CDROM Databases : an Approach for Creation of Personalized Databases

Sonia Dube Premchand T A V Murthy

Abstract

Database by definition refers to organized collection of data stored at a Central place. Appreciable technical advancements have taken place in the field of Data organization, Metadata and indexing for faster and accurate access to the information. However for applications where connectivity to the centralized database cannot be ensured and has requirement of access to specific set of the database, then the approach proposed by us in this paper can be quite useful. This paper discusses approach for information extraction and tailoring of information as per user requirements – Database Personalization. This paper also addresses the issues of database security and Intellectual Property Rights involved in creation of CDROM Databases.

Keywords : Database Search, Database Security, Content Personalization, Query Generator

0. Introduction

Integration of techniques related to Database and Web has helped in making the database available to remote users, the power of search engines have further made the job of exploring the information more easy and user friendly using the centralized database. Use of techniques for Indexing, Metadata and Intelligent databases has helped in providing quick and accurate results. The current scenario calls for access to Centralized Databases for information extraction and further processing to make it compatible for use in decision-making. Centralized databases play an important role in reliable and accurate information access. Apart from the fact of popularity of use of Database there are situations where access to Centralized Database in not available, in such cases the information required from the database can be extracted as a small database and put on CDROM to provide mobility to the database. This brings the customized information required for decision maker to their desktops and has capability of mobility. This processes of extraction of a subset of Database as per user requirements is called as Database Personalization. Personalized databases on CDROM are designed to serve specific group of persons and provides mechanism for distribution of database in more efficient, compact and secure way. To distribute information in CDROM, customized information will be extracted from the databases and stored in the form of personalized databases. The Search interface distributed along with the database on the CDROM provides users to search the database as if they are accessing the Central database server.

1. Aims and Objectives

- Database personalization : Extraction of specialized information from the Central database for creation of customized databases.

- Data distribution : The CDROM databases are efficient way of data distribution in a secured way as this provides capability of encrypting the information, the Search and retrieval program ensures that only the information, which meets the search query, is decrypted and displayed to the user.

2. Implementation

The overall process of creation of Personalized Database as shown in Fig-1 can be divided into two parts

- System for Study of user requirements, automatic query generation and facility of Visualization and Verification of results for creation of Personalized databases
- System for Secured Distribution of Database

The Software System consist of

1) Client Search Program : this program is distributed along with the CDROM Database and consists of mechanism for Decryption, license verification and Search Interface to the database. This Graphical Search Interface (GSI) has interactive user views and provides users an easy and faster way to access information and provides means to navigate the results of the search query. Results are displayed to the user using crystal reports.

Search interface as shown in Fig - 2 has 3 main options –

a) User can browse database based on various field options provided

b) Quick search provides facility to search on selected field with the desired keyword. On selection of the field to search on a given keyword, window shows one browse list with available field values and number of hits presented in the database. User can either select field value to search from that browse list or can type desired keyword to search.

c) Advanced search option provides user a facility to combine 2 or more fields to get more enhanced and exact search results. On selection of any field to search on, Selection list is displayed on the window. User can select field value to search from the provided Search list or can type desired keyword to search on.

2) Intelligent Query Generator: This software system runs on the Centralized Database Server and is used to generate optimized query [1] for extraction of Personalized Database based on the User Requirements. This System also provides facility for Visualization of the results for verification of results against the requirements. The Graphical Visualization helps in refining the query and provides facility for selective removal of the elements from the results of the generated query. The IQG has built-in search engine, which apart from searching the database provides facility for generation of query based on the search results. The Search engine has features as described in Client Search Program.

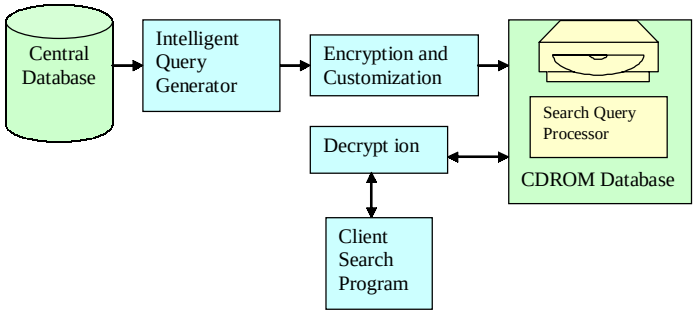

Fig. 1. Client Search Program

3) CDROM Database : A CDROM database for creating personalized database is implemented by using MS Access as database storage system. The personalized CDROM contains a self-extracting program, [2] which when executed installs the client part on user machine. The contents of the database are encrypted so as to protect the Intellectual Property Rights and avoid un-authorized use of the data, when distributed to the users.

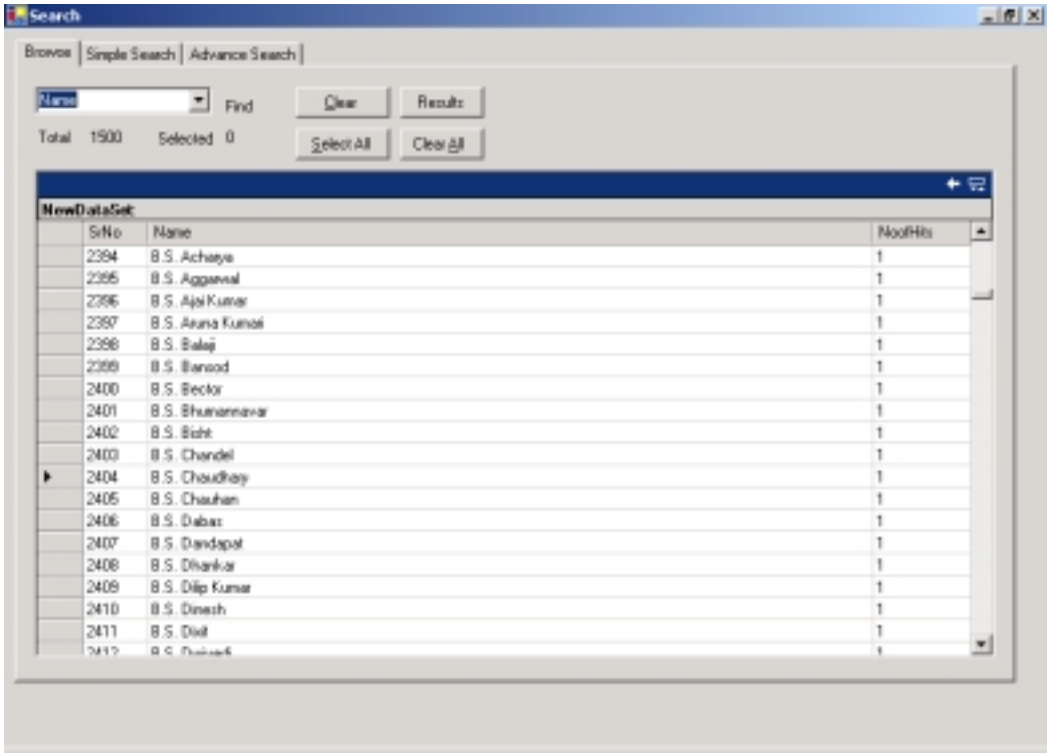

Fig. 2. Self extracting program

3. Database Personalization

Database Personalization refers to extraction of specific information from the database to make customized and personal databases. In other words, we can say that personalized databases [3] are subsets of the large databases. With personalized databases, user can access more accurate information with minimum possible time and without any need for access to costly resources. The steps involved in creation of Personalized databases are:

1) Study of Requirements
2) Use of Intelligent Query generator for extraction of data
3) Verification of Result set against the requirements
4) Creation of Encrypted Database
5) Generation of Client/Node Specific Search Program
6) Transfer of Encrypted Database and Client Search Program on CDROM

4. Database Security and Intellectual Property Rights

Database can contain information, which can be of open nature, where data can be freely distributed or can contain data, which can be distributed only with a limited rights i.e. in read-only mode. For maintaining security and protection of Intellectual Property Rights, database is secured with password and encryption techniques are applied on data to make it more secure. In our CDROM database system, MS Access database is provided with a customized menu bar, this menu bar provides limited options for the user to access and update the database. Stored procedures are used for data retrieval from the database. No other query option is open to users to make any changes in the database. User can't copy the database on his system as the Search Clients check the file-system type before accessing the database and allows to access the database only when being accessed from CDROM, this prevents un-authorized copies being created on the disk. Node locked license is used in the client program in order to ensure that the un-authorized copies cannot be used without taking license for specific computer.

5. Advantages of CDROM Databases

1. Database can be accessed without any network connectivity.
2. Specific part of the large database can be extracted and can be distributed on CDROM. For example in expert database [4] if we want to distribute publication details, research project or honor/awards then we can distribute only these information on CDROM.
3. Data of any particular organization can be extracted from large database and be distributed on CDROM.
4. Database can be distributed in a secure way.
5. Data can be easily distributed and there is no need for any database server to be installed for use of the data.
6. Helps in creating the personalized views of the database, which can be quite useful for decision makers and mobile users.

6. Conclusion and Future Scope

CDROM Databases is an effective tool for extension of the features of the Centralized database server to the desktops, where network connectivity to the Centralized server cannot be established. These databases when customized to generate Personalized Database can help in effective decision making as only focused data required by the user is made available to him, which avoids parsing of large volumes of data when dealing with the conventional search Engines. Database personalization is a better solution for fulfilling the increased demand of customized and up-to-date information. CDROM databases are one of the best ways to make these personalized databases available to specific group of persons without any network connectivity.

7. References

1. "A Roadmap to Personalized Information Systems by Cognitive Expansion of Queries "<http://www.cs.berkeley.edu/~balke/paper/enckompass02.pdf >
2. "Challenges for Personalized Online Content "<http://www.reddotsmartcontent.com/e_article000068023.cfm>
3. "A Navigation System for Personalized Databases: StarMap"< http://www.computer.org/proceedings/hicss/0493/04937/04937011abs.htm>
4. "Content Management of Web Enabled Experts Database in Science and Technology", by Sonia Dube, Premchand, Shivpal Singh Kushwah and Niyati Oza, published in CALIBER 2002.

About Authors

Mrs. Sonia Dube is Project Scientist at INFLIBNET Centre, Ahmedabad-380009, India and holds MCA. Earlier she was Programmer in Projects at INFLIBNET funded by NISSAT and WHO.
E-mail: sonia@inflibnet.ac.in

Mr. Prem Chand is Scientist- C and Group Leader- DDMG at INFLIBNET Centre, Ahmedabad-380009 and holds Master Degree in Political Science and Library Science and currently pursuing Ph D. He worked with Lal Bahadur Shastri National Academy of Administration, Mussoorie. He has published a good number of technical papers and attended conferences and workshops. He coordinated projects funded by WHO and NISSAT besides coordinating number of training programme at center and at various Universities.
E-mail: premchand@inflibnet.ac.in

Dr. T.A.V. Murthy is Director of INFLIBNET Centre, Ahmedabad-380009, India and holds B Sc, M L I Sc, M S L S (USA) and Ph.D. He is President and Fellow of SIS-India, Hon. Director of E.M.R.C., Gujarat University and Member Secretary of ADINET, Ahmedabad. He carries with him a rich experience and expertise of having worked in managerial level at a number of libraries in many prestigious institutions in India including National Library, IGNCA, IARI, Univ of Hyderabad, ASC, CIEFL etc and Catholic Univ and Case western Reserve Univ in USA. He has been associated with number of universities and has guided number of Ph.Ds and actively associated with the national and international professional associations, expert committees and has published good number of research papers. He visited several countries and organized several national and international conferences and programmes.
E-mail : tav@inflibnet.ac.in