ESANN'2007 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2007, d-side publi., ISBN 2-930307-07-2.

Identification of churn routes in the Brazilian telecommunications market

David L. García, Alfredo Vellido and Ángela Nebot

Department of Computing Languages and Systems. Technical University of Catalonia.
C. Jordi Girona, Barcelona, 08034 Spain.
e-mail: {dgarcia, avellido, angela}@lsi.upc.edu

Abstract. The globalization and deregulation of business environments are rapidly shifting the competitive challenges that telecommunications service providers face. As a result, many of these companies are focusing on the preservation of existing customers and the limitation of customer attrition damages. In this brief paper, we investigate the existence of abandonment routes in the Brazilian telecommunications market, according to the customers' service consumption pattern. A non-linear latent variable model of the manifold learning family is used to segment and visualize the data, as well as to identify typical churn routes.

1 Introduction

In the prevailing competitive market environment, anticipating the customer's intention to abandon facilitates the launching of retention-focussed actions. Nowadays, this is one of the main preoccupations among service providing companies, as it represents a clear element of competitive advantage. In this context, data mining techniques can assist churn (customer attrition) management processes [1]. This paper focuses on the treatment aspect of the efficient churn management model proposed in Fig. 1 [2] and, in particular, on discovering customer abandonment routes on the basis of the customer service consumption pattern. A non-linear latent variable model, Generative Topographic Mapping (GTM, [3]) is used to segment and visualize data from the Brazilian telecommunications market. From these results, typical customer churn routes are investigated.

2 Problem Description

Identifying a customer's intention to abandon their current service provider with sufficient anticipation has become one of the main focal points of marketing studies in recent years [1, 4]. The main trends in predicting customer behaviour and, in particular, customer abandonment (churn), are based on the direct identification of the possible churner through the use of historical variables relating to customer behaviour. Data mining methods have seldom been employed to this end [1].

In this context, we propose an indirect and exploratory approach to the prediction of customer abandonment, based on the visualization of customer data -consisting of their consumption patterns- on a two-dimensional representation map. Customers' routes across the map, with a focus on *departure gates* of customer abandonment, will be investigated. Changes in customer value across periods will also be explored. The identification of customer migration routes, despite being a novel approach in the field of churn prediction, has already been studied, using SOM-based methods, in the areas of company financial status analysis and in bankruptcy prediction [5, 6].

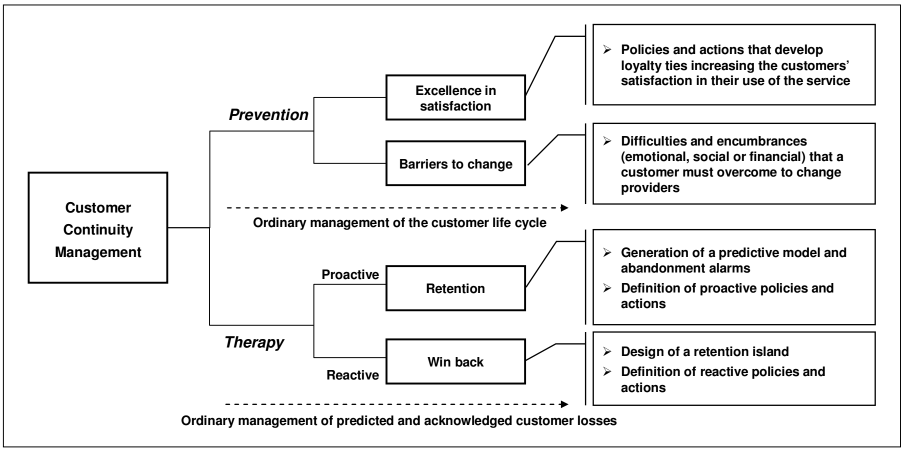

Fig. 1: Conceptual model of customer continuity management.

Our approach is based on two basic hypotheses: a) Different patterns of service consumption, regarding the type of communications established, correspond to different levels of predisposition to abandon; b) Different migration routes between time periods are likely to exist and be identifiable in the representation map, both negative: towards lower customer value and, eventually, service abandonment; and positive: towards higher customer value areas.

3 Telecommunications Customer Attrition Survey Data

For the purpose of this study, a total of 60,596 small and medium-sized companies, clients of the main fixed telephony company in São Paulo (Brazil), were considered. Measurements were taken over two consecutive non-vacational periods. *Period 1* (P1) from June to December 2003 and *Period 2* (P2) from March to August 2004.

The unreleased information available in this database (supplied by the company) corresponds to the averages of several clients' consumption variables over each period (minutes and number of calls). In addition, for profiling the segmentation results, the following information was taken into account: commercial margin, value-added services (VASs) on portfolio, length of time as client, CNAE code (National Classification of Economic Activities) and number of employees in the company.

4 Experiments

As described in section 2, in this study we aim to explore the relationship between the telecommunications consumption behaviour of small and medium-sized companies in the Brazilian market and the propensity to abandon the service provider. The experiments carried out to this end can be summarised as follows:

In a first stage, the GTM model (see description in the next section) was fitted to the variables listed in Table 1 for period P1. Each point of the GTM map (corresponding to a *micro-cluster*) was then characterized using the profiling variables described in section 3. The resulting *micro-segments* were aggregated into *macro-segments* using the *k*-means algorithm in order to make the segmentation results more easily actionable from a business point of view.

p_ll_nor	Percentage of normal timetable calls
p_ll_dif	Percentage of differentiated calls
p_ll_mis	Percentage of mixed calls
p_ll_red	Percentage of reduced timetable calls
p_ll_sup	Percentage of super reduced timetable calls
p_ll_ln	Percentage of local calls or Internet
p_ll_ita	Percentage of intrastate calls
p_ll_ite	Percentage of interstate calls
p_ll_int	Percentage of international calls
p_ll_mov	Percentage of calls to mobile
p_lc_loc	Percentage of reversed charge local calls
p_lc_ita	Percentage of reversed charge intrastate calls
p_lc_ite	Percentage of reversed charge interstate calls
p_lc_mov	Percentage of reversed charge calls to mobile

Table 1: Description of the variables used to develop the GTM model.

In a second stage, the same steps were followed for the data corresponding to P2, with the exception of clients who abandoned their service provider between periods. Comparing the position of the clients over the GTM maps in each of the periods studied, we aim to identify the micro-segments and the areas on the GTM map associated to higher probabilities of churn (the *departure gates*) , as well as those associated to loss or increase of value between periods. Customer movement between areas should provide the basis for the development of a churn warning system. The existing literature does not show comparable customer churn prediction examples and no other algorithms have been applied to analyse churn according to this approach.

4.1 A Brief Description of GTM

The GTM [3] is a probabilistic alternative to the heuristic Self-Organizing Maps (SOM) for clustering and visualization. However, unlike the standard SOM, GTM defines an explicit probability density model of the data. Its probabilistic foundations allow it to be expanded to tackle problems such as missing data imputation, outlier detection, feature relevance determination and time series analysis, amongst others. The GTM is a non-linear latent variable model of the manifold learning family that defines a mapping from a low dimensional latent space, which acts as a visualization canvas, onto the observed data space. The mapping is carried through by basis functions generating a (mixture) density distribution. The functional form of this mapping for each variable d can be expressed as:

$$y_d(\mathbf{u},\mathbf{W}) = \sum_{m}^{M} \phi_m(\mathbf{u}) w_{md} \tag{1}$$

where Φ are basis functions $\Phi(\mathbf{u}) = (\phi_1(\mathbf{u}),\ldots,\phi_M(\mathbf{u}))$ that introduce the non-linearity in the mapping; $\mathbf{W}$ is the matrix of adaptive weights w_{md} that defines the mapping; and $\mathbf{u}$ is a point in latent space. In order to provide an alternative to the visualization space defined by the characteristic SOM lattice, the latent space of the GTM is discretized as a regular grid of K latent points $\mathbf{u}_k$. The mixture density for a data point $\mathbf{x}$, given Gaussian basis functions, can be written as:

$$p(\mathbf{x} \mid \mathbf{W},\beta) = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{\beta}{2\pi}\right)^{D/2} \exp\left\{-\beta/2 \left\| y_k - \mathbf{x} \right\|^2\right\} \tag{2}$$

where the D elements of $\mathbf{y}$ are given by Eq. 1. This density allows for the definition of a model likelihood, and the well-known Expectation-Maximization algorithm can be used to obtain the Maximum Likelihood estimates of the adaptive parameters ($\mathbf{W}$ and β) of the model. See [3] for details on these calculations.

Each of the latent space points $\mathbf{u}_k$ can be considered by itself as a cluster representative (of a cluster containing the subset of service clients assigned to it). For simplicity, for the GTM we use a cluster assignment method akin to that of SOM, which is based on a winner-takes-all strategy: each data observation (client) is assigned to the location in the latent space (cluster) where the mode of the corresponding posterior distribution is highest, i.e. $\mathbf{u}_n^{\mathbf{mode}} = \arg\max_{\mathbf{u}_k} r_{kn}$, where r_{kn} is the probability of client n belonging to cluster (*micro-segment*) k, and it is obtained as part of the EM algorithm.

4.2 Results and Discussion

A GTM with an 8×8 cluster grid structure was implemented, following a standard PCA-based procedure for its initialization [3]. The 64 clusters, or *micro-segments*, corresponding to the data of the P1 period, are visualized in Fig.2; the relative size of each cluster is directly proportional to the number of clients assigned to it with the criterion described in the previous section. The 64 GTM cluster prototypes were further grouped using *k*-means in order to find segments of increased market actionability. Four segments were identified this way as: 1.- *Local companies* (83% local calls); 2.- *Commercial companies with mobile employees* (30% calls to mobiles); 3.- *National companies* (43% national calls); and 4.- *Professionals* (15% calls outside working hours). These segments are visualized in the right-hand side map of Fig.2.

We are interested in the migration routes of clients across segments over the periods P1 and P2. These results are summarized in Table 2, where the percentages of clients moving (or not) from one segment to another are displayed. Segment 2

(*Commercial companies with mobile employees*) are by far the most volatile (highest value of *relocation* in Table 2), whereas segment 3 (*National companies*) are the most resilient to change (lowest value of *relocation* in Table 2). Besides, this volatility is directly proportional to their corresponding percentage of churn.

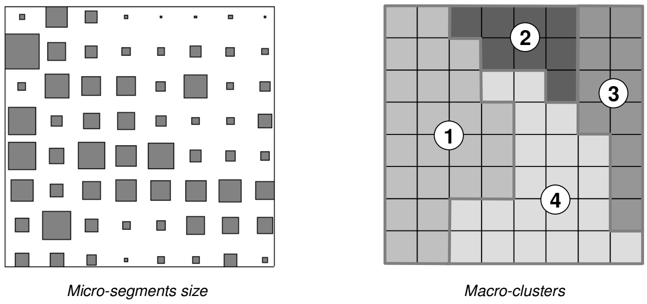

Fig. 2: GTM maps of clusters and segments for data of the P1 period.

P1 \ P2	1	2	3	4	Churn	Relocation
1	70,4%	3,9%	0,3%	14,6%	10,8%	18,8%
2	9,4%	49,0%	10,9%	16,2%	14,5%	36,5%
3	0,8%	5,5%	73,7%	10,2%	9,8%	16,5%
4	16,3%	8,2%	8,1%	56,1%	11,3%	32,7%

Table 2: Segment mobility and percentage of churn over the P1-P2 periods.

We now explore churn in more detail through the identification of the *departure gates* for abandonment: GTM (micro) segments for which the probability of churning between periods is higher. The probabilistic nature of the GTM allows the definition of a principled churn index $P_{k,churn} = \left(\sum_{\{N\}} r_{kn} \right)^{-1} \sum_{\{n*\}} r_{kn}$ for a micro-segment k, where $\{n*\}$ is the set of churning clients; N is the total number of clients; and r_{kn} has been defined in section 4.1.

This index results are colour-coded and displayed in Fig.3 (left): white corresponds to the highest churn index, whereas black corresponds to the lowest. This is accompanied, on the right-hand side map, by a display of the values of commercial margin for each cluster. This last map reveals differentiated areas of commercial margin (clusters of highest margin in white: R$740; and of lowest margin in black: R$124). Using Fig.2, we see that high margins correspond mostly to small-medium sized micro-segments, mainly belonging to the *National companies* market segment. Similarly, areas with neatly different churn index can be clearly identified. The high values of the churn index mostly appear associated to small micro-segments (the *departure gates*) belonging to two market segments: *Professionals* and *Commercial companies with mobile employees*. Interestingly, they are also micro-segments with low commercial margin, which is indicative of the commercial health of this particular operator's client portfolio.

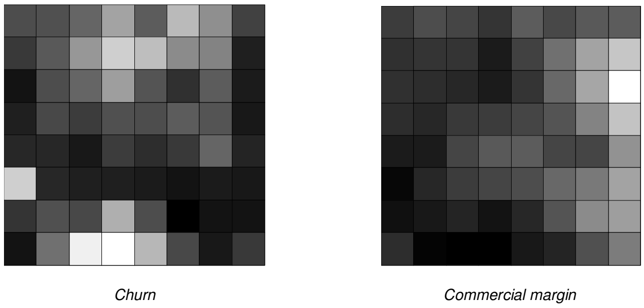

Fig 3: Churn index and commercial margin for each micro-segment.

In general practice, the combination of the churn index, the commercial margin, and the typical migration routes over periods can provide telecommunication operators with a decision support system that warned about profitable clients moving towards departure gates with higher probability of abandonment.

5 Conclusions

The use of effective models for the exploration of customer churn becomes an important challenge for service providers. In this brief study, we have used a probabilistic model of the manifold learning family to cluster and visualize the clients of a major Brazilian telecommunications provider. Different areas where the risk of abandonment is higher, or *departure gates*, have been identified on the basis of the service consumption patterns. The migration routes between market segments have also been explored. This model should provide the basis for the development of a churn warning system.

References

[1] J. Hadden, A. Tiwari, R. Roy and D. Ruta, Computer assisted customer churn management: State-of-the-art and future trends, *Computers & Operations Research*, Elsevier, in press.

[2] A. Vellido, T. A. Etchells, D. L. García and Á. Nebot, Describing customer loyalty to Spanish petrol stations through rule extraction, LNCS, Vol.4224, Springer, pages 970-977, 2006.

[3] C.M. Bishop, M. Svensén and C.K.I. Williams, GTM: The Generative Topographic Mapping, *Neural Computation*, 10(1):215-234, MIT Press, 1998.

[4] Hung S-Y, Yen DC and Wang H-Y. Applying data mining to telecom churn management, *Expert Systems with Applications*, 31:515–524, Elsevier, 2006.

[5] K. Kiviluoto, Predicting bankruptcies with the self –organizing map, *Neurocomputing,* 21:191-201, Elsevier, 1998.

[6] A. Vellido, P.J.G. Lisboa and J. Vaughan, Neural networks in business: a survey of applications (1992-1998), *Expert Systems with Applications,* 17:51-70, Elsevier, 1999.