

# Data-Driven Marketing

by Maurice Mulvenna, Marian Norwood, and Alex Büchner, University of Ulster, United Kingdom

## Abstract

Increasingly, companies view the Internet as a selling and marketing channel for their products and services. This paper outlines the context for this viewpoint, describing how the marketing model for electronic commerce players may be revolutionised through the application of data mining technologies. A major research project called MIMIC that will explore use of data mining in Internet retailing is described. Finally, the application of data mining to other Internet areas is explored, and issues of ethics and anonymity are discussed.

## Key Words / Phrases

Data Mining, Data-Driven Marketing, Electronic Commerce, Internet Marketing, Internet Retailing, MIMIC

## Introduction

Electronic commerce has been described as “one of those rare cases where changing needs and new technologies come together to revolutionise the way in which business is conducted” (ESPRIT 1996). Currently, relatively few consumers and businesses are connected to the Internet. Projections estimate over 50 million on-line by the year 2000, but is electronic commerce ready for the rigours of international trade, both retail and business to business?

Much of the excitement of electronic commerce is focused on the delivery of products and tradable services to consumers in their own homes. This is usually known as on-line retailing and is seen as the area of electronic commerce where many of the following issues will be debated first: standards; international electronic trade agreements; security; and trust. This on-line retailing model is the permutation selected in the MIMIC project (described later) to illustrate the potential of data mining. The primary reasons for selecting the on-line retailing model are:

1. Behavioural - Consumers exhibit a wide range of differing behavioural patterns;
2. Marketing - Consumer and product descriptive information provides a rich source of useful data;
3. Descriptive - Products and services in on-line retail sites are varied.

**Figure 1. Internet Retailer Web Site Processes**

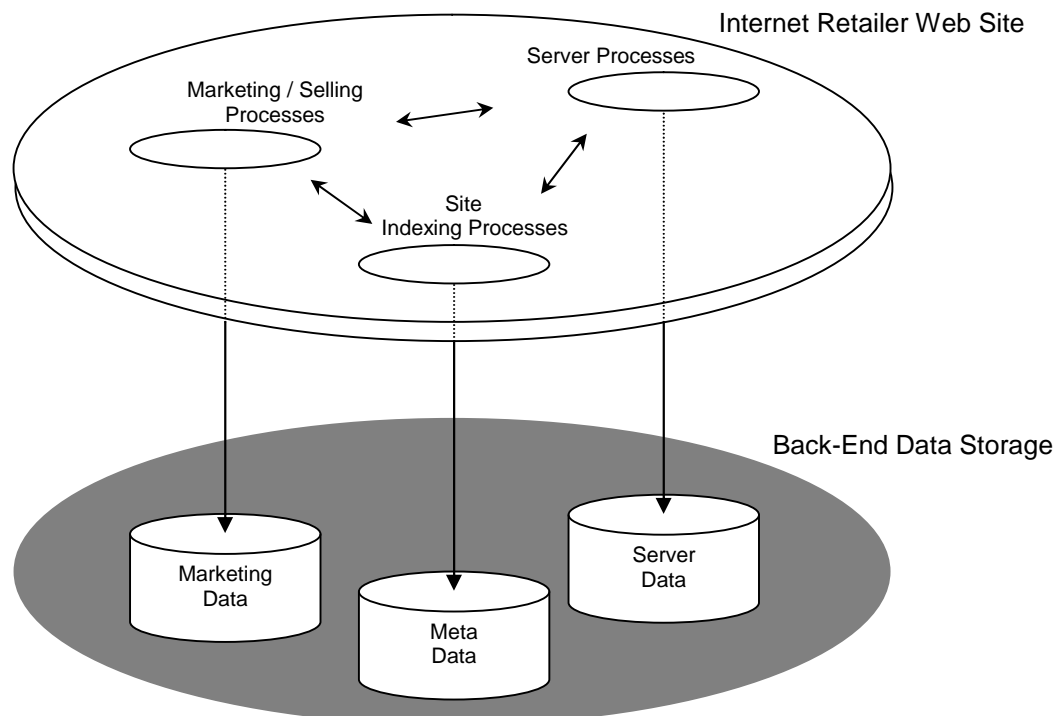


Figure 1 (above) illustrates the main processes that occur at an Internet retailing web site, along with the associated data stores. The processes in Figure 1 provide data, which is at the core of successful Internet marketing and suitable for data mining. The data may be described along the following dimensions (which correspond with the reasons given above):

1. Server data – data generated by the interactions between the persons browsing an individual site, and the web server. For example, the ip address of a computer browsing a retailing site.
2. Marketing information data – the data stored by the Internet retailer on products, customers, suppliers, etc. For example, the consumer responses to discounting.
3. Site ‘meta’ data – the data about the site, usually generated dynamically and automatically after a site update. For example, location of product pages on a site, which are ‘leaf’ pages, navigational pages, etc.

There are now hundreds of thousands of on-line Internet shopping providers world-wide, with thousands more appearing each month. Data generated by clients of the on-line retailers by their interaction with the site in browsing and buying. Clearly such large data sets, which may be distributed and heterogeneous, will contain useful information helpful to business marketing strategies, both for retrospective analyses as well as data-driven forecasting.

The potential for this new form of *data-driven* marketing is enormous – browsers and buyers of products can be identified and targeted with attractive offers and sales promotions.

### **Data-Driven Marketing**

(Dibb; Simkin; Pride; Ferrell 1997) define marketing as consisting of “individuals and organisational activities that facilitate and expedite satisfying exchange relationships in a dynamic environment through the creation, distribution, promotion and pricing of goods, services and ideas”. This definition highlights the need for the development of the ‘right’ marketing mix (product, price, promotion, distribution, etc.). According to (Kotler 1997) this is “a set of marketing tools that the firm uses to pursue its marketing objectives in the target market”. The development of the ‘right’ mix is governed by an understanding of customers’ needs and wants.

Sophisticated successful marketing takes this a step further. An organisation must understand its customers’ needs and wants, and it must satisfy them with a product or service that displays some form of competitive advantage. With an understanding of not only the customer but also the marketplace (the business environment and competition) as well as internal capabilities, a marketing strategy can be developed. In developing the strategy, an organisation attempts to identify groups of customers where each separate group, or market segment, has ‘similar’ needs. Each group of customer can then be offered a specifically tailored programme, also known as target marketing.

However, it is a well-known fact that success in marketing is never guaranteed. In today’s dynamic business environment, marketing decisions are extremely complex. Such decisions as ‘what customers to target’ or ‘what products or services to offer targeted customers’ demand reliable and accurate knowledge. As (Druker 1983) states, “The aim of marketing is to make marketing superfluous. The aim is to know and to understand the customer so well that the product or service fits him/her and sells itself”.

It is interesting to note that the Internet effectively blurs the distinction between the traditional marketing processes of a) informing customers, b) selling to customers (via advertising and personal selling strategies) and c) order taking, processing and even delivery. Increasingly, customers can react immediately to a marketing communication aid, and purchase directly. An example is the purchase of software that is downloadable upon payment. The Internet therefore facilitates instantaneous advertising, selling and distribution.

Normally, all marketing communications are subject to measurement and control in order to monitor the success or otherwise of campaigns. Traditionally effectiveness is measured by tracking awareness and image, by measuring response rate to all direct communications and of course by looking at sales volume. It involves a complex series of tasks. This is not necessarily so for Internet marketing. Organisations that market on the Internet can now measure directly the response generated by the campaign, advertisement, discounting, etc. In the case of online ‘banner ads’, the success is measured by the ‘click through’ rate. This is the number of times that, when someone is presented with an ad, they then choose to click on that ad to obtain further information. A company hosting its own merchant server can measure the number of ‘hits’, and break down this information by domain (e.g., .com for commercial).

The above examples illustrate that each interaction between consumer and seller is recorded in digital form. This is the important difference between conventional selling and selling on the Internet. At this stage in the Information Age, each actor in the information cyberspace, for example, consumers and retailers in Internet retailing, generates digital trails and data that may be stored and analysed by the organisation operating the web server. Data mining algorithms are the new actors in this information cyberspace.

### **Data Mining**

Data mining has been defined as “the automated discovery of non-obvious, potentially useful and previously unknown information from large data sources” (Frawley; Piatetsky-Shapiro; Mattheus 1991). It applies various artificial intelligence techniques to discover patterns that may be both useful and actionable.

Data mining programs comprise algorithms that can discover *associations* between data, help to *classify* items, and discover *sequences* hidden in the data (Anand 1998). An example of a discovered association could be the types of products that consumers usually buy at the one time. An example of classification would be the classification of a particular type of buyer, such as 'high frequency but low value' purchaser. An example of discovered sequences could be a buyer periodically returning to purchase other novels by a particular author.

Within the context of data analysis of Internet retailing sites, data mining can discover knowledge about the unique characteristics of identified customer segments, so that business decisions may be made in relation to predictions about customer value and appropriate loyalty incentives can be developed (Mulvenna 1997a).

Data mining allows marketers to reveal layers of information about markets (or subsets of markets) in ever increasing detail, enabling customer or prospect profiles to be built and the identification of the segments upon which marketing activities can focus.

So what are the benefits of using data mining techniques to gather intelligence about a market? It allows on-line retailers to 'fine tune' their selling strategy. This gives them a greatly enhanced insight into the number and types of lines in stock, the best electronic shop front format, and which offers should be directed at which customers and how should they be communicated (Humbry 1996). Using direct, data driven marketing, heavy buyers of a product can be identified and targeted with attractive offers and sales promotions. Targeting does not have to involve discounted offers alone. It could mean making certain categories of customer aware of products or services that might be of interest to them, or inviting them to special on-line events. The application of data mining techniques to the data of on-line retailers provides high-level knowledge – for example, in the form of rules - that describes consumer navigational and purchasing behaviour. These rules capture trends and behaviour patterns that may be applied within a marketing strategy. The authors propose that the high-level, descriptive, behavioural rules supply the marketing specialists with the means to feedback directly the aggregated behavioural responses of the clients of an online service.

The high-level mined rules may be incorporated into the architecture of an on-line retailer. When each person interacts with the on-line service to navigate and purchase goods or services, their on-line behavioural patterns are identified, and the on-line retailing system may react to change dynamically the information presented to that consumer.

### **The MIMIC Project**

MIMIC ("Mining the Internet for Marketing IntelligenCe") is a project funded by ESPRIT IV Framework Thematic Programme for Electronic Commerce (MIMIC 1997). The MIMIC Consortium consists of four partners, represented by the University of Ulster, and three small-to-medium enterprises (SME).

The technical objective of the MIMIC project is to develop a data mining toolkit that will make possible the mining of the data generated by online retailers (Mulvenna 1997b). The business objective is to provide European on-line shopping retailers (of which many are SME companies) with advanced technology that will enable them to maintain their competitiveness in the global Internet marketplace. The approach of the MIMIC project is to capture the requirements from SME on-line shopping providers, and develop specialised Internet-capable data mining algorithms. These will be validated by the users and incorporated ultimately into a special version of existing data mining software.

The target market segment is on-line shopping retailers, and the goal of the MIMIC project is to give these retailers the ability to market their products and services effectively. Such customised and directed marketing communication, however, relies on accurate information about the customer. MIMIC will provide, for the first time, the capability for retailers to consider a cost-effective means of customised data-driven marketing communication (Norwood 1997).

Although it is widely believed that mass marketing techniques such as advertising are a major proportion of a marketing budget, many companies now spend more on sales promotion methods, which include direct marketing techniques. The benefits of data-driven marketing are the ability to measure directly the response to the marketing effort, and the resultant cost-effectiveness of this technique. Data-driven marketing also enables on-line retailers to monitor their customers' behaviour patterns and to detect and entrap potential 'switchers' – those people who show no allegiance to a particular retailer. Electronic Commerce dispenses with geographic boundaries, and empowers customers to shop where they obtain the cheapest/best bargain. The marketing intelligence information from the MIMIC data mining toolkit may be used to provide a competitive edge for on-line shopping retailers, allowing them to capture, retain and satisfy customers.

The three areas of data mining outlined in the previous section – classification, association and sequences – will require additions and specialised extensions to cope with the specific nature of Internet data.

## Conclusions

This paper has outlined how data mining may be applied to assist the marketing communication process for Internet retailers. The algorithms used in data mining are capable (with modification) of processing of large, distributed and often heterogeneous databases that store the transactional, behavioural and other data gathered by Internet retailers.

The use of data mining in data-driven marketing for on-line retailing is clear, but its market application may be tempered by customers and consumer organisations who may react negatively to the collection and 'mining' of aggregated or personal information (Toronto Star 1998). It is impossible to predict the future landscape of the Internet and the roles of the actors in this information cyberspace. However, it is becoming apparent that consumers value their rights and will only participate in marketing relations that benefit them as much as the retailer (Financial Times 1998, Easton; Parker 1998). The onus is therefore on the retailers to provide assistance and engender trust. This may be done positively, as a value adding action (Thirkell 1997).

Internet retailers who do not attempt to address the trust issue directly face low turnover at best and vigorous collective reactions from consumers at worst. Ultimately, consumers and other actors in cyberspace will seek out lateral solutions to what they view as privacy intrusions (Hagel III; Rayport 1997). These may include a desire for anonymous transactions; for example, using digital cash equivalents.

Data mining algorithms and techniques may be applied to more than Internet retailing data. Other areas that involve Internet transactions include business-to-business electronic commerce. If the electronic commerce market grows as some commentators predict, there will be an explosion in the growth of digital transactional and behavioural data, and an increased demand for data mining tools and techniques. The results of the MIMIC project should demonstrate the usefulness of data mining techniques, and explore the implications of the application of these techniques to retailers and consumers alike.

## References

Anand, S.S., Büchner, A.G., Decision Support Using Data Mining, Financial Times Pitman Publishers, ISBN 0-273-63269-8, 1998

Dibb, L. Simkin, W.M. Pride, O.C. Ferrell: Marketing Concepts and Strategies, 3<sup>rd</sup> European Edition, Houghton Mifflin, Boston, New York, 1997.

Drucker, F., Managing for Results, Heinemann, London, 1983.

"Electronic Commerce - An Introduction", <http://www.cordis.lu/esprit/src/ecomint.htm>

Frawley, W.J., Piatetsky-Shapiro G. and Matheus, C.J., "An Overview: Knowledge Discovery in Databases", in Knowledge Discovery in Databases, AAAI/MIT Press, 1991, pp. 1-27

"Knowing When to be of Service", Financial Times, London, 12 February, 1998

Hagel III, J., Rayport, J.F., "The Coming Battle for Customer Information", Harvard Business Review, January-February 1997

Humbry, C., Digging for Information, Marketing, November 21, pp.41-42, 1996

Kotler: Marketing Management: Analysis, Planning, Implementation and Control, 9<sup>th</sup> Edition, Prentice Hall International Inc, pp. 466-467, 1997.

Mining the Internet for Marketing Intelligence – MIMIC, Esprit Proposal N°:26749, 1997

Mulvenna, M.D., Büchner, A.G., Data Mining and Electronic Commerce, in Overcoming Barriers to Electronic Commerce (OBEC '97), Malaga, Spain, 1997

Mulvenna, M.D., Büchner, A.G., Norwood, M.T., Grant, C., "The 'Soft-Push': Mining Internet Data for Marketing Intelligence", Invited paper: Working Conference on Electronic Commerce (EUROMED), Metsovo, Greece, July 1997

Norwood, M.T., MULVENNA, M.D., Büchner, A.G., Grant, C., "Competitive Advantage on the Internet Through Relationship Marketing", 1<sup>st</sup> Berlin Internet Economics, October 1997

Thirkell, P.C., "Caught by the Web: Implications of Internet Technologies for the Evolving Relationship Marketing Paradigm", Conference of American Marketing Association, Dublin, Ireland, June 1997

“Data Mining’ Privacy Threat”, Toronto Star, Toronto, Canada, 14 January 1998

Easton, J., Parker, J., “Seeking the *There* There: Building Relationships is the Key to Building Profits”, ZD Internet Magazine, February 1998

### **Pen Portraits**

Maurice Mulvenna – Maurice is a lecturer in informatics at the School of Computing & Mathematics in the Faculty of Informatics, University of Ulster. He has published over thirty papers on his research specialities - knowledge based systems and electronic commerce. He is interested in research and technology transfer with small companies in Northern Ireland, and is the project leader of the CEC MIMIC electronic commerce project.

Marian Norwood – Marian is a lecturer in marketing and marketing research at the School of Commerce and International Business Studies in the Faculty of Business & Management, University of Ulster. She publishes in the areas of Internet marketing and relationship marketing. She is currently undertaking several market research projects in the North-west Ireland. She is also a fanatical Manchester United supporter.

Alex Büchner – Alex is a research fellow in the Northern Ireland Knowledge Engineering Laboratory in the Faculty of Informatics at the University of Ulster. He works in the area of Data Mining and is involved in several research projects in Europe and the Far East. His main research interests are in the fields of artificial intelligence, object-oriented technologies and multi-databases. He has written more than 15 publications in these fields.

Address for correspondence:

Maurice Mulvenna  
School of Computing & Mathematics  
Faculty of Informatics  
University of Ulster, Magee College, Northland Road,  
Derry, Co. Londonderry,  
BT48 7JL, UK

Email: [MD.Mulvenna@ulst.ac.uk](mailto:MD.Mulvenna@ulst.ac.uk)

Tel: +44 1504 375437