The role of intelligent agents and data mining in electronic partnership management

Merrill Warkentin [a], Vijayan Sugumaran [b,c,*], Robert Sainsbury [d]

[a] Department of Management and Information Systems, Mississippi State University, P.O. Box 9581, Mississippi State, MS 39762-9581, USA
[b] School of Business Administration, Oakland University, Rochester, MI 48309-4401, USA
[c] Department of Global Service Management, Sogang Business School, Sogang University, Seoul 121-742, Republic of Korea
[d] Gravity Jack, Spokane, Washington, USA

ARTICLE INFO

ABSTRACT

The marketspaces of the "New Economy" and the eServices revolution have enabled the formation of new types of partnerships which are electronically mediated. Web-based electronic commerce has also brought a tremendous increase in the volume of data that can be mined for valuable managerial knowledge. The data mining procedures used in this process can be enhanced by employing intelligent agents. This paper describes emerging electronic partnerships between players in developing electronic marketspaces and identifies typical data flows between such players, with an analysis of the potential role of data mining and intelligent agent technology. By identifying the complex nature of information flows between the vast numbers of economic entities, we identify opportunities for applying data mining techniques that can lead to knowledge discovery. In particular, we show how a Generic Agent-based data Mining Architecture (GAMA) can be customized to support managerial decision-making and problem solving in a networked economy. A prototype implementation of GAMA is presented, along with a demonstration of the some of the capabilities of the system. Finally, we explore the role of agents in promoting and maintaining strong automated relationships between various strategic partners.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Millions of individuals surf the Web every day and interact with electronic commerce Websites around the world. While many sites capture user activity, most do not capture all interactions with "etail" (electronic retail) consumers, suppliers, and partners, and they do not maximize the potential uses for such data (Liu, Cao, & He, 2011; Willow, 2005). According to Forrester Research, only 18% of the companies it surveyed use their Web data for marketing purposes, and only 16% use it for customer support. The Forrester study also indicates that 72% of the companies that collect Web data admit that they do not analyze this data or use it in any application. However, organizations are beginning to realize the value of this Web data and are allocating vast resources for creating the necessary infrastructure to analyze this data, which would enable them to learn more about their customers and gain a competitive advantage (Rajan & Saravanan, 2008; Yarom, Rosenschein, & Goldman, 2003; Zuo & Hua, 2012).

The sheer volume of data generated by the activities of visitors to a company's site (their "digital footprint") poses various problems in the storage, management, and analysis of this data as well as new opportunities. Some companies rush to set up their electronic storefronts, focusing more on transaction processing, online inventory, shopping carts, and ad banners, without giving sufficient consideration to the data management issues (Gomes & Canuto, 2006; Holmes, Tweedale, & Jain, 2012). In order to get the most mileage out of this data, each company must decide: (a) what data to collect and how to organize it; (b) what kind of analysis to perform on the data; (c) how frequently to perform data analysis; and (d) how to validate and integrate the results into decision making and planning.

As data warehousing and data mining technologies mature, an increasing number of organizations are employing these technologies in their problem solving and managerial decision making in the business to consumer context (Rao, 2010). Through data mining, a company can synthesize consumer Website patterns into meaningful information, enabling it to understand and engage customers and prospects over the Internet (Chaimontree, Atkinson, & Coenen, 2012; Nassiri, 2009). The mining of Web-based data and the implementation of the business intelligence it represents is the key to creating a lasting relationship with online customers

* Corresponding author at: School of Business Administration, Oakland University, Rochester, MI 48309-4401, USA. Tel.: +1 248 370 2831; fax: +1 248 370 4275.
E-mail addresses: m.warkentin@msstate.edu (M. Warkentin), sugumara@oak-land.edu (V. Sugumaran), sainsrob@hotmail.com (R. Sainsbury).

and establishing a successful online storefront. But in the future, mining the Web service interactions between company Websites will also enable them to create lasting relationships with their strategic partners (Jain, 2012; Marik & McFarlane, 2005).

There are several commercial software products available to analyze Web traffic data such as NetTracker, WebTrends, NetIntellect, HitList, and SurfReport. However, these products are limited to analyzing server activity based on the data stored in log files. By unifying the log data with personal information supplied by vendors such as Equifax, Experian, TransUnion, MetroMail, and others, one can develop a more complete customer profile. This integrated information can then be mined to gain insight into who is buying what products, what products are the most popular, buying patterns, and so forth (Gao, Yan, & Dang, 2012; Jayabrabu, Saravanan, & Vivekanandan, 2012; Kehagias & Mitkas, 2007). The dynamic nature of the online environment dictates that this analysis should be performed promptly to enable companies to quickly respond to changes in customers' buying behavior.

There are myriad data mining tools available in the market that employ a variety of data mining algorithms and techniques. For a novice user, it is often difficult to determine which tools or techniques are appropriate for a particular data analysis or data mining scenario. Companies are beginning to employ "intelligent agents" (Chan, Fan, Prodromidis, & Stolfo, 1999; Gannon & Bragger, 1998; Gorodetsky, Karsaev, Samoylov, & Serebryakov, 2008; Grimes, 1998; Sugumaran & Bose, 1999) to reduce some of this cognitive load. These agents can automate some of the mundane activities such as data cleansing and data transformation and can help the user in the selection of appropriate tools and data mining methods (Li & Li, 2011; Lee & Liu, 2004; Moemeng, Zhu, Cao, & Jiahang, 2010). Typically, intelligent agents act on behalf of the human user in problem solving activities and decision making.

The objective of this research is to: (a) study the information flow between various entities in different electronic markets; (b) investigate how data warehousing and data mining techniques can be applied for discovering new relationships and nuggets of knowledge that could be incorporated into managerial decision making; and (c) develop a generic architecture for an intelligent-agent based data mining environment; and (d) apply this architecture to various eCommerce marketspaces to help the user validate and interpret the results, thereby enabling the discovery of valuable knowledge.

The remainder of this paper is organized as follows. The next section ("Emerging eCommerce Marketspaces and Data Flows") presents a vision of relationships between strategic partners facilitated by the implementation of interoperable automated processes. The following section ("Relationship Management in the New Economy") focuses on collaborative commerce (cCommerce) activities such as outsourcing and establishing strategic partnerships through the use of intelligent agents. The next section ("Emerging Technologies") provides an overview of emerging technologies such as data warehousing, data mining, and intelligent agents, and shows some examples of their use on the Web. The following section ("Generic Architecture for Agent-Based Data Mining") proposes an architecture for a generic agent-based data mining environment. This Generic Agent-Based Data Mining Architecture (GAMA) can be customized to support managerial decision-making and problem solving for a particular application. The penultimate section ("Agent-based Data Mining Applications in eCommerce") provides a detailed discussion of the application of intelligent agent-based data mining in different electronic marketspaces such as etailing and B2B exchanges. The final section ("Managerial Implications and Future Directions") concludes the paper by discussing the issues in agent-based data mining and their managerial implications.

2. Emerging eCommerce marketspaces and data flows

The term "electronic commerce" has been used to describe a disparate variety of business activities. The focus in this paper will be on relationships among various economic entities (individuals and Websites) which are supported by the standard communication and data representation protocols of TCP/IP, HTTP, HTML, and XML.[1] Many such relationships are built on data flows in the background of the business-to-consumer (B2C) channel or between two business computers (business-to-business or B2B) which are transparent to the human user (Zhang, Wang, & Shen, 2012).

The adoption of electronic commerce technologies has enabled entirely new economic market models that connect many firms to consumers (etailing or B2C eCommerce) and to other businesses (B2B eCommerce). Associated with these emerging markets are new strategic business models, frequently built upon strong strategic alliances with firms that provide critical services in the value chain (or value web).

Business-to-Consumer eCommerce (or etailing) has evolved from an obscure commercial niche to a widely-accepted marketing channel for selling countless retail items. A number of factors distinguish etailing from traditional markets, but the most important one is its implementation of electronic customer relationship management (CRM). Etailers can effectively analyze buyers' clickstream data and purchase data in order to offer personalized services and product offerings (Tuzhilin, 2012). An online bookseller, for example, may display an opening screen that features items similar to those a specific customer has purchased or researched in the past, and it may send email reminders when new items arrive that may interest the buyer. The online seller may also enable flexible ordering processes, online shipment tracking, and quick checkouts.

Business-to-business eCommerce is evolving from a proprietary EDI based, closed, expensive, and non-scalable system to an open, inexpensive, and scaleable system where multiple suppliers and buyers are connected to each other by Web-based exchanges. Because procurement activities comprise such a significant portion of many companies' budgets,[2] the greater selection and transparency (price, availability, supplier, product) that the B2B system brings has been largely welcomed by the capital markets. *Vortals* (industry-specific vertically integrated portals) and *e-procurement* represent the majority of B2B activity.

In a later section, we analyze the role of data mining and intelligent agents in *B2B eCommerce* where multitudes of buyers and suppliers interact to share information regarding products and conduct transactions using a variety of market mechanisms like contract procurement, auctions, reverse auctions, and bid-ask spot markets. The presence of multiple players and complex information flows between them creates interesting possibilities for knowledge discovery using data mining. Agent-based data mining can facilitate the establishment of automated relationship management by identifying trends in purchase activity, profiling business partner activities, and offering support for various online exchange activities (Liao, Yang, & Geng, 2011).

3. Relationship management in the new economy

In today's rapidly changing eBusiness environment, it is imperative that organizations have flexible organizational structures that can accommodate partnerships with a variety of external

---

[1] Transmission Communication Protocol / Internet Protocol, HyperText Transfer Protocol, HyperText Markup Language, and eXtensible Markup Language.

[2] US-based Fortune 1000 companies spend approximately half their budgets on various categories of procurement, ranging from major acquisitions (such as raw materials and sub-assemblies purchased by heavy manufacturers) to routine office supplies.