

3.3.2 FUTURE CASES: DATA MINING IN VIRTUAL ORGANIZATIONS

Lex Kwee², Walter Kusters³

INTRODUCTION

In this article we describe two cases of future companies. These companies are virtual in two or even three aspects. First, in having a virtual business model, in which all non-core activities have been outsourced to specialist providers. Secondly, in being active in electronic business wherever possible, reducing physical operations to a minimum by making maximum use of information technology. Finally, the second company does not sell physical products. Its business is based on bits and bytes of intellectual property that can be distributed electronically. The only physical product aspect is packaging to make the core product more accessible.

.....
2 Drs A.Y.L. Kwee, lk@zi.nl, New Business Associates, Abcoude, The Netherlands

3 Dr W.A. Kusters, Kusters@wi.leidenuniv.nl, University Leiden, LIACS, Leiden, The Netherlands

The virtual organization

The concept of the virtual organization was launched in the early nineties. Davidow and Malone described how companies focus on providing the ultimate customer service by applying organizational concepts such as mass customization in a virtual business model [Davidow, 1992]. Mowshowitz presented a theoretical framework for business design and operational decision making in virtual organizations [Mowshowitz, 1994; Mowshowitz, 1997]. Both approaches depend on the availability of large quantities of operational data and explicit business rules that transform operational data into planning recommendations and fulfillment execution. It will be evident that a virtual organization depends heavily on data mining.

Supply chain outsourcing

A virtual organization is set up around selected groups of customers. Every activity that does not directly contribute to the customer's experience and satisfaction may be outsourced. Examples can be found at non-manufacturing companies such as Nike and Ikea, which focus on product design, marketing and distribution, leaving actual production outsourced to low cost providers, and, in the case of Ikea, rely on the customer for final product assembly. It is predicted that traditional manufacturers will increasingly adopt this approach, as can already be seen in the automotive and IT hardware industries, in which the traditional manufacturers have outsourced the majority of manufacturing work, turning their traditional factories into final assembly shops and warehouses. Some companies, such as Hewlett Packard, have even outsourced final assembly and warehousing, making it part of the delivery process, which can be outsourced to global logistics providers and supply chain fulfillment companies. A similar approach in retail was pioneered by Wal-Mart, that sends cash register sales data directly to its suppliers and allows them into its stores to restock the shelves. This type of strategic outsourcing is dependent on highly integrated design, product management and supply chain planning systems, in which data mining applications constantly monitor behavior according to preset schedules and exception criteria.

Collaborative electronic commerce

The latest virtualization trend is the rise of collaborative electronic commerce, in which trade partners cooperate to serve markets, manage suppliers and logistics and implement electronic match making mechanisms, using a shared external ICT infrastructure and a common marketplace service. All these activities are highly data dependent, requiring a high degree of system and process connectivity. The increases in possible trade relations and electronic transaction volumes generate significant operational complexity. Companies need to connect, exchange and translate information electronically, providing instant feed-

back and situation specific pricing, terms and conditions. This forces companies to align processes and adopt emerging standards to achieve the benefits of fully automated collaborative electronic commerce. End customers can interact with the business systems of all players in a supply chain, leaving valuable data traces, which can be used for future marketing and customer service purposes. As the level of supply chain cooperation increases and the volume of electronic commerce rises, the volume of search, transaction and commerce related data will explode. Market tactics and business operations are becoming increasingly dependent on data mining.

CASE 1: AUTOMATED SUPPLY CHAIN

A corporate Purchaser usually buys from a pre-selected set of suppliers. When a good offer is found at an electronic marketplace, an order is initiated. The information is automatically sent to an electronic supplier clearinghouse. The clearinghouse transaction hub:

- applies the Purchaser’s business rules to match the supplier profile to the Purchaser’s criteria;
- translates the order to a processing standard that is used on the shared supply chain management system;
- performs a credit check and
- sends the information to the selected supplier.

The supplier receives the order and sends a confirmation, via the hub, where it is translated and forwarded to the Purchaser’s planning system. Given the value of the order, a business rule instructs the transaction hub to send a notification to the Purchaser’s mobile phone for confirmation.

Fulfillment is monitored automatically and when the predefined situation is reached, another business rule triggers the transaction hub to send an email to inform the Purchaser that the order is ready to be shipped and that all export documents have been cleared. If no exceptions occur, a payment transaction is triggered automatically.

During these transactions, a lot of data is produced. Mining these data will expedite future transactions. The Purchaser will combine process and fulfillment data to manage vendor rating and volume purchase contracts. The Supplier will combine process and payment data to manage a buyer rating and update pricing policies. Unexpected patterns will be used to activate exception handling and notification rules. Purchaser and Supplier may decide to outsource data mining to the marketplace, which has already outsourced this to the transaction hub. Demand analysis and forecasting will be used to determine optimum levels of stock, spare parts and production capacity.

CASE 2: MULTIMEDIA DISTRIBUTION

Background

A global multimedia conglomerate deals with all aspects of entertainment. The firm is large, and its activities take place throughout many different countries. In addition to internally produced entertainment goods, the firm sells external products, and uses external providers for specific technologies (e.g. video games). The firm's core business in this situation is branding. Distribution of media is outsourced, but all logistic and service data are collected for internal analysis.

In this approach, the firm does not know its final customers. Therefore, it makes heavy use of coupons and web forms, encouraging consumers to exchange personal data for discounts.

Also, the firm has a live entertainment division that provides customers with the opportunity to interact with the firm and its products in an attractive environment.

Future

Logistics

For an entertainment company, logistics are crucial. If products are ordered using the Internet, both the customer and the firm need to keep track of the status of the order and its payment. As many intermediaries (banks, logistics providers, warehouses) are usually involved, order tracking used to be a major problem. Agents can take care of this by searching the appropriate databases (which can be of totally different structure), extracting the knowledge about the product at hand, and presenting this to both customer and company. Data mining will predict logistic bottlenecks (by examining records from the past) and will come up with alternatives (by examining the existing resources). Finally data mining provides tools to deal with fraud. Data from transactions from the past can be used to detect fraudulent patterns.

Improved techniques allow for 'just in time' delivery on all levels. Since many products are digital, network behavior and server load prediction deserve attention. New techniques combine recent and expected behavior into on-line monitoring of the whole process. This may facilitate dynamic pricing, again based on customer observation. It is expected that in many industries pricing will follow the dynamic revenue enhancing approaches that have been developed for airway companies and have led to the myriad of price variations for simple airplane tickets. All available parameters are used to find a match between the buyer's willingness to pay a price and the seller's risk of having to fly with empty seats.

These approaches rely on heavy computing power and can support a limited number of transactions per second.

Marketing

The main area where modern data mining techniques can be applied is marketing. Here the questions are quite diverse. One may ask for customer profiles, but possible connections between purchases and predictions of sales are also interesting. A database of past transactions is extremely useful for customer profiling, in combination with the web-log database this may yield powerful tools. It may also be possible to automatically combine different databases (data fusion).

It may be imagined that adding computing power will lead to a situation in which the user's information environment continuously adapts, maintaining a personalized commerce space in which multiple offers are available, and adapted in real time. This will not only include personalized pricing, but also personalized delivery, packaging and even personalized products, in a future extension of Dell's seemingly 'made to order' PC assembly model or Unilever's Rituals 'made to order' cosmetics approach. Soon personalized shops will be available on the Internet for all types of configurable and customizable products. New graphical techniques will be used to spice up the presentation.

Amidst these, mainly technology driven, 'enhancements', the high level goal is a consistent and meaningful view of customers. For this purpose all existing databases need to be examined and their contents should be combined. Existing models will be further developed and improved using extensive calculations. On the one hand black box methods get better and better, on the other hand the management stresses the importance of understanding behavior. Location specific personalization, using inputs from mobile communications and computing devices, will result in real-time situation specific personalization. The entertainment company will not send its offer for a comprehensive vacation package (entertainment included) to a user's mobile phone during a car drive in an unknown inner city. The offer will be held until the prospective customer has reached his hotel, and switches on the television, where he can watch a highly personalized version of what his next vacation could look like. The trigger for this specific offer came after mining location data from his mobile phone. If he decides not to switch on the television, the phone will point him to a bookstore, museum or restaurant, all in accordance with the customer profile, preferences, situation and location.

When consumers prefer to be entertained at home, suppliers will be ready to serve their home entertainment centers. Entertainment on demand such as games and virtual trips in space can be provided, using equipment such as body suits and 3D visualization. These techniques offer a lot of opportunities for the entertainment business, but also raise a lot of questions concerning logistics, customer profiles, privacy, etc. Participants in such a virtual reality game or trip want to be amused in a way which fits — among other things — their emotional state and personal feelings. In order to determine the emotional state one has

to investigate the behavior of the customer. The virtual entertainment center gives a lot of information on this behavior; the physical reactions and the reflexes of a person to a certain situation, how the person proceeds, alterations to the scene, and all sorts of other information. All this information is, of course, encoded in raw data. Data mining can be used to analyze, decode this enormous amount of data and to determine the part of the current customer profile relevant for the virtual entertainment of the customer. The books or CDs a person buys can (and must) also influence the profile, since it gives information about the kind of entertainment a person likes. The ending of a person's favorite stories gives relevant information on the endings for the virtual entertainment games. So one also has to mine this data in order to get the very specific customer profile needed to offer interactive entertainment that fits the desires of the customer. This customer profile together with the perceived temporal emotional state of the participant in the virtual entertainment session can be used to determine the continuation of the virtual interaction in such a way that the participant is very satisfied at the end of the interaction. In this way personalized entertainment is within reach. On the other hand the behavior in the virtual entertainment center can be used to extend the customer profile for the more classical channels of entertainment, and makes it possible to offer certain (electronic) books or CD's that correspond to the behavior in the virtual world.

Accounting

In business areas where available data are more complete and precise, such as accounting and financial reporting, data mining may be used in semi-autonomous mode, for instance, in decision systems. In areas where data is scarce, incompatible and or unstructured, human supervision and analysis will be needed.

One can expect that the methods mentioned above will be extended and improved. In a sense the techniques might still be called classical, since the main new ingredient is computing power.

RELATED ISSUES

On the basis of these previous examples of future companies, we may predict several issues that deserve our attention.

Privacy

The enormous quantity of customer data is unfortunately prone to politically incorrect approaches. It is also clear that border crossing activities, especially when dealing with information in general and entertainment in particular, may lead to unwanted behavior. For this topic we refer to Part 4.

Unclear pricing mechanisms and information overload

We expect to see many autonomous competing firms, who are eager to use novel technologies in a rapid way. Technology will result in new combinations of product, packaging, personalization and delivery. Pricing will have to reflect these new possibilities. Some authors argue that the decisive factor in buying decisions will not be product-based. Rather, experience (immediate, user controlled delivery) will be the determining factor that sets new price levels. In addition to the pricing models of airlines, companies will have to adopt menu and serving suggestions from restaurants and other leisure companies.

As a countermeasure against any possible information overload driven by data mining, consumers can use data mining themselves, to compare products and prices, or they can use tools like I AM⁴ to put their personal preferences in control over external stimuli and inputs. And finally, consumers will adopt multiple (electronic) personalities to allow them to continue to browse, buy and consume anonymously where they prefer to do so.