_____

A Analysis of Different Type of Advance database System For Data Mining Based on Basic Factor

Kamlesh Kumar Pandey[1]
UTD - Department of Computer Science
MCRPSVV, Amarkantak Campus
Amarkantak, India
kamleshamk@gmail.com

Rajat Kumar Yadu[2]
Department of Computer Science
Mahant Laxmi Narayan Das College
Raipur, India
rajyadu28@gmail.com

Anshu Dwivedi[3]
UTD - Department of Computer Science
MCRPSVV, Amarkantak Campus
Amarkantak, India
Id_anshudubey62@gmail.com

Pradeep Kumar Shukla[4]
UTD - Department of Computer Science
MCRPSVV, Amarkantak Campus
Amarkantak, India
pradeepshukla2512@gmail.com

Abstract— Normal databases are unable to handling such as large range and large amount of data. Then we need database to support creating, storage, indexing and retrieval of large and wide variety of data for mining. This research paper presents different ways of data mining for advance data as multimedia, spatial, Time-series and heterogeneous data and management of database is given to help to creating, storing, indexing and retrieval. This includes advanced data structures and use of metadata to store advance data like multimedia, spatial, Time-series and heterogeneous database. This research paper claim the database management systems should be extended to arrange new type of data and enable to search based on their contents. Media, Geometry, Time, Calendar objects and all type objects are modeled as attributes of abstract data types. This paper will be describe multimedia, spatial, time series and heterogeneous database as a point of data mining methods, database management technique, data type and application for advance database.

Keywords- *multimedia database, spatial datbase, time- series database, heterogeneous database, legacy database, data type, data mining.*

_____*****_____

I. INTRODUCTION

Database systems are designed to manage large bodies of importance information. Management of data involves defining structures for storage of information and providing mechanism for the operation of information. In this research paper are presented to advance database as a type of relational database and object oriented database.

Object-relational databases have become the new standard for addressing the growing data management and analysis needs of non-traditional database applications such as Multimedia and Spatial Database Systems. Extensible database management systems main object are to make data management easier and more natural to users or applications such as web content management, location-based services, urban planning, utility and remote sense. Databases are traditionally used in business and administrative applications. These new object-relational and extensible databases can be used to solve the problems posed by multimedia database management, spatial database management, time series database management and Heterogeneous Legacy Data Base Management.

Application-oriented database systems like as spatial, temporal, multimedia, engineering databases, knowledge/information bases, and office news bases have flourish as related to the distribution, diversification and sharing of data. Heterogeneous database systems and World Wide Web (WWW) play a very important role in the information industry.

A database system is called a database management system (DBMS) consists of a group of interrelated data known as a database. We define to DBMS as Another word a set of software programs to manage and accessing on the data. Relational database is a collection of tables and each table name as a unique name. Each table of the relational database is consists of a set of attributes and its stores a large set of tuple. Each tuple in a relational table represents as object and object identified by a unique key. A type of semantic data model like entity-relationship (ER) data model is used of constructing for relational databases. An ER data model can be show the database as a set of entities and their relationships. In this research paper presented to advance database and all database is a type of relational database.

II. ADVANCED DATABASE AND THEIR FUNCTION

Basic database are relational database, transactional databases, data warehouses, object oriented database, flat files and World Wide Web. Advanced database systems are include

456

_____

object-relational databases and specific application-oriented databases, like as spatial database, time-series database, heterogeneous database, and multimedia databases.   In this section we discuss to all advance database based on definition, management, data mining, data type, 3-tier architecture, application and relational of other database.

A.   *Multimedia database:*

Multimedia database is a collation of multimedia data and this data are interrelated to each data   Means multimedia database are stores, manages and represented a large collection of multimedia data, like audio, video, image, graphics, speech, text and hypertext data, text markups. It attributed are related to continue data like picture content, voice-mail, video, World Wide Web, and speech-based recognize.

Multimedia Database is a generation of multimedia data and management of Multimedia Database is called Multimedia Database Management System (MMDBMS) it's function are storage, and management, distribution, receiving, consuming, editing, and sharing of multimedia data. MMDBMS are Handel and deal with huge amount of multimedia data and multimedia data types. A multimedia DBMS provides services for the multimedia database creation, storage, access, query, and management.

Multimedia database are hold to one or more multimedia data and Object oriented multimedia database, distributed multimedia database are use to may be one or more multimedia database.

Every multimedia database has a data type. Text, graphics, audio, video, animation are a basic data type of it database. If we implemented in Relational data base system or Data Base Management System like SQL, MS-ACCESS,MYSQL etc.

Multimedia data mining is the discovery of interesting patterns from multimedia databases. Multimedia data mining include classification and prediction analysis, mining associations, and focus on image mining, text mining, audio mining and video data mining. Issues of multimedia data mining are hold contain based retrieval and similarity search, and generalization and multi dimensional analysis

Multimedia databases are popular in many applications like as making a film and song, libraries, process control system, video conferring, game etc.

We tried to show 3-tier structure of multimedia database. In this structure have divided three part first part only view by user, second part and three part is view by database administer. Physical data level have a indexing, compression, feature extraction and media data is a main interface component and interface of conceptual data view have temporal model, data model, data access component
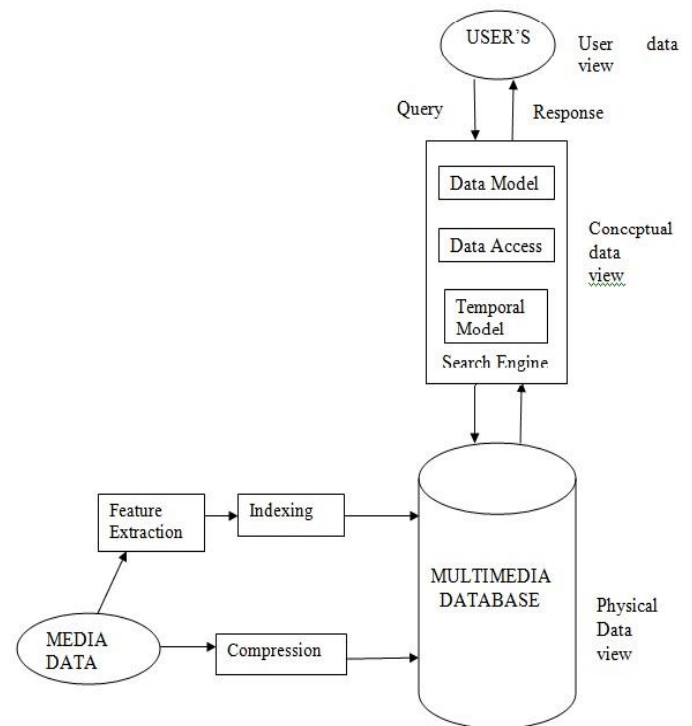

Fig:-1 Three-tier Architecture of Multimedia Database

B.   *Spatial database: -*

This database is consisting of spatial-related information. Spatial data are connected to a place in the Earth. in general it data and attribute are a geographic related. Spatial data may be represented in raster image format, consisting of *n*-dimensional pixel maps or bit maps. we represented   to this type of data in computer screen on diagram such as points, lines and polygons by 2D or 3D tech.

Spatial Database Management System (SDBMS) play a prominent role in the management of query of spatial data. It is a system that can work with an underlying database management system e.g. Geographic (map) databases, very large-scale integration (VLSI) or computed-aided design databases and satellite image databases. SDBMS support multiple spatial data model, spatial indexing, spatial operations and domain specific rules for query optimization. Object oriented Spatial database, distributed Spatial database are use to may be one or more multimedia database.

Special data types needed to model geometry and to suitably represent geometric data in database systems. spatial data types or spatial algebra can be capture the fundamental abstraction for point, line, region partitions (maps) and graphs (networks) are described above together with relationships between them and operations for composition (example form the intersection of many regions).

Spatial data mining is the discovery of interesting patterns from large geospatial databases Means Spatial data mining

457

refers to the extraction of knowledge or interesting information , spatial relationships or othe pattern not explicitly stored in spatial databases Spatial data cubes that contain spatial dimensions and measures can be construct. Spatial Online Analysis Processing (OLAP) can be implement to facilitate multidimensional spatial data analysis. Spatial data mining includes mining spatial association and co-location patterns, clustering, classification, and spatial trend and outlier analysis. Many approaches are use to indexing on spatial data is R-tree approach, Grid Files, Quad trees, and techniques based on "space-filling" curves.

Geographic databases have numerous applications, ranging from forest and ecosystem planning to provide public service information regarding the location of telephone and electronic cables, pipe, and sewage systems. Geographic databases are commonly used in vehicle navigation and dispatching systems.

We tried to show 3-tier structure of spatial database. In this structure have divided three part first part only view by user, second part and three part is view by database administer. Physical data level have a digitizer, load, geocode and spatial data is a main interface component and interface of conceptual data view have a topology, metric geometry, temporal, data model, data access component .
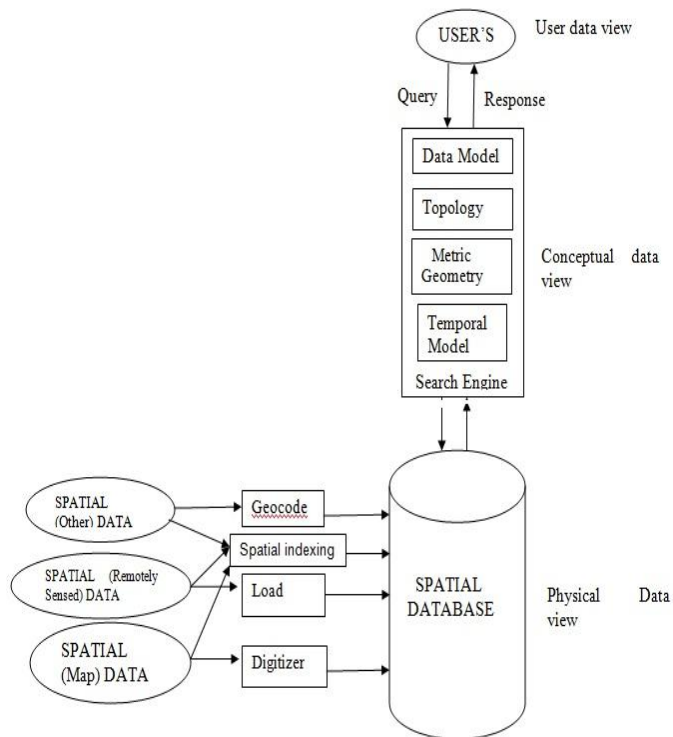

Fig:-2 Three-tier Architecture of Spatial Database

C. Time Series Database:

A time-series database stores sequences of values or actions are obtain over repeated measurements of time like a hourly, daily, weekly means A time series database is a collection of time series data, arrays of numbers indexed by time and this data are interrelated to each one. The TSDB allows to user to create, enumerate, update and destroy various time series and arrange them in some way. These series may be arranged hierarchically and optionally have companion metadata available with them. The implementation of a time series database can be deployed in a conventional SQL-based relational database. That provides to the database for binary large objects (BLOBs) and user-defined functions.

A time-series database is some time called a sequence database. A sequence database is consists of sequences of ordered events with or without notions of time. Temporal database are subset of   time-series database and it's stores relational data that include time-related attributes. These attributes involve a lot of several timestamps, each having different semantics. Some advance time-series database are Object oriented time series database, distributed time series database and relational time series database.
Systems that can be manage to time series information that can be called a Time Series Data Base Management System (TSDBMS). TSDBMS is provide to system define function for calculation of time related data and Conversion of between date, time, date time value. A Time Series DBMS provides services for the time series database creation, storage, access, query, and management.

Date, Time, Month, hours, minute, second, Calendar is basic data type of time series database. Time series database are use to time series algebra for calculation of time variable and time constrains.

Time-series mining is finding on inserting pattern/ sequence information for time related. It use to indexing, clustering, regression analysis, Anomaly Detection, Segmentation, Summarization technique. After mining on time series data are represented to data adaptive and non data adaptive method. in data adaptive method Symbolic, tree, curve, single form are most popular technique in representation of data. Time-series mining use multi dimensional data.

Time-series databases are popular in many applications like as stock market examination, economic perdition and sales forecasting, budget analysis, utility study, inventory study, give way projections, working load projections, process control and quality maintained , observation of natural phenomena scientific and engineering experiments, and medical treatments.

Applications include stock market analysis, economic and sales forecasting, cardiogram analysis, and the observation of weather phenomena.

458

We tried to show 3-tier structure of time-series database. In this structure have divided three part first part only view by user, second part and three part is view by database administer. Physical data level have a script, TDS, indexing is a main interface component and interface of conceptual data view have a data Getaway API, RAM, blob maker, data model and data access.

Fig:-3 Three-tier Architecture of Time-Series Database

D. Heterogeneous Databases:

A heterogeneous database consists of a set of interrelated autonomous module databases. The modules communicate in order to swap information and answer query. Objects in one module database may be differing greatly from objects in other module databases making it difficult to assimilate their semantics into the overall heterogeneous database. In short heterogeneous database store different type of data or module in one place and this data are interrelated to each one.

A *legacy database* is a special type of heterogeneous databases and that combine different kinds of data systems and different type of behavioral data structure such as relational database, object-oriented databases, hierarchical databases, object relational oriented database, multimedia databases, time series database etc. and file systems. The heterogeneous databases in a legacy database may be connected by computer networks.

In summary heterogeneous database are consists of interconnected autonomous database and legacy database is a

group of heterogeneous database. All results are depends on all of the stored data in both database.

We tried to show 3-tier structure of heterogeneous and legacy database. In this structure have divided three part first part only view by user, second part and three part is view by database administer. Physical data level have a TDS, indexing, compression and different type of data is a main interface component and interface of conceptual data view have a data Getaway API, RAM, data model, data access, topology, temporal model component.
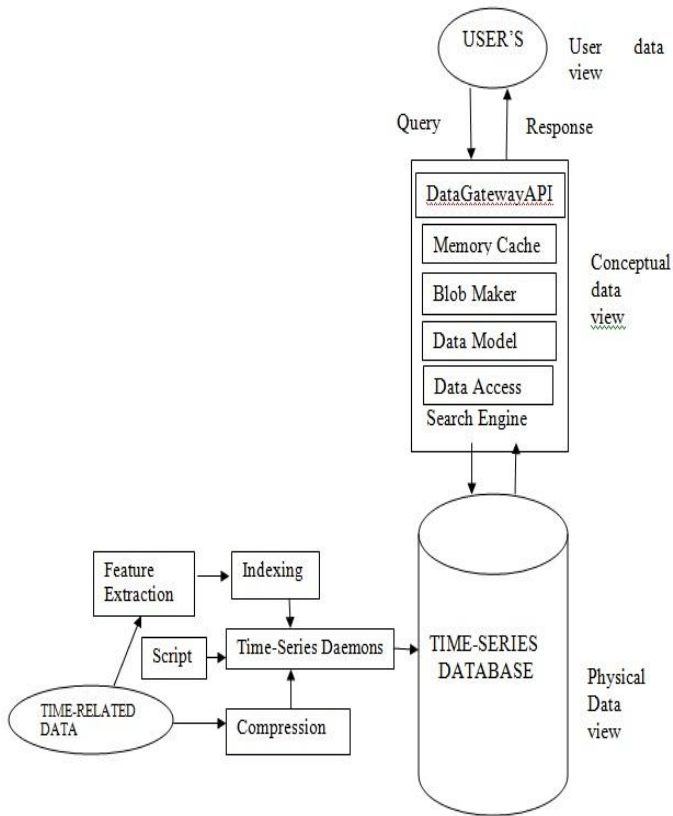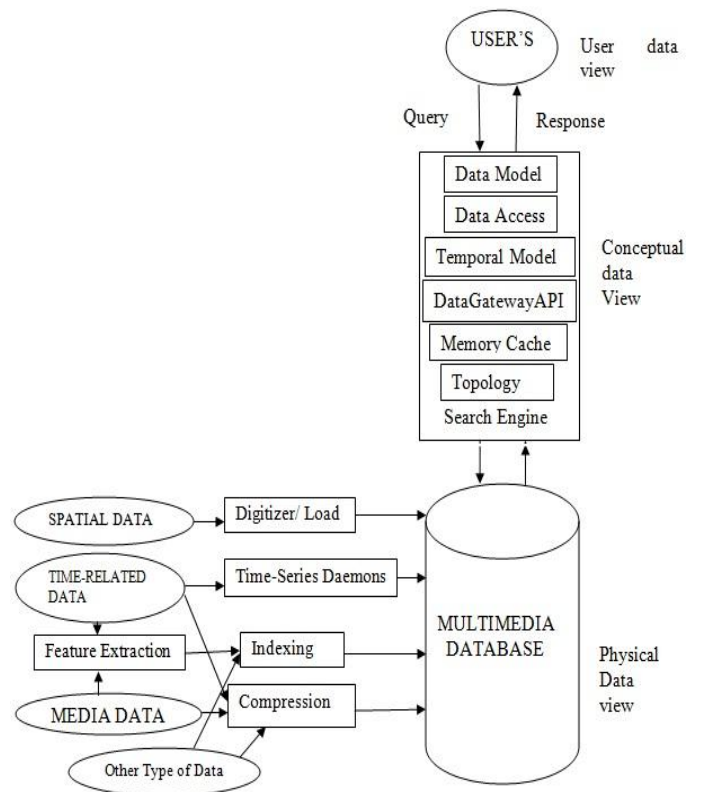

Fig:-4 Three-tier Architecture of Heterogeneous and Legacy Database

.

Systems that can be manage to heterogeneous database and legacy database that will be called a Heterogeneous Legacy Data Base Management System (HLDBMS). HLDBMS is providing a function for communication to different type of database and combination of different type of data. A HLDBMS provides services for the heterogeneous database and legacy database creation, storage, access, communicating, query, and management.

Every Heterogeneous and legacy database have a different kind of data type. Like a Text, graphics, audio, video, animation, time, date, integer, point, line are a basic data type of it database.

Local area network and wide area networks connect many sources of data, distributed database and heterogeneous databases. The discoveries of knowledge from different data sources are structured information or unstructured information and both type information are challenge to data mining in legacy database and heterogeneous database. In Heterogeneous and legacy data mining is use to Classification technique like a decision tree use for taking a decision, neural network use for predicted a class, Bayesian classification use for mathematical classification..

III. CONCLUSION

This paper discusses advance database system based on data mining technique, data type, application and relational of other database and 3-tier Architecture. This research paper cover to various point of advance database system. In first we focus on what is a database second we focus on Data Base Management System, third we focus on relational of other special advance database like distributed database, object oriented database. Fourth we focus on different data type of data base. Fifth we focus on data mining technique of advance database. Six we focus on application of all databases. Basic aim of this research paper is mining on advance database system.

REFERENCES

[1] Jiawei Han University of Illinois at Urbana-Champaign Micheline Kamber - Data Mining: Concepts and Techniques Second Edition, ISBN 13: 978-1-55860-901-3 and ISBN 10: 1-55860-901-6

[2] Agarwal,R and Psaila G, Active Data Mining. In Proceedings on Knowledge Discovery and Data Mining (KDD-95), 1995, 3-8 Menl

[3] Shunge Li, Xiangning Liu, Bharat Bhargava "A Conceptual Model for Multimedia Database Systems" Computer Sciences Department Purdue University West Lafayette, IN 47907

[4] Madiha Waris, Farooque Azam, Abdul Wahab Muzaffar "A Survey of Issues in Multimedia Databases" International Journal of Computer Applications (0975 – 8887) Volume 46– No.7, May 2012

[5] B. Thuraisingham. Security and privacy for multimedia database management systems, Multimedia Tools and Applications, 33(1): 13-29, April 2007.

[6] Multimedia Database Applications: Issues and Concerns for Classroom Teaching, The International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.1, February 2011

[7] R. Ramakrishnan and J. Gehrke , Database management Systems 3ed, Chapter 28- Spatial Data Management

[8] Ralf Hartmut Güting "An Introduction to Spatial Database Systems" Praktische Informatik IV, FernUniversität Hagen, D-58084 Hagen, Germany, E-mail-gueting@fernuni-hagen.de

[9] Christopher B. Jones "DATABASE ARCHITECTURE FOR MULTI-SCALE GIS" Department of Computer Studies, The Polytechnic Of Wales Pontypridd Mid Glamorgan, CF37 1DL, UK

[10] K. Li, R. Laurini: The spatial locality and a spatial indexing method by dynamic clustering. Proc. 2nd Int. Symp. on Large Spatial Database, Lecture Notes in Computer Science #525, Springer-Verlag, 207-224(1991).

[11] Michail Vlachos, Gautam Das "MINING TIME SERIES DATA" University of California, Riverside and IBM T.J. Watson Research Center

[12] Chapter 2-"Working with Time Series Data", SAS/ETS User's Guide, Version 8 Copyright © 1999 by SAS Institute Inc., Cary, NC, USA. ISBN 1–58025–489–6

[13] J. Aßfalg, H.-P. Kriegel, P. Kr¨oger, P. Kunath, A. Pryakhin, and M. Renz. Threshold similarity queries in large time series databases. In Proceedings of the 22nd International Conference on Data Engineering, ICDE 2006, 3-8 April 2006, Atlanta, GA, USA, page 149, 2006.

[14] Wook-Shin Han "Ranked Subsequence Matching in Time-Series Databases" Department of Computer Engineering, Kyungpook National University, Republic of Korea, E-mail- wshan@knu.ac.kr

[15] Keogh, E. "A Decade of Progress in Indexing and Mining Large Time Series Databases," In VLDB, Tutorial, 2006.

[16] Rafiei, D. et al., "Querying Time Series Data Based on Similarity," IEEE TKDE, Vol. 12, No. 5, 2000.

[17] Ted Dunning and Ellen Friedman "Time Series Databases New Ways to Store and Access Data" Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. ISBN: 978-1-491-91702-2

[18] Batini, C., Lenzirini, M., & Navathe, S. B. "A comparative analysis of methodologies for database schema integration" ACM Computing Suroey. 18,4 (December 19861, pp. 323 - 364.

[19] Breitbart, Y., Olson, P. L., & Thompson, G. R. Database integration in a distributed heterogeneous database system. In Proceedings of tk 2nd International Conference on Data Engineering, (Los Angeles, CA, February, 1986), pp. 301-310.

[20] William W. Cohen "Integration of Heterogeneous Databases Without Common Domains Using Queries Based on Textual Similar" AT&T Labs-Research 180 Park Avenue, Florham Park NJ 07932 E-mail wcohen@research.att.com