarities with biological nerve tissue. It has a large number of neurons and can store a large amount of information. Therefore, the efficiency of neural network for data processing is particularly high, and it also has a strong ability to deal with uncertain information. Neural network simulates human thinking logic, so even if the information is incomplete, the neural network can extract information from the past memory to think. As long as the samples are input to the neural network, the system can calculate the correct conclusion according to the similar pattern with the stored information. The processing output of neurons is not arbitrary. Only when the synthesized calculation results of neuron for all the input information exceed a certain threshold will the neural network draw a conclusion on the input problem.

BP algorithm (error back propagation) is a kind of algorithm similar to neural network. The learning process is backward propagation, which processes a set of selected training samples by iteration, and then compares the prediction of input sample information with the actual knowledge to learn. By modifying the threshold of experimental samples, the difference between predicted value and actual knowledge is the minimal. This "backward" propagation mode is carried out from the output layer through the hidden layer to the input layer. This mode divides learning into two parts: first, the input information is inputted forward (input layer - hidden layer processing - output value). If the output value is not the expected value, then the error between the output value and the real value is calculated step by step. Adjusting the full value according to the error value is a reverse process. Because of its robustness, self-organizing adaptability, distributed storage, parallel processing and high fault-tolerant rate, the neural network is very suitable for solving data mining problems.

Decision tree is also a common machine learning algorithm. It is a decision-making process based on tree graph, usually for making a decision. This algorithm first needs to determine a classification-oriented model, classify the problems, classify them according to branches, and gradually solve the problem of data mining. The advantage of the decision tree algorithm is to show the whole judgment process, which is very intuitive and easy to understand.

Second, statistics. Statistics can support data mining technology, which includes sampling, prediction algorithm (regression), experience-based design and so on.

Third, decision support system (DSS). DSS is based on management science, operational research, cybernetics and behavioural science. DSS means are computer technology, information technology and simulation technology. Aiming at semi-structured decision-making problems, it supports intelligent man-machine system for decision-making activities. DSS can provide a series of data services for enterprises, including establishing decision objectives, determining research problems, constructing or improving decision models, providing alternatives and evaluating and choosing alternatives. It can also provide a series of scientific, effective and reliable decision schemes for decision makers through human-computer interaction, which can meet the requirements of decision-makers and achieve the goal of scientific decision support.

DSS generally consists of six systems, including interactive language system, problem system,

database, model base, method base and knowledge base management system. In some special DSSs, there may be no separate database and management system, but there must be model base and method base. There are many kinds of problems that need to be analyzed, so according to the actual situation, there are many kinds of structure of DSS.

Some people mistakenly think that DSS is to automate decision support for decision makers, which is wrong. DSS emphasizes decision support more than automated decision support. It can support decision-making at all levels, such as strategic level, functional level and business level.

Fourth, data warehouse. Data warehouse is a special form of database. It is a data set that can support the decision-making process at all levels of enterprises. It is built for decision support and specific analysis process and it stores a single data body and provides business process improvement, cost savings and quality control for enterprises requiring artificial intelligence.

Data warehouse is a data environment, a structured data environment for DSS and on-line analytical application data sources. Data warehouse is to obtain data from database to solve related problems. Its main characteristics are subject-orientation, stability, time-varying property and integration. Data warehouse is a data analysis process, not just a project, but a system that can provide decision information. Data warehouse system obtains data from business processing system. Data storage is mainly stored in the form of snowflakes and stars. It can provide users with various ways to obtain information and knowledge from data. The structure of data warehouse system mainly consists of three parts: data acquisition, data storage, and data access.

The architecture of data warehouse includes the following parts:

Data source: It is the basis of the whole data warehouse architecture. The source of data is the data source, including all the data inside and outside the enterprise. The data inside the enterprise is stored in RDBMS, including all kinds of business processing data and all kinds of document data. The external data is mainly the information of competitors, markets and potential customers and national laws and regulations.

Data storage and management: The core of the whole data warehouse system is data storage and management, which plays a key role in the whole data management. Different from traditional databases, data warehouse has its own unique organization and management framework and different expression forms of external data. Starting from the technical characteristics of data warehouse, suitable products and technologies are chosen to build data warehouse. Enterprises can build different data warehouses for business purposes, extract, clean up and integrate data effectively. Data warehouses can be organized according to the research issues. The data is divided into two parts: enterprise data warehouse and department data warehouse, among which department data warehouse is also called data mart.

OLAP server: Online Analytical Processing, whose function is to integrate data effectively according to multi-dimensional model, carry out multi-level and multi-angle analysis, find out the existing laws or trends, and then present them to users through more convenient use. ROLAP (Relational Online Analytical Processing), MOLAP (Multidimensional Online Analytical Processing) and HOLAP (Hybrid Online Analytical Processing) are three types of data storage. The basic data and aggregated data of ROLAP are stored in relational database, the basic data and aggregated data of MOLAP are stored in multidimensional database, the basic data of HOLAP are also stored in relational database and aggregated data are stored in multidimensional database.

Front-end tools mainly include various reporting tools, data analysis tools, query tools, data mining tools and various application development tools based on data warehouse or data mart. In particular, data analysis tools are mainly for OLAP servers, while report tools and data mining tools are mainly for data warehouse.

3. Results and discussion

The composition of strategic precise planning information analysis environment is the result of the integration of data warehouse, OLAP, data mining technology and so on. Strategic precision planning information system involves the display of interface, human-computer interaction, organization and access of planning data, query and statistics of planning information and so on.

Considering the shortcomings of traditional two-tier structure mode in operational efficiency and function, the system uses a three-tier structure with clear function division and logic independence, that is, data layer, logic layer and application layer. As shown in Figure 2, compared with the traditional two-tier model, the three-tier structure embodies the integrated database management system, GIS and MIS application components and functional modules based on secondary development language, and the integrated system application platform, which constitute the back-end support, middleware and front-end display of the system physical framework. The hardware structure is more flexible, and each layer can choose hardware for parallel development and implementation which is suitable for its processing load and processing characteristics, so that the logic of the program becomes simpler and the maintainability is greatly improved.
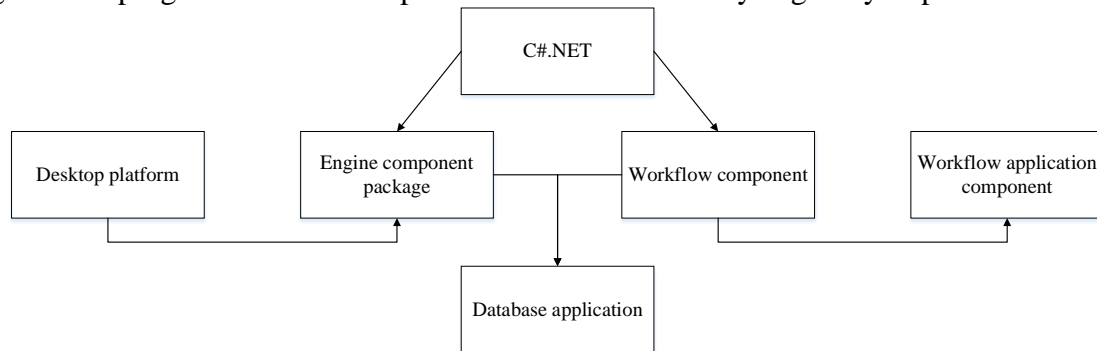

Figure. 2 Schematic diagram of system framework

Data layer. Data layer is the bottom structure, which is composed of the original database management system. It integrates and manages multi-source heterogeneous spatial data and attribute data in the form of a unified database to provide basic data support for the whole system. Strategic precise planning business needs to support large-scale relational database with mass storage and high security as data layer. Therefore, the system constructs integrated database based on SQLserver2005 relational database management system, centrally manages various types of sub-databases. Moreover, it processes the access commands of application layer users through interconnecting large-scale spatial database engine user ArcSDE, open database ODBC and secondary development components, calls and analyses the spatial and attribute data of multi-database, and returns the result information to users through the form in the system interface.

Logic layer. Logic layer is the core part of three-tier structure, in which components and modules encapsulate all functions of operating integrated database. Users pass command requirements to logic layer through application layer, call C. NET secondary development of layered application components, and use user JcSDE and ODBC to interact with data layer. The results will be displayed in the application layer for users' use or further interactive processing. The application layer is generally designed as a simple and easy-to-use graphical user interface. If the user's needs change and the operation function interface changes, it is only necessary to modify the application code in this layer, without involving the other two layers. On the basis of ensuring the data safety, the work efficiency is greatly improved.

Application layer. Application layer, as the human-machine interaction carrier of the system, is a set of user-oriented applications. In order to enable users to operate intuitively, the application layer should be designed as a simple and easy graphic menu mode. Users can send requests and get feedback information to the logic layer through menu and window operations, without paying attention to the specific implementation process, which is conducive to users concentrating on the required information and interactive operations. The application layer of the integrated system adopts the mainstream Windows main form mode, which is convenient for users at all levels and departments to use.
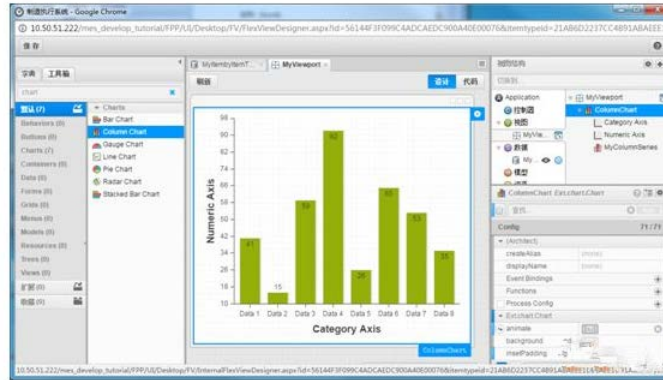
Figure. 3 System implementation diagram

4. Conclusion

Data engine and hierarchical component structure are introduced into the integrated model of strategic precise planning. The key technologies such as computer and database used in the implementation of the integrated system are discussed in detail. On this basis, the framework and functional modules of the integrated information system of strategic precise planning are designed, and the integrated database of strategic precise planning is constructed. The model completes the sharing and calling of multi-source heterogeneous data and different system functional modules. Finally, the secondary development and programming of integrated system is achieved.

References

[1] Wang Z, Wang Z, Li Y. Strategic planning for national biomedical big data infrastructure in China [J]. Quantitative Biology, 2017, 5(3), pp. 272-275.

[2] Sun L, Wang Z. Identifying the Core Competitive Intelligence Based on Enterprise Strategic Factors [J]. Journal of Shanghai Jiaotong University, 2015, 20(1), pp. 118-123.

[3] Dobrzykowski D D, Leuschner R, Hong P C, et al. Examining Absorptive Capacity in Supply Chains: Linking Responsive Strategy and Firm Performance[J]. Journal of Supply Chain Management, 2015, 51(4), pp. 3–28.

[4] Zhang S, Construction D P. An Analysis of the "Internal Relations in Four Comprehensive Strategic Framework"- Based on the In-depth Thinking - the Eighteen Session of the Fourth Plenary Session Comprehensively Promote the Rule of Law [J]. Journal of Ningxia Communist Party Institute, 2015, 52(3), pp. 143-151.