ISSN XXXX XXXX © 2018 IJESC

Research Article Volume 8 Issue No.2

Crime Prevention with Data Warehouse using OLAP through Business Intelligence

Hanamant B. Sale[1], Deepali Patil[2], Sonal Thube[3]
Assistant Professor[1], BE Student[2, 3]
Department of Information Technology
Bharati Vidyapeeth College of Engineering, Navi Mumbai, India

Abstract:
There Abstract—in the recent past there has been an expeditious rise in the crime rate. In modern criminal law the term "crime" doesn't have any uninvolved and universally accepted definition. Crime is regarded as a dangerous act not only to individuals but also to society, community or state. It is considered as an act of breaking human laws. Various factors leading for crime to takes place are unemployment, depression, social and mental conditions, family condition, politics, etc. It is not easy to recognize the patterns of crime by the authorities. Crime Detection has become a mayhem task. In their role of catching criminals, the cops are required to remain convincingly ahead in the external race between law breakers and law enforcers. There is need for user interactive interfaces based on current technologies in order to provide them the much required edge and to fulfill the new emanating responsibilities of the police authorities. So, we require new tools in order to utilize the complete capacity of the temporal as well as spatial dimensions of a data warehouse. OLAP holds a certain capacity to assist spatio-temporal scrutiny has been shown. The paper exhibits multidimensional database design for prevention of crime containing agglomerated data from different crime registers and many other different sources for e.g. unemployment, properties, health, population, facilities etc. Introduction of spatial data in multidimensional model elevates major issues from implementation as well as theoretical point of view. There are some open issues in SOLAP (Spatial OLAP). Thus we propound GIS-OLAP integrated solution which supports geographical dimensions and measures, and provides interactive pivot tables and diagrams displays in order to efficaciously support decision makers. Thus assisting us understand convoluted factors and impacts of changes, and finally also for strategic crime prevention.

Keywords: URL, Pattern matching algorithm, Node.js, Angular.js, Web scraping and Web Crawler.

I. INTRODUCTION

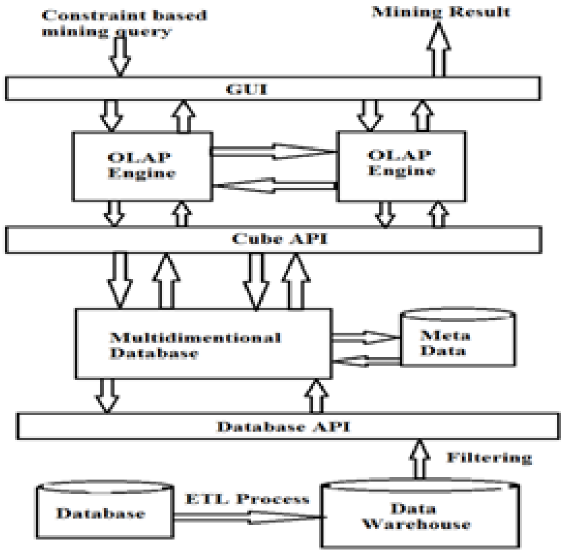
Increasingly remarkable volumes of data are collected by the organizations. They form the foundation for data analysis processes once stored in data warehouses and escort the organization's tactical decisions. Nevertheless, data are not always used to their complete capacity and part of their richness is sometimes left out, that is, their spatial component. "Hidden in most data is a geographical component that can be tied to a place: an address, postal code, global positioning system location, (…) region or country" [ESRI 2000]. Emphatically, it has been reckoned that about 80% of all data stored in corporate databases has a spatial component [Franklin 1992] that can be distinguished by shape, position, orientation or size. Time can also be regarded as a pith component of data: "...without a record of the time of the observation, the useful content of the information may be minimal" [Sinton 1978]. The significance of the temporal dimension for decision making is pivotal and it is reflected in the definition of data warehouse: "...a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management's decision making process" [Inmon 1996]. But time is only one of the two logical components, in addition to the explanatory components containing the features of the data, needed to take complete advantage of a data warehouse; As mentioned previously, space is the other [Gonzales 1999]. A Data Warehouse (DW) is a concentrated repository of data obtained from external sources of data and assembled following a multidimensional model (Inmon, 1996) in order to be scrutinize by On-Line Analytical Processing (OLAP) applications. OLAP tools provide the ability to interactively explore multidimensional data presenting detailed and aggregated data. The outcome of the scrutiny is the foundation of strategic business decisions. Spatial dimensions, just like temporal ones, should then be considered standard for any data warehouse implementation. In fact, to date, the spatial dimension has been extensively integrated in data warehouses, but usually in a nominal, non-cartographic manner (i.e. using solely place names). Seldomly, one may use coordinates (as is, or reorganized such as with HHCodes) inside a data warehouse for displaying map and drilling, but it is rarely used to its complete capacity for data exploration. To gain better advantage of the spatial and temporal dimensions in decision making, suitable tools must be used. Geographic Information Systems (GIS) are the obvious potential candidates for such tasks. While having some spatio-temporal analytical capabilities, it identified that existing GISs per se are insufficient for decision-support applications when it is used alone and that alternative solutions must be used [Bédard et al. 2001]. Amid the possible solutions, the coupling of spatial and non-spatial technologies, GIS and OnLine Analytical Processing (OLAP) for instance, may be a fascinating choice. OLAP is a class of decision-support tools frequently used to impart access in an systematic and instinctive way to a data warehouse. Data warehouse and multidimensional modeling may be applied for different purposes not only for business strategies. The public sector generally frequently coalesces data from regularly provided data sources and needs effective tools for repeated spatio-temporal analysis of large volume of incremental data. One of such application can be the database for crime prevention. Data for crime prevention database comes from various data sources different in purpose, structure, formats, data quality, and maintenance, legal and ethical aspects (license policy, sensitivity). The data value is hidden in its organization, accessibility and usability.

II. LITERATURE REVIEW

The GeoMondrian Project (2011) aims to develop an open-source implementation of a SOLAP analysis server. Currently, it provides a spatially enabled version of the Mondrian OLAP server (Pentaho Analysis Services: Mondrian Project, 2011). However, in our view the GeoMondrian system lacks a clear underlying spatial data model to help integrate spatial objects into OLAP systems. Instead, it is essentially built ad-hoc, by using a combination of the Java Topology Suite (2011) (which provides spatial operations according to OGC standards) and Mondrian (which provides the OLAP operations on thematic attributes) with PostGIS (which provides the spatial data types). These together create a functional spatial data analysis toolkit supporting the integration of spatial data and operations in an OLAP server. The architecture of Business Intelligence (BI) containing many technologies such as Data Warehousing, Data Mining, data integration, OLAP etc. have been proposed in many fields. It is used to establish decision support that prioritizes efficient querying of huge amounts of historical, clean and consolidated data [10]. However, a spatial perspective is needed for strategic analysis: the integration of spatial component in BI is called Geospatial Business Intelligence (Geo-BI) in order to reveal relationships based on the geographical positions associated with the information. The main aim of this section is to review the use of BI technology (such as data warehouse or/and OLAP) with GIS in two ways, either by integrating spatial and warehousing information in a single framework or treating tools separately. There are previous studies in this area. Some approaches and research tools form a good basis for a multidimensional spatio-temporal decision support tool [11]. The Geomatics Center of Laval University in Canada provides Spatial OLAP (SOLAP) as a tool for support rapid and easy spatio-temporal analysis, exploration of multidimensional data and offers a set of visualization techniques such as maps, tables and diagrams [12]. The Informatics Center of the Federal Pernambuco University in Brazil has proposed GOLAPA (Geographic Online Analytical Processing Architecture) providing users with an abstraction of the complexity involved in querying both analytic and geographic data for decision support [13,14]. Among other tools that combine BI with GIS, we cite Cube View and Polaris proposed by Voss [15] to facilitate the observation of spatial patterns and temporal trends in large volumes of data. Furthermore, in [16], the authors have developed a robust, easily accessible and user-friendly Geo-BI solution for the local governance of infrastructure services. Moreover, in [17], the authors proposes an extensible GIS-OLAP integrated solution supporting geographical dimensions and measures, based on a formal multidimensional model named GeoCube. More than that, development of OLAP- GIS tool [18] was proposed by the University of Pittsburgh for numerical-spatial problem solving in community health assessment analysis. GIS, and the added value in the combination of both tools propounded in the university study [18]. With regard to the research discussed above, we shall now conclude that the approaches proposed are limited in the use of data warehouse models, as these models are composed solely of quantitative data, while the necessary data for decision-making in complex problems are often quantitative and qualitative in type. This leads to the difficulty in evaluating the quality of geospatial information. Hence, it is normal to regard several types of heterogeneous data.

III. SYSTEM ARCHITECTURE

*The front end system gives graphical user interface (GUI).The data is gathered from various sources into the database which further passes through ETL process and is stored in data warehouse. Whenever query is fired a call is made to data warehouse by database API for the requested data which is then filtered and converted into multidimensional data. Then call to. Multidimensional data is made by cube API and session is generated. The required data is fetched and is processed by OLAP engine. It processes the data by drill down and roll up approach and the requested data is acquired in the form of table, graph or pie chart



IV. PROPOSED SYSTEM

The authority/public will first register into the application. After registration there will be 2 logins. One for the authority people and other for common public. If any authorized person logs in then he/she can add data, view data, query data. And if common people logs in they can just view and query data and cannot modify the data. While adding data the authority have to select the table based on various factors or crime type. Then the entry of criminals information is done. This added data is stored in the database. Whenever the authority or the public wants to view data they click on view data and a map is displayed on the screen. The person the selects the area from the map and grids of that area is displayed. Eg : the person selects kharghar area from the map and further selects sector 12 from grid, the data of that area is displayed in the form of table, graphs, pie chart. Another approach is querying data. The person will select query data and write the query according to their need and the data will be displayed in the form of table, graph, pie chart. Thus in this way crime data can be analyzed by taking into account previous records which will help the authority to take necessary actions in future

V. CONCLUSION

The Business Intelligence provides appropriate effective and user friendly applications not only for business, commercial activities, but also for public sector. Main advantages can be seen in ETL processes for repeated data integration from various sources and SOLAP analytical tools for simple spatio-temporal analysis of data in OLAP cubes. The design of multidimensional database for crime prevention has been described and appropriate rules were established. The core of the database represents various fact tables. Each fact table is documented including basic operations for implementation as well as main dimension tables.

VI. REFERENCES

[1]. Bédard, Y., T. Merrett and J. Han. 200 1. Fundamentals of spatial data warehousing for geographic knowledge discovery. In: Geographic Data Mining and Knowledge Discovery. Fdited by H. Miller and J. Han, Taylor & Francis.

[2]. Bédard, Y, R. Devillers, E. Bernier and B. Moulin. 2000. Automatic Generalization or Multiple Representations: Towards a combination of both approaches.

[3]. Codd, E. F., S. B. Codd and C. T. Salley. 1993. Providing OLAP (On-Line Analytical Processing) to User- Analysts: An IT Mandate. Hyperion white papers, 20 p.

[4]. Bimonte, S., Tchounikine, A., Miquel, M.: GeoCube, a Multidimensional Model and Navigation Operators Handling Complex Measures: a Spatial OLAP Application. In Proceedings of the 4th Biennial International Conference Advances in Information Systems, Izmir, Turkey, October 2006. LCNS 4243, Springer-Verlag Berlin Heidelberg: 100-109, 2006.

[5]. Bimonte, S., Di Martino, S., Ferrucci, F., Tchounikine, A.: GeOlaPivot Table: a Visualization Paradigm for SOLAP Solutions. In Proceedings of the VLC Workshop of 12th International Conference on Distributed Multimedia Systems, Grand Canyon, USA: 181-186, 2006.

[6]. Bimonte S.,Wehrle, P., Tchounikine, A., Miquel, M. GeWOlap: a Web Based Spatial OLAP Proposal, Second International Workshop on Semantic-based Geographical Information Systems

[7]. Kimball, R., Ross, M. (2002). Data Warehouse Toolkit (2nd ed.). New York: John Willey and Sons.

[8]. Song, I. (2009). Data Warehouse. Liu L., Özsu M. T. (Ed.), Encyclopedia of Database Systems (1, p. 2731- 2735). Berlin: Springer.

[9]. Abelló, A., Romero, O. (2009). On-Line Analytical Processing. Liu L., Özsu M. T. (Ed.), Encyclopedia of Database Systems (20, p. 2731-2735). Berlin: Springer.

[10]. Ferraz, V.R.T., Santos, M.T.P. (2010). GlobeOLAP: Improving the geospatial realism in multidimensional analysis environment. Proceedings of the 12th International Conference on Enterprise Information Systems, 5, HCI, Funchal, Madeira, Portugal, June 8 - 12.

[11]. Shekhar, S., Lu, C.T., Tan, X., Chawla, S., Vatsavai,R. (2001). Map cube: A visualization tool for spatial data warehouse. As Chapter of Geographic Data Mining and Knowledge Discovery. Harvey J. Miller and Jiawei Han (eds.), Taylor and Francis.

[12]. Rivest, S., Bédard, Y., Marchand, P. (1997). Towards better support for spatial decision-making: Defining the characteristics of Spatial On-Line Analytical Processing (SOLAP). Geomatica, The Journal of the Canadian Institute of Geomatics, 55 (4) 539-555

[13]. Fidalgo, R., Silva, J., Times, V., Sousa, F., Barros, R. (2003). GMLA: A XML Schema for Integration and Exchange of Multidimensional Geographical Data. GEOINFO In: Proceedings of the 5th Annual Brazilian Symposium on Geo Informatics.

[14]. Fidalgo R. N., Times, V. C., Silva, J., Souza F.F., Salgado, A.C.(2004). Providing multidimensional and geographical integration based on a gdw and metamodels. Brazilian Symposium on Databases (SBBD).

[15]. Voss, A., Hernandez, V., Voss, H., Scheider, S. (2004). Interactive Visual Exploration of Multidimensional Data: Requirements for CommonGIS with OLAP. Database and Expert Systems Applications, 15th International Workshop on (DEXA'04), p. 883-887, California, IEEE.

[16]. Wickramasuriya, R., Jun Ma, Berryman, M., Perez, P.(2013). Using geospatial business intelligence to support regional infrastructure governance. Knowledge-Based Systems, 53. 80–89.

[17]. Bimonte, S., Tchounikine, A., Miquel, M. (2007). Spatial OLAP: Open Issues and a Web Based Prototype. 10th AGILE International Conference on Geographic Information Science Aalborg University, Denmark

[18]. Scotch, M. L. (2006). An OLAP-GIS system for numerical-spatial problem solving in community health assessment analysis. University of Pittsburgh.