

## Intelligent Data Mining and Personalisation for Customer Relationship Management

Kok Wai Wong<sup>1</sup>, Chun Che Fung<sup>2</sup>, Tamás Gedeon<sup>3</sup>, and Douglas Chai<sup>4</sup>

<sup>1</sup>School of Computer Engineering, Nanyang Technological University, Nanyang Ave, Singapore 639798

<sup>2</sup>School of Information Technology, Murdoch University, Murdoch, WA 6150, Australia

<sup>3</sup>Department of Computer Science, The Australian National University, Acton, ACT 0200, Australia

<sup>4</sup>School of Engineering and Mathematics, Edith Cowan University, Joondalup, WA 6027, Australia

### Abstract

Customer Relationship Management (CRM) initiatives have gained much attention in recent years. With the aid of data mining technology, businesses can formulate specific strategies for different customer bases more precisely. Additionally, personalisation is another important issue in CRM – especially when a company has a huge product range. This paper presents a case model and investigates the use of computational intelligent techniques for CRM. These techniques allow the complex functions of relating customer behaviour to internal business processes to be learned more easily and the industry expertise and experience from business managers to be integrated into the modelling framework directly. Hence they can be used in the CRM framework to enhance the creation of targeted strategies for specific customer bases.

### 1 Introduction

Computational intelligence is an emerging technology used to solve many problems of high complexity. This paper examines the use of intelligent techniques in Customer Relationship Management (CRM) to create specific targeted strategies for different customer bases, and also to perform personalisation. In [1] and [2], intelligent techniques have already been explored for the purpose of CRM, this paper focuses more on using these CI technologies in electronic CRM which could improve the E-business side of a company. With the popularisation of E-commerce in recent years, web services that can provide better CRM for the company through electronic means are also gaining importance.

When businesses first used computers to store data, data mining technology started to evolve as a new approach in assisting navigation through the database. Its purpose is mainly helping businesses to focus on important and useful information by extracting the hidden predictive information from large databases [3]. Basically, data mining techniques perform these predictive features based on modelling. The previously known situations are used to build the model, and then the model is applied to other unknown situations. The objective of the data mining technique in such application is typically for the purpose of automating the process of detecting relevant patterns in a large database.

From data mining, we may find out the interest of a small group of customers. However, each customer is different in many aspects. Even in the sub small group, customers need to be identified as individuals through some personalisation agents. With the use of electronics business model, this can be done more effectively [4]. When using Internet to provide personalisation for CRM, it is basically providing a customised content to individual based on their interests and behaviour. This involves three main logical stages: (1) to collect customer information and purchasing behaviour, (2) to perform analysis of the customer information, and (3) to present relevant information that the customer will be interested.

In recent years, CRM initiatives have gained much attention. Although CRM involves technology, the important success factor involves customer-focused strategy. A study in the US indicates that many businesses are dissatisfied with their current CRM initiative involving multi-function CRM software [5]. The report cited the lack of customer focus and less adaptation to their unique requirements as reasons for their dissatisfaction. This paper posits that with the aid of data mining or targeted techniques, businesses can formulate specific customer-focused strategies more easily and scientifically and therefore be more satisfied with their CRM initiatives. As the term CRM suggests, there are three main areas of focus: (1) the Customer, (2) the Relationship, and (3) the Management of the relationship [6]. Loyal customers are valuable assets for a business. Studies have shown that a 5% increase in customer retention can lead to a 25-100% increase in customer value [7]. Relationships with customers are driven primarily by the value the customers perceive from the relationship. Heskett *et al.* [8] have offered a model of customer value as shown below:

$$value_i = \frac{results + process\ quality}{price + acquisition\ cost} \quad (1)$$

From the above model, we can see that value for customer *i* has several components. The first component, *results*, refers to the idea that customers buy results and not products or services. When a product or service enhances the desired results, it increases customer value. Similarly, *process quality* also increases customer value. The way in which a service is delivered

is often as important as the result itself. *Price* is also a component of customer value but not the only component. The costs of acquiring a product or service can sometimes overshadow the price itself. Data mining technology can enhance the understanding of different components of customer value as well as the needs and background of the customer. Different components of customer value provide opportunities for enhancement and management of the relationship with individual customers. From (1), we can see that value is defined at the individual level (hence the subscript  $i$ ). Therefore it is important to identify the components of value that are unique to each customer or customer base in order to create unique value propositions to that customer base and manage those relationships appropriately.

One method of identifying components of value and opportunities for relationship enhancement is the identification of customer segments. In marketing research, this is normally known as market segmentation [9]. Market segmentation breaks down a heterogeneous market into a number of smaller homogeneous markets where special treatment and care can be used to address a more precise satisfaction factor of the customer needs. Segmentation can normally be classified as a-priori and post-hoc approaches [10]. In CRM, the market to be segmented is the customer base.

In order to understand a customer, the process has to start by analysing all the relevant data belonging to the customer. Thus, data mining techniques can be considered as the intelligence behind a successful CRM strategy [6]. This technology is to transform data into useful information for a business to focus on its customers. There are basically two main types of data mining: descriptive and predictive. Descriptive data mining generates information about the data so that we can realise some interesting underlying information. Predictive data mining makes use of past patterns and information in predicting what a customer will buy in the future.

There are five main steps in the process of implementing a successful data mining solution for CRM [11]: setting goals, data collection, data preparation, analysis and prediction, and measurement and feedback. During goals setting, identification of the market segmentation model is important. Proper goals under each segment will permit a more precise address of issues such as retention, risk avoidance as well as possible cross selling. In data collection and preparation, it is important to address issues like feature selection, parameter identification and handling of missing data. When building analysis and prediction models, different methods may have to be used in each segment in order to meet the intended goals. A crucial point in gaining business confidence in establishing a model is to avoid a total "black box" method that eliminates the contributions of the expert in the business.

In this paper, data mining and personalisation tool using intelligent techniques are proposed in part of the CRM building block.

## 2 CRM Model

In this section, we propose a CRM model that can facilitate the use of E-commerce. After the proposed CRM model has been discussed, intelligent data mining techniques that can aid the formulation of the CRM strategy will be presented in the subsequent sections. As CRM model for this application could be a very complex model, the strategy behind may be manifold. Consequently, we limit the scope of this paper by presenting only the modules that serve the purpose of providing personalisation using E-commerce. The purpose of the model is to assist the organisation or business in developing some forms of CRM strategy that could effectively provide personalised service. We emphasise that this model is strictly developed for this research, and the authors have no intention of copying any existing model.

The building blocks of the CRM model provide a framework to deliver value to the customer and profitability to the company. This framework starts with differentiating customers along two dimensions, by the customers' value to the firm, and by their wants and needs. The first dimension focuses on customer valuation measurement by understanding the value a customer represents to the company. This measurement identifies revenue growth potential from repeat purchases, referrals, and expanding scope of business. The second dimension looks at the customer preference to determine what are the wants and needs of customers, i.e., the customer value proposition. This measurement examines the results that a customer is trying to achieve, and determines whether the company can bundle products to deliver the results. This could also be able to translate to some level of personalisation in the E-commerce environment.

Our proposed model aims to integrate two main aspects of the CRM, which are intelligent data mining and personalisation, in order to provide personalised services by recommending products or services which the customer will most likely be interested in. The main objectives of the integration can be divided into three main categories: market segmentation and product matching, customer preference learning, and product filtering.

In a CRM data mining model, the first step is to collect knowledge from the past and existing customers so that their wants and needs can be studied. This first stage of the CRM model is to perform segmentation in the large pool of customer profiles within the data warehouse. This data warehouse normally contains information about the customer, information gathered via purchase made, surveys, as well as other point of contacts information. After segmentation of the large data

warehouse, it is possible to identify a few smaller subsets. Ideally, after the initial segmentation, the subsets data pool will be more homogeneous within each segment, which is more suitable for target analysis. Beside the data warehouse that contains customer profiles, there will be another data warehouse that is specially used to store information about products or services provided by the company. Some automatically generated rules and may be human interaction and analysis is required at this stage to perform some product matching. Normally, in this step, the human analyst will also take away any input variables that are irrelevant to the objectives. For example, input variables such as Post Codes, Address etc may not be very useful for some subgroup of the data mart. Figure 1 gives a graphical illustration of this data mining stage of the CRM model.

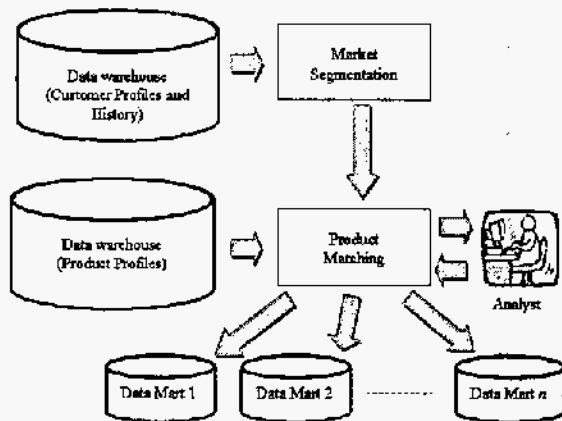


Figure 1: Data mining for the CRM model.

After the data warehouses have been merged and broken down into data marts, these data marts will be used in the personalisation module of the CRM as shown in Figure 2. When customers first register, based on the information collected from the registration process, the customers will be pre-classified according to the knowledge generated from the previous data mining module. Every time when a particular customer logs in with their account, the relevant data marts will be identified as the information that the personalised system will work on. This will speed up the personalisation process by working on a subgroup of the mass amount of the data.

Beside this pre-classification of the customers when they registered, a customer preference learning module will learn the customers' profile online. Every time, when a customer interacts with the E-business site, the preference gathers by his behaviour will be logged. From this logbook, the customer preferences like price, quality, and brand can be learned. At the end, rankings on the products will be performed and save in a database. The information on the database is then used to verify the pre-classification stage, and if necessary re-

classification is performed to extract the more relevant data marts for better personalisation.

When a customer is interested in purchasing a product or searching for a product, an intelligent agent will searches through the data marts based on the customer preference learned. Based on the products attributes and customer's preference, a filter is used to select those products that the customer are most likely to be interested. It is then presented to the customer in such a way that the higher ranked products will appear to the customer first.

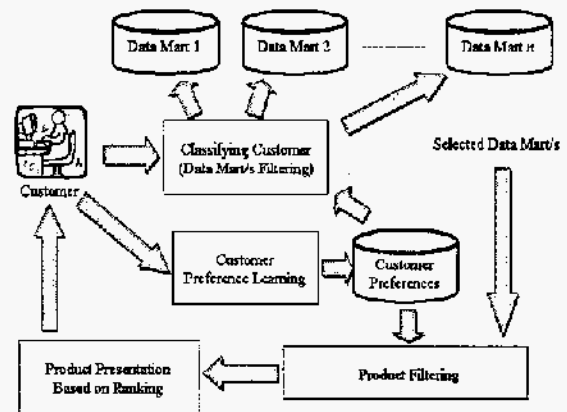


Figure 2: Customer personalisation for the CRM model.

### 3 Computational Intelligence

#### 3.1 Artificial Neural Networks

In the last decade, Artificial Neural Networks (ANNs) have emerged as a useful option for inferential data analysis and solving complex data analysis problem. The observation sample that is used to derive the predictive model is known as training data in ANN development. The independent variables, or the predictor variables, are known as the input variables and the dependent variables, or the responses, are known as the output variables.

In supervised learning, an ANN makes use of the input variables and their corresponding output variables to learn the relationship between them. Once the relationship is found, the trained ANN is then used to predict new output variables given new input data set. Back propagation Neural Network (BPNN) as shown in Figure 3 is the most widely used neural network system and the most well known supervised learning technique [12]. Back propagation is a systematic method for training multilayer ANN. It has been implemented and applied successfully to various problems. A basic BPNN consists of an input, an output and one or more hidden layers. Each layer is made up of a number of neurons that are connected to all the neurons in the next layers. However, the output layer will only generate the results of the network.

The objective of training BPNN is to adjust the weights so that application of a set of inputs will produce the desired set of outputs. A training set containing a number of desired input and output pairs is used. The input set is presented to the input layer of BPNN. A calculation is carried out to obtain the output set by proceeding from the input layer to the output layer. After this stage, feed forward propagation is done. At the output, the total error (the sum of the squares of the errors on each output cell) is calculated and then back propagated through the network. The total error,  $E$ , can be calculated using:

$$E = \sum_{k=1}^K \left( \frac{1}{2} \sum_{i=1}^{N_i} [T_i(k) - O_i^L(k)]^2 \right) \quad (2)$$

where  $K$  is the number of patterns,  $L$  is the layer number,  $T$  is the expect target, and  $O$  is the actual output.

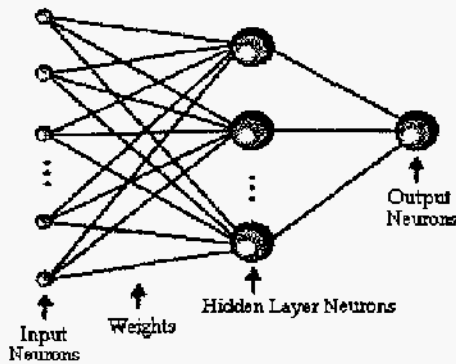


Figure 3: Back propagation neural network.

A modification of each connection weight is done and new total error is calculated. This back propagated process is repeated until the total error value is below some particular threshold. At this stage, the network is considered trained. After the BPNN has been trained, it can then be applied to predict other cases.

For unsupervised learning, an ANN will only make use of the input variables and attempts to arrange them in a way that is meaningful to the analyst. Self-Organising Map (SOM) is a popular unsupervised neural network technique mainly because it is a fast, easy and reliable unsupervised clustering technique [13]. SOM is designed to simulate the organisation found in various brain structures and is related to brain maps. Its main feature is the ability to visualise high dimensional input spaces onto a smaller dimensional display, usually two-dimensional as shown in Figure 4. In this discussion, only two-dimensional arrays will be of interest. Let the input data space  $\mathcal{R}^n$  be mapped by the SOM onto a two-dimensional array with  $i$  nodes. Associated with each  $i$  node is a parametric reference vector  $m_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{in}]^T \in \mathcal{R}^n$ , where  $\mu_{ij}$  is the connection weight between node  $i$  and input  $j$ . Therefore, the input data space  $\mathcal{R}^n$

consisting of input vectors  $X = [x_1, x_2, \dots, x_n]^T$ , i.e.  $X \in \mathcal{R}^n$ , can be visualised as being connected to all nodes in parallel via a scalar weight  $\mu_{ij}$ . The aim of the learning is to map all the  $n$  input vectors  $X_n$  onto  $m_i$  by adjusting weights  $\mu_{ij}$  such that the SOM gives the best match response locations.

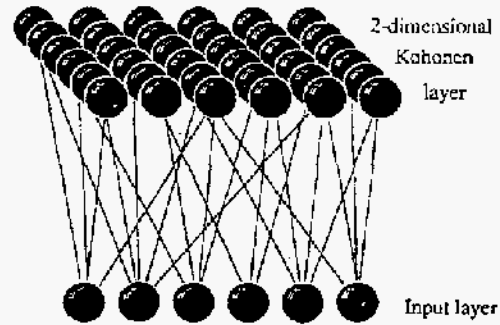


Figure 4: Visualisation of self organising map.

It can also be said that SOM is a nonlinear projection of the probability density function  $p(X)$  of the high dimensional input vector space onto the two-dimensional display map. Normally, to find the best matching node  $i$ , the input vector  $X$  is compared to all reference vector  $m_i$  by searching for the smallest Euclidean distance:

$$\|X - m_i\|, \text{ indexed by } c, \quad (3)$$

$$\text{i.e., } \|x - m_c\| = \min_i \|x - m_i\|$$

During the learning process, the node that best matches the input vector  $X$  is allowed to learn. Those nodes that are close to the node up to a certain distance will also be allowed to learn. The learning process is expressed as:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[X(t) - m_i(t)] \quad (4)$$

where  $t$  is a discrete time coordinate, and  $h_{ci}(t)$  is the neighbourhood function.

After the learning process has converged, the map will display the probability density function  $p(X)$  that best describes all the input vectors. At the end of the learning process, an average quantisation error of the map will be generated to indicate how well the map matches the entire input vectors  $X_n$ . The average quantisation error is defined as:

$$E = \int \|X - m_c\|^2 p(X) dX \quad (5)$$

### 3.2 Fuzzy Theory

Fuzzy theory works on the basis derived from fuzzy sets [15]. A fuzzy set allows for the degree of membership of an item in a set to be any real number between 0 and 1, this allows human observations, expressions and expertise to be modelled. The membership function of a fuzzy set  $A$  is denoted by:

$$A: X \rightarrow [0,1] \quad (6)$$

Once the fuzzy sets have been defined, it is possible to use them in constructing rules for fuzzy expert systems and in performing fuzzy inference. Fuzzy systems can produce more accurate results based on the basic idea of defuzzification. A defuzzification technique is used to calculate the conclusion by evaluating the degree of matches from the observation that triggered one or several rules in the model. This will lead to a better result by handling the fuzziness in the decision making process. Thus, fuzzy technique can improve on statistical prediction in certain cases.

Fuzzy sets allow human expertise and decisions to be modelled more closely. It is suggested that Fuzzy sets will play an important role in the CRM model. In this paper, with the availability of vast amounts of data in each subset, it will be useful to extract knowledge from the data directly. This has the advantage of discovery of new knowledge or relations underlying the data. In extracting fuzzy rules from the data, the first step is to translate all the available data into linguistic fuzzy rules using linguistic labels. The following algorithm outlines the steps in extracting the fuzzy linguistic rules from the available data. For  $k$  inputs, the given input-output data pairs with  $n$  patterns:

$$\begin{aligned} & (x_1^1, \dots, x_k^1; y^1) \\ & (x_1^2, \dots, x_k^2; y^2) \\ & \vdots \\ & (x_1^n, \dots, x_k^n; y^n) \end{aligned} \quad (7)$$

The number of linguistics terms  $T$  and the distribution of data in the regions of the whole domain are first determined. For ease of interpretation and computational simplicity, the shape of the membership function used in this algorithm is triangular. In this case, we will obtain for every  $x \in X$ ,

$$A_t \in F(x) \rightarrow [0,1] \text{ for all } t \in T \quad (8)$$

After the fuzzy regions and membership functions have been set up, the available data set will be mapped accordingly. If the value cuts on more than one membership function, the one with the maximum membership grade will be assigned to the value:

$$R_n \Rightarrow [x_1^n(A_{i1}, \max), \dots, x_k^n(A_{ik}, \max); y^n(B_t, \max)] \quad (9)$$

After all the data sets have been assigned with a fuzzy linguistic label, Mamdani type fuzzy rules are then formed and centroid defuzzification is used. After fuzzy rules have been generated from each data point, repeated rules are removed. In the event that there are repeated fuzzy rules, the number of repetitions of the fuzzy rules and the firing strengths of the rules will be examined to resolve conflicts.

## 4 Intelligent CRM

In this section we present the computational intelligence methods used in blocks shown in Figures 1 and 2. There are basically two stages in the CRM model that we have proposed in Section 2; there are the data mining stage and the customer personalisation stage. In the data mining stage, the intelligent techniques used are similar to the one published earlier in [15]. Here we summarise the intelligent technique used as follows.

For market segmentation, SOM is used to cluster the data warehouse. After which, the products are mapped to the SOM 2-dimensional map. The assigned products to the customer profiles clusters generated from SOM will be divided into data mart which consists of a subset of information from the data warehouse. Fuzzy rules extraction will be used to extract fuzzy relationships between the customer profiles and the products profiles. With this rules, a human analyst can determine the best way to generate the data mart.

For the second stage of the CRM model, the customer personalisation is performed through three main blocks, classifying customer, customer preference learning, and products filtering.

When a customer first registers into the system, the customer will be pre-classified based on the information collected from the customer during the registration process. This is done by mapping the customer information to the SOM generated from the data warehouse during the data mining stage. After which, the corresponding data marts that this customer will most probably fit in will be selected for personalisation. When the customer logs in, the customer profile and history learned during the further interaction will be used to perform the classification of the customer. Thus, this should improve the accuracies of selecting the appropriate data marts when dealing with the customer.

In this CRM model, to reduce the complexity and to improve interactivity, customers' ranking towards products is the only content interested in the profile. In order to get the ranking information, customer's online behaviours need to be collected and learned. The recording method similar to those proposed in [16] is used to collect customer's behaviour. The information of the customer recorded for our CRM model are customer login ID, products ID, the number of times the products are shown to this customer [*times*], and accumulated customer ranking on the products [*acc\_rank*].

The number of times that a product is shown to the customer is important to calculate the ranking for the customer. A product may be shown to a customer multiple times during one session or in multiple sessions. Whenever a product is presented to a customer, the customer may show interest in it or ignore it. Based on the customer's behaviour, the rank is calculated in the following ways:

- When a new customer behaviour record is created, both the customer ID and product ID are set. Both *times* and *rank* are initialised to zero.
- If the customer clicks on the product to see its detailed description, the *rank* for that product will increase by 1.
- If the customer adds the product to shopping cart, the *rank* will increase by 2.
- If the customer makes a purchase on the product, the *rank* will increase by 2 again.
- After the customer preferences are collected, the customer ranking can be calculated using the formula:

$$acc\_rank = rank / times \quad (10)$$

After the customer preference records are stored, a BPNN can be used to learn the behaviour of the customer. In this BPNN delegated to a customer, we will have four inputs namely Price, Quality, Guarantee, and Brand, and the output will be ranked. As we want to give a more precise ranking, the rank generated will be in real number rather than integer, as it will assist in generating ranks of the products when preparing for product filtering. The filtering of the products is mainly based on the trained BPNN for a specific customer. The inferred rank will determine which products should be filtered. A higher rank value implies a higher chance for the product to be preferred by the customer. When presenting to the customer, the products are sorted according to their ranking values. The products with low ranking values and below the threshold will be filtered out. Figure 5 gives a cross plot of the predicted ranks as compared to the actual ranks collected for a specific customer. It has shown that this intelligent CRM model could be an alternative of implement CRM strategy suitable for E-business environment.

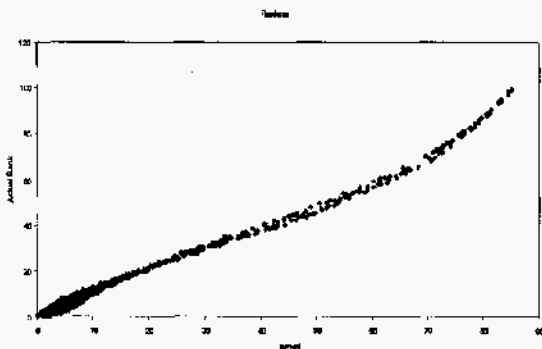


Figure 5: Cross plot of predicted ranking versus actual ranking.

## 5 Conclusion

This paper has examined the possibility of using intelligent techniques in the CRM model specially designed by the authors for use in E-business context. The paper also highlighted two areas in a typical CRM model where the use of intelligent techniques can

improve the whole process. The advantage of using intelligent techniques in CRM is that the business analyst can perform the CRM strategy better and at the same time gain in-depth understanding into the CRM model. With the understanding of the model, the analyst can modify and add-on knowledge and experience into the model. Furthermore, fuzzy theory can handle uncertainties in the data more efficiently than traditional data mining techniques.

## References

- [1] K.W. Wong, B. Griner, W.R. Dillon, and T.D. Gedcon, "Intelligent Data Mining for Customer Relationship Management," *Australian J. of Intelligent Information Processing Systems*, vol. 7, no. 1/2, pp. 64-71, 2001.
- [2] K.W. Wong, "Data Mining Using Fuzzy Theory for Customer Relationship Management," *Proc. Western Australian Workshop on Information Systems Research*, Australia, Nov. 2001.
- [3] A. Berson, S. Smith, and K. Thearling, *Building Data Mining Applications for CRM*, McGraw-Hill, 2000, pp. 123-199.
- [4] S. Murugesan and A. Ramanathan "Web Personalisation - An Overview," *Active Media Technology*, pp. 65-76, 2001.
- [5] CRMGuru.com, High-Yield Marketing & Mungen Research Associates, "Multi-Function CRM Software: How good is it?" 2000.
- [6] M. Nykamp, *The Customer Differential: The Complete Guide to Implementing Customer Relationship Management*, Amacom, New York, 2001.
- [7] R. Riechheld, "The Loyalty Effect: The Hidden Force Behind Growth, Profits and Lasting Value," Harvard University Press, Boston, 1996.
- [8] J. Heskett, W.E. Sasser, and L. Schlesinger, "The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction, and Value," The Free Press, New York, 1997.
- [9] M. Wedel and W.A. Kamakura *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer Academic Publishers, Massachusetts, 1998.
- [10] Y. Wind, "Issues and Advances in Segmentation Research," *J. of Marketing*, vol. 15, pp. 317-337, 1978.
- [11] T.J. Siragusa, "Implementing Data Mining for Better CRM," *Customer Interaction Solutions*, vol. 19 no. 11, pp. 38, 2001
- [12] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representation by Error Propagation" in *Parallel Distributed Processing*, vol. 1, MIT Press, Massachusetts, pp. 318-362, 1986.
- [13] T. Kohonen, *Self-Organising Map*, Springer-Verlag, 1995.
- [14] L.A. Zadeh, "Outline a new approach to the analysis of complex systems," *IEEE Trans. on Sys. Man and Cybernetics*, vol. 1, pp. 28-44, 1973.
- [15] K.W. Wong, C.C. Fung, and T.D. Gedcon "Data Mining Using Neural Fuzzy for Student Relationship Management," *Proc. Int. Conf. of Soft Computing and Intelligent Systems* Tsukuba, Japan, Oct. 2002.
- [16] Z. Wang and C.K. Siew "A Novel Approach to Personalization Product Filtering in Electronic Commerce", *Australian J. of Intelligent Information Processing Systems*, vol 7, no 3-4, pp. 81-95, 2001.