

AN APPROACH TO TARGETED MARKETING USING SPATIAL DATA MINING AT CAMPBELL SCIENTIFIC, INC

Pamela Jean Pratt, Utah State University, Pam.Pratt@sdl.usu.edu
Dr. David H. Olsen, Utah State University, dolsen@b202.usu.edu

ABSTRACT

The discussion presented here is based on the concepts of data warehousing, data mining, and Geographic Information Systems (GIS) for a local manufacturing company with a worldwide customer base.

Keywords: Data Mining, Geographic Information Systems, Relational Database, Marketing, Targeted Marketing, Business Information Systems

INTRODUCTION

The concept of this article focuses on merging two technical information systems (IS) fields: data mining and geographic information systems (GIS). Initially, a literature review is presented followed by a brief background of Campbell Scientific, Inc., a Logan, Utah based company. Next, the framework for spatial data mining is presented followed by a proposal for improvement. Data mining of relational databases is explored and combined with the functionality of geographic information systems (GIS) with an emphasis on business GIS processes.

The objective of data mining is to discover some new pattern of information from the information recorded in a database (7). Data mining adds value to business information by extracting invisible but meaningful patterns and rules in corporate data (4). These patterns enable organizations to better understand and predict customer behavior to develop a targeted marketing schema.

A GIS is an analytical tool used to integrate tabular information and graphical information. A geographic component can be identified in business data by an address or zip code and mapped accordingly.

Campbell Scientific, Inc., (CSI), a major manufacturer of scientific weather monitoring equipment and data loggers, has implemented an Oracle Enterprise Resource Plan (ERP) to integrate its business processes. CSI is used throughout this paper as an example of how the spatial data mining strategy can be applied in a business environment. CSI has a strong understanding of their business and its science, their customers, and their partners. By implementing the processes proposed here, CSI will advance its marketing strategy and seize new business opportunities.

LITERATURE REVIEW

A data warehouse is a large database where data from many different sources (e.g., order entry, production, financials) are gathered together and organized into a consistent format (13). Approximately 95 percent of Fortune 1000 companies are implementing data warehouses to organize the vast amount of information they require to operate and plan their business and marketing strategies (17). Nearly 80% of data stored in corporate databases has a geographical component (8). The introduction of a spatial component is the most powerful aspect in a data-warehousing environment (8).

On-line Analytical Processing

On-line analytical processing (OLAP) is a data analysis technology that presents a multidimensional, logical view of data to the business analyst (10). OLAP tools can sort, forecast, track trends, and perform other complex analyses on data contained in a data warehouse. OLAP tools also allow users move from one query to another and obtain results quickly and easily (10). An OLAP query would be stated as 'How many weather stations with CR-10 data loggers did we sell last month in each region, Europe and Asia/Pacific, compared to the same month last year?' (10). A common example for OLAP is sales performance measured by time, geography and customer profile (12). Non-OLAP queries extract basic, non-dimensional information from the database. A non-OLAP query, for example, would be 'How many weather stations did we sell last month?'

Geographic Information Systems

A GIS is an organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. GIS consists of data input, data management, data manipulation and analysis, and data output. Data output can be in the form of maps, tables, or reports.

Applications in which GIS can assist are corporate site selection and analysis, demographic analysis, product pricing analysis, data access and reporting, geographic locator, and geocoding customer and facility locations. Implementation of these applications can assist with the market implementation strategy design, custom application development, and database evaluation, design, and construction (5).

Many corporations have implemented geographic information systems to direct their marketing programs at a particular audience. Bari Horton, director of direct marketing at Chico's, explains, "The best benefit of the system is that I can flag customers who will receive a specific promotion in [the GIS software] and then bring those names back into [our software]. If a customer responds to a specific promotion, I know [they were] flagged for it and where [they] came from geographically," (14). "We probably use only a sliver of the GIS software functionality, but it has helped us," said Horton (14). "I can specifically map our current and prospective customers,

which helps Chico's decide new store locations and appropriate ZIP codes for promotional mailings.

Business Intelligence

Business intelligence is the process of gathering and analyzing business data to assist a company in understanding the weaknesses of competitors. Business intelligence enables a company to take advantage of those weaknesses when planning a marketing strategy (5). Business data have several important aspects (5). First, the customer is an individual and the attributes that represent that customer include the products that they might buy. Second, there are rows and rows of customer data and the similarities and differences among customers are represented in this data, which can be displayed graphically. Finally, the spatial relationship or location of customers with respect to a place of business as well as their relationship to each other is an important aspect of business intelligence.

Corporate organizations can plan successful marketing and relocation campaigns based on spatial analysis of business intelligence. Tony Burns, manager of corporate sales at ESRI, states that GIS works with the databases organizations already have. GIS can map total sales by ZIP code or sales territory, which can expose spatial patterns. He continues with the idea that it can link customer data to additional data like business clusters and determine the buying behavior of a specific demographic based on those patterns. The user can query, or geographically mine, the database to learn more information about customers.

CAMPBELL SCIENTIFIC BACKGROUND

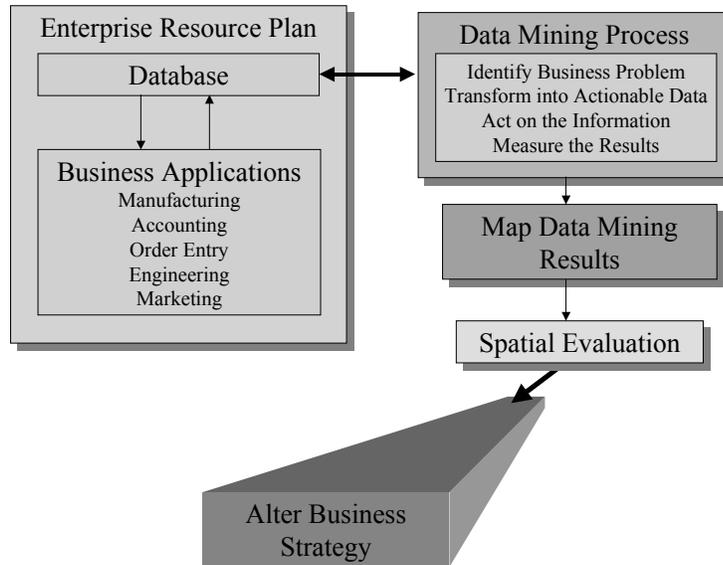
Campbell Scientific, Inc. incorporated in 1975 and has made numerous technological advancements in the areas of data loggers, data acquisition systems, and measurement and control products. An important feature of any enterprise resource plan is the ability to share information among multiple department, databases, and applications. Implementing an Oracle ERP in 1998 allowed CSI to integrate its business functions and provide better service to its customers while ensuring efficient manufacturing processes. CSI uses the Oracle ERP database and business applications to store and access customer, sales, and product data. The company does not have a strong marketing strategy at this time.

Benefits to Campbell Scientific, Inc.

GIS market analysis tools can be used to determine which products and promotions match the requirements and buying patterns of customers (5). With these tools, CSI can create a multidimensional snapshot of trends to create trade areas, predict equipment sales, and evaluate and update the design of sales territories. CSI can use this information to turn statistical and geographic data into meaningful information for business decision-making (5).

However, many additional benefits can be gained through the integration of a data mining application and geographic information systems. CSI can increase its market share of scientific equipment sales significantly through combining their business data with a GIS. Sales territories could be redefined based on information culled from combining the information extracted during the data mining process with the spatial component.

FRAMEWORK FOR COMPETITIVE SUCCESS



Framework for a Spatial Data Mining Mapping Application (DaMMAp)

Data warehousing

Based on the DaMMAp shown above, CSI stores their corporate data in the Oracle database, which is located within the Oracle ERP. The data mining processes are applied to the database to extract the information. The results are then mapped and a spatial evaluation is conducted to determine whether the corporate business strategy should be altered. By analyzing corporate business information, companies can effectively compete in the market and increase their profits. Business information, such as customer location and sales, is the most valuable asset a company can use to drive market competitiveness.

Data warehousing applications include the analysis of sales results and marketing programs and product profitability based on an individual products produced by CSI such as the CR-10 data logger. The Oracle ERP integrated CSI's departmental databases. However, corporate asset management, quality assurance, and statistical analysis are additional applications that a data warehouse can assist with that a standard ERP implementation does not address (10).

The design, construction, and management of a data warehouse determine the efficiency and usefulness of the data warehouse as part of a corporate strategy. Sigal (16) states several common sense development strategies for designing and building a data warehouse. Initially, the organization can reduce complexity of the database via standardization of the database schema. Prerequisites for implementing a data warehouse include personnel with a solid understanding of files and databases and a solid understanding of the organization's needs and practices. A

background in quantitative business statistics or quantitative methods for managers is helpful during the analysis process (13).

Data Mining

Data mining can easily accomplish several data organization tasks (1). Classification of data is one of the more common data mining tasks performed and consists of assigning a feature to one of a set of predetermined classes. Classification results in a discrete outcome. Estimation, on the other hand, results in a continuous outcome such as a value between 0 and 1. Prediction is similar to classification except that the records are classified based on some future expected value. Affinity grouping is used to define groups of records based on a frequently occurring item. Clustering groups heterogeneous records into more homogeneous groups based on a chosen value in the database. Clustering does not have any predetermined classes as in classification technique. Description is used to describe what is happening with the data in a database. Data mining analysis can benefit the organization by preventing customer attrition, cross-selling to existing customers, assisting in acquiring new customers, detecting fraud, and identifying the most profitable customers (4).

Knowledge Discovery

The potential for discovering, or mining, knowledge in corporate databases depends on the costs of the data mining application and the benefits of the data mining application. Brachman (2) lists several factors that affect this potential. There should be no *alternative* solutions that would make the process simpler. The key *relevant* factors need to be present in the data from the outset. The *volume* of data should provide a sufficient number of cases (several thousand at least) to provide a low statistical error probability. The *complexity* of the knowledge discovery can be increased as the number of variables (key factors) increase. Time-series, or versioned, data also increase processing complexity. The *quality* of the data should encourage a low statistical error rate. The data should be easily *accessible*; accessing data or merging data from different sources increases the cost of an application. The application should easily incorporate *change*. The expertise to run the application and interpret the results can have a major affect on the potential for mining knowledge from the data. Brachman emphasizes that expertise on the format and understanding the meaning of the data is as important as the knowledge of problem solving in the domain. Overall, knowledge discovery technology promises to improve the ability to cope with and exploit the growing abundance of data.

GIS

A major advantage of a GIS is that it allows an organization to identify the spatial relationships between tabular (sales) and spatial data (customer location) features. CSI already has their customer and sales spatial data housed in their Oracle ERP. Every time an Applications Engineer adds a customer to the database, the address information is included. The address information is used as the spatial component or location attribute in the business information. It can be used to create, manage, integrate, analyze, and disseminate data for every organization. A geographic information system can help organizations visualize their sales and marketing opportunities by

populating a map based on both corporate information (14). Any information that can be matched to an address can be crossed with demographic or census data through a GIS system.

Because graphics are very easy to interpret, maps depicting business data mined from a corporate database are helpful when faced with millions of customer names, addresses, and sales transactions (14). Spatial data mining is useful for businesses of all sizes that understand that the latest computer technology, along with the expertise to implement and analyze it, is critical.

CONCLUSION

Limitations

Once the data warehouse is deployed, it must deliver a high level of performance and maintain fresh data for fast, accurate decision-making. With a solid and reliable data warehouse foundation, a business intelligence system can turn high quality data into valuable information in the form of reporting, querying, and sophisticated analysis (12).

Future Research

Research indicates that combining the tools of data warehousing, data mining, and geographic information systems is a rare occurrence. The power of this combination has yet to be experienced in most corporate organizations.

Awareness of GIS has been growing steadily since its inception in the late 1960s. It is now becoming the next big development in data mining. GIS software allows complex demographic information to be displayed in colorful, easy-to-see format. The effectiveness of GIS lies in its ability to show, rather than tell, but, ironically, the word is out that data mining software is essential to the whole operation (18). Based on the discussion presented here, CSI can easily increase their worldwide customer base by incorporating a GIS into the Oracle ERP that CSI implemented several years ago.

REFERENCES

1. Berry, Michael J.A.; Linoff, Gordon. *Data Mining Techniques for Marketing, Sales, and Customer Support*. 1997.
2. Brachman, Ronald J.; Khabaza, Tom; Kloesgen, Willi; Piatetsky-Shapiro, Gregory; and Simoudis, Evangelos. Mining business databases. *Communications of the ACM* 39(11), pp. 42-48. November 1996.
3. *Campbell Scientific, Inc.* Retrieved November 2002 from <http://www.campbellsci.com>.
4. Cheevers, Sandra. ORACLE9I DATABASE SUMMARY *An Oracle White Paper May 2002* Oracle9i Database. Oracle Corporation World Headquarters 500 Oracle Parkway Redwood Shores, CA 94065 U.S.A.
5. ESRI. www.esri.com. Retrieved information November 2002.
6. Gardner, Stephen R. Building the data warehouse. *Communications of the ACM* (41)(9), pp. 52-60. September 1998.

7. Glymour, Clark; Madigan, David; Pregibon, Daryl; and Smyth, Padhraic. Statistical inference and data mining. *Communications of the ACM* 39(11), pp. 35-41. November 1996
8. Indulska, M. and Orłowska, M.E. On aggregation issues in spatial data management. *ACM International Conference Proceeding Series Proceedings of the thirteenth Australasian conference on Database technologies - Volume 5*, 2002, pp. 75-84. Melbourne, Victoria, Australia, Publisher Australian Computer Society, Inc. Darlinghurst, Australia.
9. Inmon, W. H. The data warehouse and data mining *Communications of the ACM* 39(11), pp. 49-50. November 1996.
10. Jones, Katherine Dr. Research Article: An introduction to data warehousing: what are the implications for the network? *International Journal of Network Management* 8(1), pp. 42-56. Published Online: 4 Dec 1998.
11. Murray, Chuck. Oracle Spatial User's Guide and Reference Release 8.1.7. September 2000. Part Number A85337-01, Oracle Corporation World Headquarters 500 Oracle Parkway Redwood Shores, CA 94065 U.S.A.
12. Patel, Mona. Oracle9i Database for Data Warehousing and Business Intelligence. *An Oracle Business White Paper*. Ensure the Success of Your Business Intelligence Project With Oracle. May 2002. Oracle Corporation World Headquarters 500 Oracle Parkway Redwood Shores, CA 94065 U.S.A.
13. Pierce, Elizabeth M. Developing and delivering a data warehousing and mining course. *Communications of the AIS* 2(3es) November 1999.
14. Roussel-Dupre, Stephanie. Build A CRM Road Map. Retrieved November 2002 from www.ismretail.com/articles/2002_03/020301.htm
15. Sen, Arun, and Jacob, Varghese S. Industrial-strength data warehousing. *Communications of the ACM* 41(9), pp. 28-31. September 1998.
16. Sigal, Michael. A common sense development strategy. *Communications of the ACM* 41(9) pp. 42-43. September 1998.
17. Spatial Data Warehousing for Hospital Organizations *An ESRI White Paper*. March 1998. Environmental Systems Research Institute, Inc., 380 New York Street, Redlands, CA 92373-8100 USA.
18. Totty, Patrick. SERVICE SOLUTIONS Pinpoint members with new data-mining tools. Geographic information systems help your CU map its future. Page 32. CREDIT UNION MAGAZINE, April 2002. www.cuna.org
19. Watson. Hugh J. and Haley. Barbara J. Managerial considerations. *Communications of the ACM* 41(9), pp. 32-37 September 1998.