TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.2, February 2014, pp. 1100 ~ 1108
DOI: http://dx.doi.org/10.11591/telkomnika.v12i2.4411

Research on Collaborative Mechanism of Data Warehouse in Sharing Platform

Quanhong Sun*, Qi Xu
China University of Water Resources and Electric Power, China
*Corresponding author, e-mail: sunquanhong@ncwu.edu.cn

Abstract

With the development of distributed computing technology, there remain some important problems of collaborative control in distributed multi-user data sharing: heterogeneous environment, data consistency and concurrent operation. To solve the sharing issues in drilling engineering, this paper introduces a method of realizing information sharing by combining the advantages of multi-agent collaborative working mechanism and data warehouse technology. Firstly, the structure of multi-agent collaborative working mechanism is put forward based on distributed data warehouse, in which the underlying data layer adopts the distributed data warehouse to organize the information, and the model of that is set up by dimensional modeling technology. Then the information sharing platform and multi-user access controlling method are analyzed, and the collaborative model is put forward. Finally through the multi-agent interaction, the management and sharing of all kinds of information in DDW can be realized. So the application of multi-agent and DDW technology can maximize the value of drilling information and play an important role in improving data sharing degree.

Keywords*: multi-agent, collaborative working mechanism, distributed data warehouse (DDW), collaborative drilling, information sharing platform*

1. Introduction

The drilling projects have the nature of large investment and high risk, so the technology used in the decision analysis and target solutions made by senior professional, technical staff and manager are directly related to the efficiency and success of construction. In addition, with the development of drilling technology and the continuous expansion of drilling scale, there is a steady rising in the quantity of drilling information, therefore the information is getting much more valuable undoubtedly. But most information is operational data which is of redundancy, at the same time, for complexity drilling engineer, the information from different oil fields or different departments in the same oil field is distributed and heterogeneous, even the same kind of data in different drilling information systems has a bigger difference, so the level of information sharing is low and the information management is difficulty settled [1]. With the development of distributed computing technology, to make full use of a mass of historical information in drilling engineering to reflect the value of historical information, relevant departments of each oilfield fully share information to provide accurate, scientific and reasonable decision information for the decision makers. This article puts forward building sharing platform by combining the advantages of multi-agent collaborative working mechanism and data warehouse technology to integrate the distributed information and maximize the value of a lot of complex information of drilling.

2. Distributed Data Warehouse

Data warehouse (DW) is defined as a subject-oriented, integrated, and stable and variation data which sets along with time to support enterprises or organizations to make decision by W.H. Inmon who is known as "the father of data warehouse". He believes that there are many kinds of forms of the distributed data warehouse. One of them is applied in drilling engineering with the typical structure shown in Figure 1, which indicates the local and global data warehouse [2].

Received August 17, 2013; Revised September 12, 2013; Accepted October 12, 2013

2.1. Local Data Warehouse

Local data warehouse contains various history and integrated data that comes from the actual operating systems in oil field or departments. The detailed degree of data is relatively high, so the level of granularity is smaller.

2.2. Global Data Warehouse

The data of global data warehouse mainly comes from all local data warehouse which is the public and integrated data in the system, including external data sources and internal public, history and integrated information in the whole system. Data source of global data warehouse is not only used to accumulate, but also to reorganize, synthesize and integrate data of the local data warehouse.

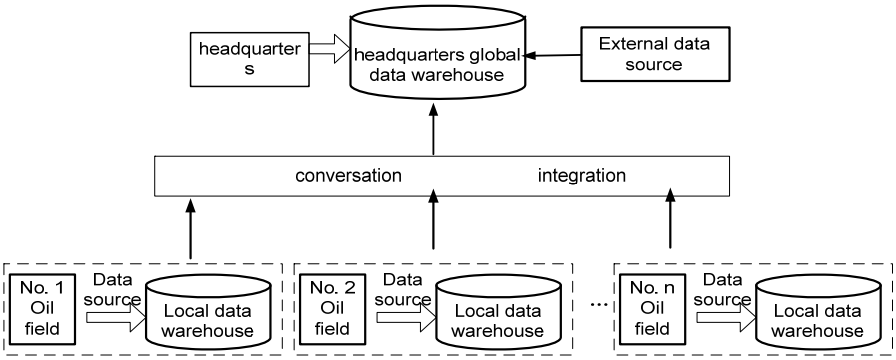

Figure 1. Structure of Distributed Data Warehouse

3. Design of Drilling Cooperative Working Mechanism Based on Multi-agent

3.1. Structure Design of Cooperative Working Mechanism

The computer network and data warehouse technology can solve the information sharing problems to a certain extent, but far from meeting the needs of users and enterprises, so it is necessary to establish a drilling cooperative working mechanism to achieve the information sharing by multi-agent. According to the analysis on existing systems and the new needs of the drilling engineering, this article puts forwards establishing drilling cooperative working mechanism on the distributed data warehouse. The drilling cooperative working mechanism is a distributed information sharing environment which can realize document management, collaboration management, application program sharing and user management functions, and be responsible for data sharing services for the headquarters or subsidiaries of each oilfield and relevant departments of government. The drilling cooperative working mechanism adopts a multi-layer framework structure model, which is shown in Figure 2.

The model relies on a distributed data warehouse including collaborative query and collaborative modification of drilling data warehouse by multi-agent collaboration. From Figure 2, the collaboration layer (application server layer) mainly consists of document management agent, application program sharing agent, collaboration management agent and user management agent. Application sharing agent is responsible for coordination, management and distribution of the shared information among the various many experts, and such information will be given the highest priority during transmitting; Document management agent is mainly responsible for the unified organization and standardized management for all documents in the system. In drilling engineering, documents are generated by the collaborative group, and all documents in the data warehouse are managed from the overall perspective and assigned the appropriate and allowable level that includes the browsing right, the right to modify and the right to write in order to effectively share documents. Collaborative management agent is mainly responsible for the allocation of communication resources and service quality management for all members. User management agent is mainly in charge of user groups and their rights management.

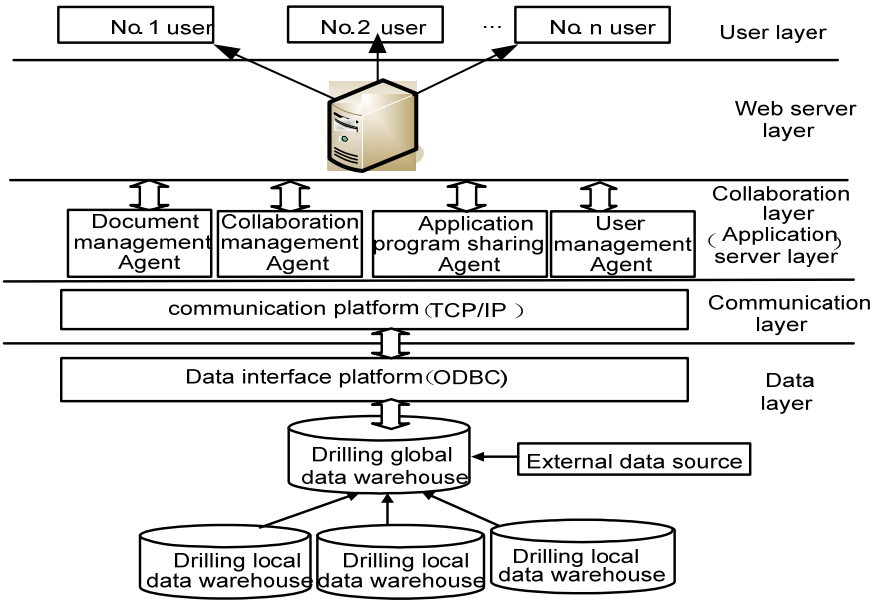

Figure 2. Structure of Drilling Cooperative Working Mechanism

For each specific member, manager can establish its account and manage related information. In addition to the addition, deletion, modification, and query of the accounts of regular members, the more important is to carefully arrange the role of various members. Collaboration drilling engineering is achieved through mutual cooperation among the several agents [3-4].

The data layer is the data source of information platform, which provides an underlying distributed data warehouse system for the platform operation.

According to the important degree of the information, communication layer creates a multi-function hall dedicated to drilling for comprehensive exchange by using different control strategies in a limited network bandwidth, and the platform mainly uses reliable connection-oriented data transmission channel based on the TCP/IP protocol.

In this model, the status of each local data warehouse is equal, so a local data warehouse will not affect the operation of the entire system. The local data warehouse nodes of system must firstly apply to system administrator for adding the resource provider. After the system administrator confirms that, it can become a node of the system (local data warehouse).

3.2. Cooperative Working Model of Multiple Agents

The multiple agents and many staff work together to complete the drilling process, based on the information sharing platform model. A collaborative model of multiple agents and many users can be obtained, which is shown in Figure 3 [5].

In this model, the functions of all collaborations and agents are listed as follows.

Human-human cooperation: humans who participate in collaboration use collaborative tools to work with the aid of design professionals to jointly complete a design task. According to the thoughts of concurrent engineering and collaborative engineering, people are from the different departments and professional of collaborative groups, who are project director, designers, technologist, maintenance staff, and even ultimate users of products. Different people have different professional knowledge and experience to complete different collaborative tasks.

Tool-tool cooperation: through respective communication interfaces of agents, collaborative tools can achieve network communication and information exchange in order to realize tool-tool cooperation. Because different staff complete different tasks, they need different collaborative tools of difference functions, accordingly, a collaborative system should provide a variety of application tools with different types and different function for collaborative users, such as electronic whiteboard, chat rooms, file transfer, audio video collaboration tools.

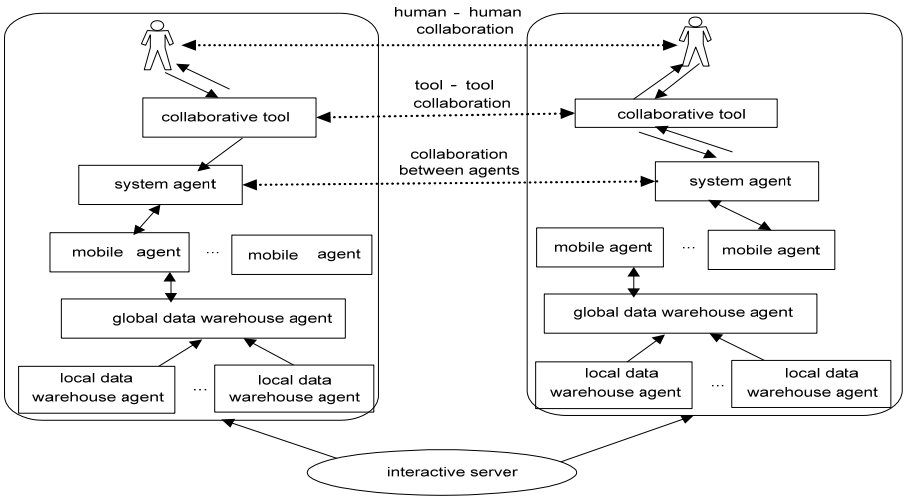

Figure 3. Cooperative Working Model of Multi-agent

Cooperation between agents: communication agent achieves messages exchange through the interactive server, the interactive server is a relay station of all communication messages.

System agents are static agents that resident in the interactive server, they are the operation center of the system, which can assign many mobile agents to each local data warehouse by the global data warehouse agent, and return the results to the collaborative users.

Mobile agents are assigned to the all local data warehouse by system agents of the server and carry information to be processed, which can communicate with local data warehouse agents, and can return the results to system agents.

Global data warehouse agents are responsible for receiving mobile agents assigned by system agents, and responsible for handing out the query information to multiple local data warehouse agents.

Local data warehouse agents always resident on the server, and receive all the mobile agents carrying access requests, which can execute the requests and return the results to mobile agents by message.

After a user sends an operation request to the system agent, according to the global agent, the system agent sends multiple mobile agents to the local data warehouse agents, and then the mobile agent sends the results of operation request back to the system agent by a global data warehouse agent, finally the system agent shows the operating results to the users, realizing the information sharing among the multi-user. Agents achieve message exchange through interaction server, and multiple users cooperate to operate on the distributed data warehouse, and what you see is what I see (WYSIWIS) is realized in the resulting output [6]. Specific working process is described in 5.

In Fig.3, as the collaborative assistant, system tools and agents can help people to accomplish the collaborative tasks.

4. Information Sharing Platform

The core issue of collaborative working mechanism is how to provide effective support to help user to use computer technology to cooperatively complete common design tasks and achieve goals, so the information sharing platform is necessary to make people conveniently share the information resources locally or on the network [7-8].

4.1. Information Classification of Sharing Platform

Seven types of information referring to member, task, collaborative design activity, document, multimedia interactive, text interactive and operation are found in the system by analyzing the work flow and the design model.

According to the request of information transmission, the information of sharing platform can be divided into real-time information and non-real-time information. Among the above seven categories information, the first four categories are non-real-time information, the last three categories are real-time information, so they should be treated differently in the establishment process of information sharing platform.

The non- real- time information management system is also known as Collaboration Management Information System (CMIS), it is an eight-tuple, that is described below.

CMIS = {MI, NI, OI, UI, RI, CRI, CI, DI}

In CMIS, MI is a members information set that is an information database for all registered users, which includes user names, websites, tools by users and other personal information (such as experience, interest, etc.). NI is a node information set that is an information database of all registered sites, which includes site name, address, site information and site resources (such as multimedia equipment etc.). OI is an object information set that is responsible for the representation, organization and decomposition of cooperative tasks, and the arrangement and distribution of cooperative activities. UI is a tool information set that is a CSCW (computer supported cooperative work) tool information database registered which includes the name of tools and resource requirements of tool etc. RI is a resources information set in supporting environment that includes the use of various tools and use state of network resources etc. CRI is a cooperative rule information set that stores all kinds of available coordination strategy and control rules. CI is a cooperative information set that is an information database of all running cooperative design activities, which includes cooperation activities users set, collaborative tools set, cooperation object set, coordination resources set, workflow model and operation state etc..DI is document information set on the file information and the result of collaborative design, which includes results information produced in cooperation design activities, site material information of design objects and historical data kept in the system.

The information can be divided into static information and dynamic information. The information sharing platform can not only provide convenient, fast static information query tools, but also reflect dynamic information to all members instantaneously, in addition, the information sharing platform of the system must maintain the consistency of sharing information.

The real-time information includes interactive type information and operation type information. According to the basic characteristics of real-time information, it can be divided into text information and multimedia information.

The text information is a tetrad that consists of SI, CI, WI, and OI. In the above tuples, SI is system information that includes user request information, server address that collaboratively work, group number, port number for multimedia communication and bandwidth etc. CI is chat information that is information of ideas interchange among users who attend collaborative work by collaborative tools. WI is whiteboard information, OI is operating information that is the collaborative operation information on sharing object.

This kind of information can be transmitted in a reliable connection-oriented transmission mode (TCP), and then be processed by the corresponding agents.

The multimedia information is a two-tuple that consists of AI and VI, in which AI is the audio interactive information of users who participate in collaborative design, VI is video information of users.

Interactive multimedia information is characterized by a large amount of real-time data, but allowing the loss of a portion of information according to a certain quality of services [7]. This kind of information can be transmitted by user datagram mode (UDP protocol) that is connectionless and efficient, and then be processed by agents.

4.2. Basic Model of Information Sharing Platform

A specific design task can be divided into several subtasks, each subtask can often be completed through many collaborative activities, and each time many personnel may participate in the collaborative activities, at the same time in the process of completing a task, a human may not only be responsible for task management, but also participate in a specific design, who is playing a different role. Thus if the whole collaborative processes is carried out smoothly, the system must efficiently manage task, cooperation, personnel and documents and files produced in collaboration. So the basic model of the information sharing platform in collaborative design system is shown in Figure 4.

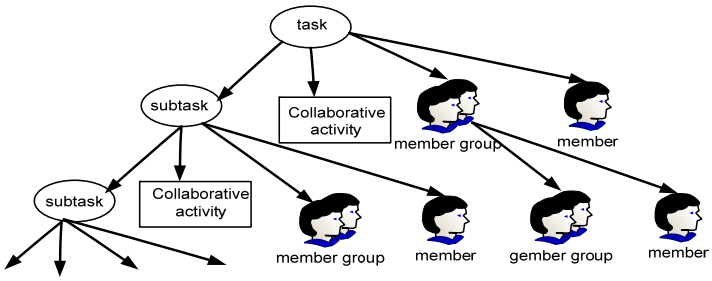

Figure 4. Basic Model of the Information Sharing Platform

5. Control Method of Multi-user Collaborative Access on Distributed Data Warehouse

To build a shared platform, the design of the underlying data warehouse is essential. According to the above shared platform, this part mainly studies the design of the distributed data warehouse. As the more complex drilling information, the drilling fluids are considered as the “blood" of drilling, so this article considers the drilling fluid as the theme of the distributed data warehouse.

During the model design of data warehouse, local data warehouse must firstly be designed, and then a global data warehouse is generated by reorganization, colligation and integration of information in local data warehouse. The data models of the local data warehouse and global data warehouse are described as follows.

5.1. Data Model of Local Data Warehouse

There are two kinds of multi-dimensional logical data model commonly used, that are the star mode (the star schema) and snow mode (the snowflake schema) [8-10]. Because the snow mode is extension of the star mode, this system adopts the snowflake schema to organize and analyze the drilling fluid information. From the perspective of the statistical analysis of the drilling fluid, it typically includes the themes of the consumption of drilling fluid, drilling fluid type, well type, well structures, oil field and the geological structures. Each theme corresponds to a snow mode. Taking the using of drilling fluid for example, the snow mode structure is shown in Figure 5.

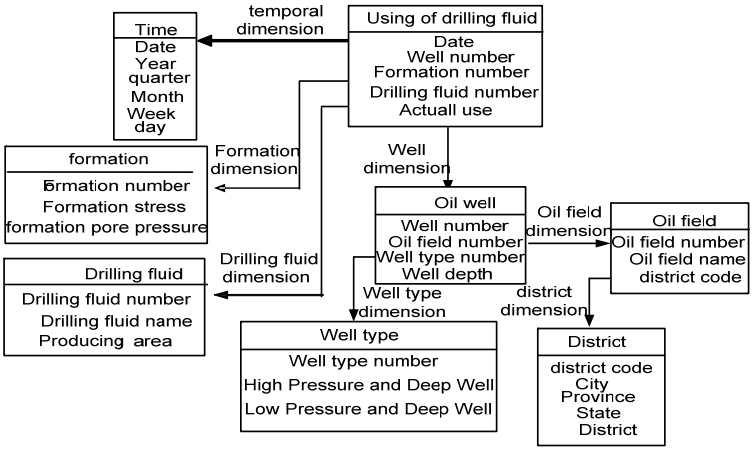

Figure 5. Snow Mode Structure of using of Drilling Fluid

5.2. Data Model of Global Data Warehouse

According to the oilfield production, the data of the global data warehouse comes from the low-grade integrated data of the local data warehouse, but the detail data still exists in each local data warehouse. So the fact table of the global data warehouse is the same as the local data warehouse, the difference is only of the higher grade of data granularity in the global, the data granularity of the system is shown in Figure 6.

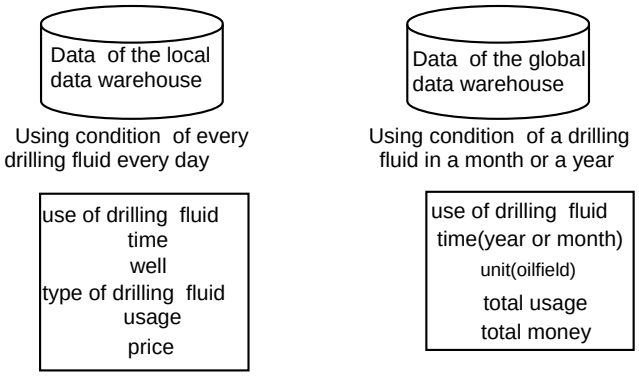

Figure 6. Data Granularity

The local data warehouse of oil fields stores meaningful data on the local decision-making, which is extracted from the local operational environment and provides support for local decision-making. The global data warehouse stores data from the local data warehouse, local operational data of itself and the headquarters. The distributed data warehouse is logically a whole, the data of which is actually distributed, when the data is accessed by the sharing platform, the distributives of which cannot be perceived by users. The global data warehouse just mildly synthesizes the data of the local data warehouse, and its detailed data corresponds to the sources of data through the source data table. The table is shown in Table 1.

Table 1. Source Data Table

field name	field code	type	length	unit	remarks
data source	SJLY	char	30		
server name of every units	FWQMC	char	10		
server IP	FWQIP	char	10		

Thus, the global data warehouse is very easy to acquire the summary information by synthesizing the details data of local data warehouse, for example, on the basis of the information about the using of drilling fluid in all the local data warehouse, the global data warehouse can analyze the using of drilling fluid, and see their specific information through the source data table.

5.3. Controlling Method of Multi-user Collaborative Access

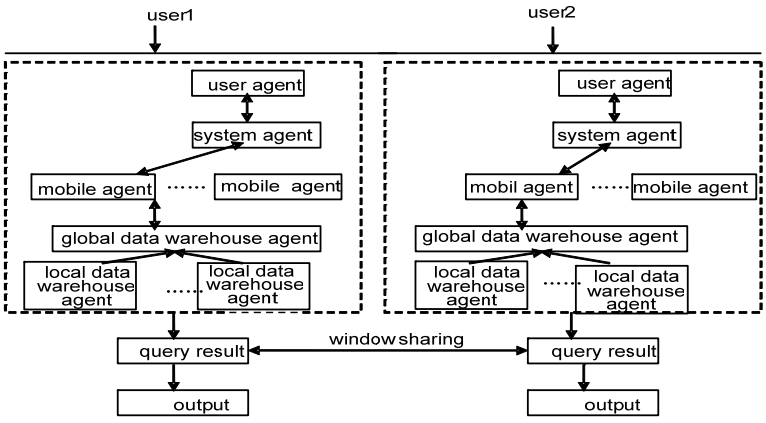

Figure 7. query structure of collaborative DW by multi-agent

When multi-users acquire drilling information by the platform, they need to synchronize when querying information. For example, when two users query, they must log in simultaneously, and then operate simultaneously. The concurrency control strategy must be adopted to control collaborative operation to realize WYSIWIS (What You See Is What I See). The structure diagram is shown in Figure 7.

The process of collaborative query operation can be realized by multi-agents. The query structure is made of user agent, system agent, overall data warehouse agent, local data warehouse and mobile agent, among which user agent answers receive user's query information, and then analyze it, finally transfer the query message to the system agent. The function of system agent, overall data warehouse agent, local data warehouse and mobile agent can be found in Figure 3. The output of collaborative query result can be realized by the synchronization screen sharing.

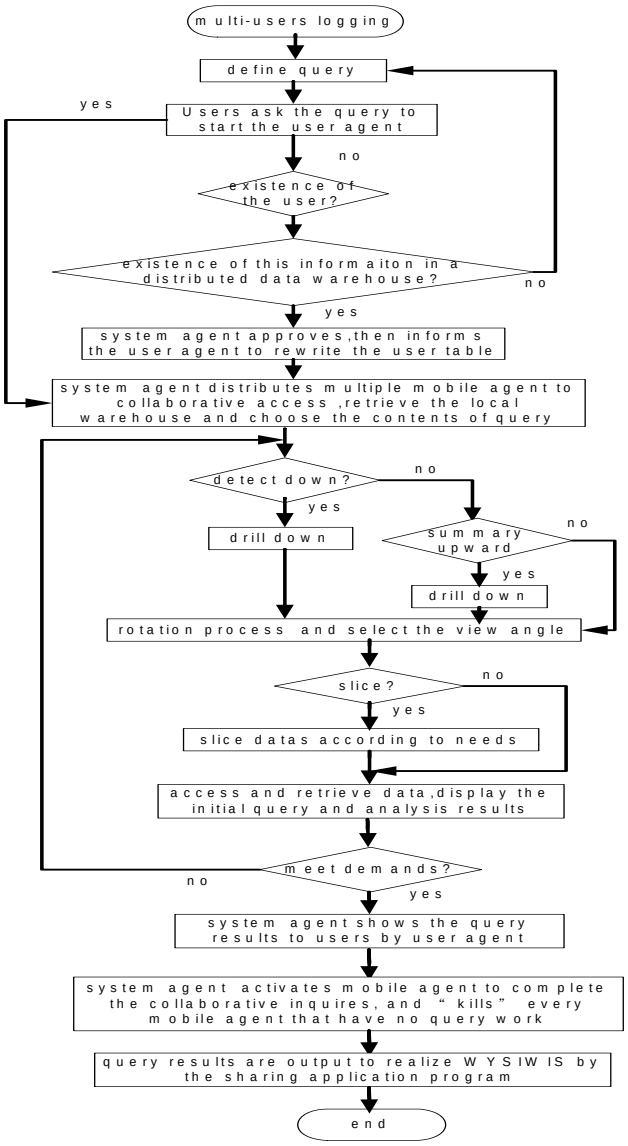

Figure 8. Collaborative Query Flow Chart

In the process of query, if the low detailed data are just needed, the data can be extracted from the global data warehouse to make analysis. If you need a relatively detailed data, firstly you find out where the data resides by the source data table in the global data warehouse, then drill (rollup, drill-down) data to reach the appropriate level, and rotate multi-

angles to observe data, and finally select the important data for the slice analysis. If the slice data is too much, it is divided into smaller data to analyze, the query results are shown to multiple users through the web server [11-13]. The query flow chart is shown in Figure 8.

The query on the distributed data warehouse is achieved through multi-agent to realize inquiry about different granularity information of local data warehouse and the global data warehouse. The query results allow multiple users to see real-time by means of the network to achieve the information sharing.

6. Conclusion

This paper combines multi-agent technology with computer network technology to build the model of cooperative working mechanism on the basis of the distributed data warehouse, and describes the information sharing platform, which can enable the latest technology in various fields apply in time at each production site and share all the information distributed in different places, and even the inventor can provide off-site real-time guidance. In this way, the technician, management experts or software systems in different locations are gathered together on the sharing platform to work together, maximizing the value of a variety of complex information of drilling. As a result, the level of information sharing can be improved, which can apply the latest drilling technology to dig the best quality wells safely and accurately improving the production rate of oil field.

References

[1] Wang Kui-sheng, et al. The Research on The Intelligent Application Syestem of A Battle for Distance Drilling Based on Network. *Acta Petrolei Sinica*. 2001; 22(2): 79-82.
[2] WH Inmon. Building the Data Warehouse. New York: John Wiley & Sons. 1996: 3-10.
[3] Pei Yunzhang, Shi Yuanchun, Xu Guangyou. Cooperative authoring system in distributed computing model. Journal of Tsinghua University (Science and Technology). 2000; 40(1): 76-79.
[4] QiMing Chen, Parvathi Hundi, Umeshwar Dayal, Meichun Hsu. Dynamic Agents. *International Journal of Cooperative Information Systems.* 1999; 8(2&3): 195~233.
[5] Xue Jianbin, Zhao Yanqin, Qin Lijing, Liang Bo. Research on Relay-Assisted Cellular System Coordination Mechanism Based on Energy-Efficient. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(9): 5502-5508.
[6] Xiao Bo, Zhang Dong, Zhu Hongwen. Concurrency Control Scheme for CSCW. *Journal of Shanghai Jiaotong University*. 1999; 33(1): 101~103.
[7] Leslie Monplaisir. An integrated CSCW architecture for integrated product/process design and development. *Robotics and Computer-Integrated Manufacturing.* 1999; (15): 145~153.
[8] Yang Yan-yan, Mei Hong. Data Warehousing Technology and Reusable Component Library System. *Computer Science*. 1999; 5: 56~60.
[9] Jiang Xu-dong, Zhou Li-zhu. A Multi-Table Join Algorithm for Data Warehouse Query Processing. *Journal of Software.* 2001; 12(2): 190~195.
[10] Bian Guangrong, Li Hongsheng, He Ninghui. Multi-sensor Data Processing and Fusing Based on Kalman Filtering. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(3): 1266-1277.
[11] Ted Gannon, Don Bragger. Data Warehousing with Intelligent Agents. 2001; 82~100.
[12] Elke A Rundensteiner, Andreas Koeller, Xin Zhang. Maintaining Data Warehouse over Changing Information Sources. *Communications Of The Acm*. 2000; 43(6): 57~62.
[13] Frank SC Tseng. Design of a multi-dimensional query expression for document warehouses. *Information Sciences*. 2005; 174(2): 55-79.