

Telco churn analysis classification using a wavelet and RBF approach

Â. M. Cister¹ & N. F. F. Ebecken²

¹*FACC - COPPE/UFRJ, Brazil*

²*COPPE/UFRJ- Federal University of Rio de Janeiro, Brazil*

Abstract

The goal of this paper is to present a quantitative study about Telco churn analysis and the inherent problems of its application and computer processing.

An independent customer characterizes the approach and the pre-processing step uses the wavelet transform concept; the pattern classification problem is solved by an artificial neural network (RBF- Radial Basis Function).

The signal representation considers different levels of resolution (multi-resolution). This idea is frequently used in image/video processing where general image versions with a low resolution are stored in different levels of resolution. The concepts of multi-resolution have been implemented by a digital filter bank, Hi-pass and Low-pass. Each sub-band will give the best resolution, for each spectral region. The idea is to look for a spectral transient.

A scalable Daubechies' function db 2 and db 3 were tested for this study and the results were used as input for the RBF neural network. Daubechies' function has shown important properties as for instance, orthogonality and scalability, which are not presented by Morlet's function.

Some results of the experiment were presented to point out the performance of the wavelet pre-processing approach.

Keywords: churn analysis, wavelet transform, RBF neural networks.

1 Introduction

There are many literary works that are aimed at the new political-economic order imposed by diverse countries in the last decade of the 20th century. The present theoretical study is not meant to incite criticism against, or in favor, of the new economic order. The truth is that companies from different segments, mainly



from the area of services, have found themselves obligated to modify their treatment of the new kind of customer that has emerged from the impact of global economics.

The new customer has become very demanding, as a result of the fact that salesmen of products and services, from many different parts of the world, have come knocking at the door offering their truly marvelous items that can make life easier and simpler.

In this case, new companies, influenced by the new winds of the global economy, see themselves obligated to change their behavior towards these new and demanding consumers.

2 The CRM system

In spite of the fact that the implantation of the CRM [1] system is not different from the implantation of other far-reaching corporative systems, it also can fail or end up with serious problems.

These problems can be caused by an excessive focus on technology and applications, and not enough attention on the details. In the case of CRM initiatives the most important detail is the customer.

The fact is, what is dealt with in the majority of CRM projects has little to do with the customer. The company is always thinking, initially, in reducing costs or increasing loyalty of the customer, because in this way it is better and cheaper for the company. The company does not perceive that the necessary focus should be on what would be the benefits for the customer, and that this would consequently lead to benefits for the company.

Benefits, that would facilitate business with the company, such as: access to better quality information, facility to buy from and communicate with the company, economy in terms of time and money, elimination of solicitations of redundant data, etc.

CRM has to do with more profound concept: each customer is distinct, different, and should be treated in a different way.

The implantation of CRM projects should always have as the most important point, the changes in the processes that involve the customer, or in other words, to see the company through the critical eyes of the best customers. Without this, you are not implanting CRM, but just another software package.

3 Retaining the customer

In research conducted in 1987, by "The Forum Group", an American consulting company specialized in administration, employees from American and Canadian companies declared that attracting customers was considered a priority where they worked, twice as important as keeping those who already were with the company. This could be considered a very costly error.

According to Kotler [2], recommends that a great deal of care should be taken to avoid losing customers. The loss of customers causes great damage to organizations and demands measures to reduce the index of desertion.



The motive for the loss of consumers needs to be investigated, seeking to correct the manner of proceeding in order to stop this abandonment.

There is not much that can be done with the customers who have already changed to another company; however those who have stopped buying because they are dissatisfied might be able to be recuperated, if a solution to their problem can be attained through dialogue.

Presently, the most successful companies strive with all their energy to maintain their customers. Many markets have achieved their maturity and there are not many new customers to be found. Competition is always growing and the costs to attract new customers are constantly climbing.

It is very important to accompany those who are already customers. Reducing the churn of customers, even if it is by 5% a year, stimulates a profit growth of up to 85%.

In the insurance area, according to the Life Insurance Marketing and Research Association, those insured the longest with a certain company, have a 50% more chance of keeping their accounts active than those who have recently opened accounts with the same company. Brokers who maintain contact with their best customers have more success and less likelihood of failing in their attempts at cross selling. In this way other companies can also benefit.

The older the relationship with the customer, the better. Data from the consulting firm Bain & Co., with its headquarters in Boston, indicate that a profit of 10%, resulting from one year with a customer, can grow up to 60% when the relationship is prolonged for more than seven years.

Customers loyal for many years are more apt to recommend the organization that serves them, to family and friends. At this time the world offers many options to customers. Because of this, to observe and follow the process of retention is fundamental, and represents a large piece of the pie in the process of surviving as a company. The management of this relationship is a process based on three pillars:

- Knowledge – refers to the quantity and depth of information that one should collect about the customer (CRM-Analytic).
- Technology to collect and analyze information that permits agility in making decisions (CRM-Operational).
- Construction of the relationship – through effective communication (CRM-Collaborative).

The first two pillars can be resolved easily. It is enough for the entrepreneur to surround himself with suppliers, who are known to be competent in the area, track down references from companies, bibliographies, and participate in congresses, courses, and events. The bottleneck appears in the third pillar, the constructing of relationships. The company needs trained and motivated people, because a relationship assumes, naturally, personal contact between the company's representative and the customer.

The team needs to be prepared to establish relationships with customers based on trust, because the results obtained from the actions set in motion by the relationships, are a major factor in the survival of the company.

Finally, there is a direct relationship between fidelity, the average retention of customers and their life time value (LTV). To be able to calculate the value of the LTV of a customer requires an estimate of the cash flow that the customer will provide if they reach a medium level of loyalty.

4 Telco churn analysis

Studying about churn is extremely important, because the large part of organizations, telecommunication, and computers, all fall under the same responsibility. And each day that passes makes it more difficult to separate them. So what is churn?

In the computer world, some authors to express the accelerated renovation of products use churn or the known programmed obsolescence. But churn, object of this study, which has its basis in marketing, treats the loss of customers by a company to the competition, or in other words, the measuring of the disloyalty of customers. This is the concept that is used most among telecommunication companies or in other companies that offer services. In Brazil, with the deregulation that has occurred in telecommunications, this has been one of the sectors most affected by this phenomenon.

Other sectors of economics also have had to administer the churn. Banks and administrators of credit cards are two well-known examples. For consumers, this vast gamut of options signifies greater freedom of choice. Hence, whether it is a home telephone, cellular phone, some kind of communication of data, or Internet, anyone can change from one company to another easily.

At this time, the analysis of the churn has become a challenge for a wide variety of businesses, although few recognize its real implications. The Telecommunications sector is a pioneer in the search for tools and systems that help in this analysis.

How is the analysis of a churn accomplished? Basically, the analysis of a churn [3] is the analysis of historic data that permit those who are analyzing to foresee which customers could leave a company, with the taking of action by the company, to avoid this churn. In the majority of cases, the study is concentrated on the behavior and/or profile of the customer.

When we say that the cost of the churn of a company varies between 5% and 10% annually, we really wish to say that this company is losing from 5% to 10% of the customers that are found in its database. This signifies that it is losing customers that do or have done business with it, and for some reason, became dissatisfied and left, disappearing from their data base.

This raises two points. The first point is that an analysis of the churn has to be detailed enough to discover the variables that make these dissatisfied customers leave and to discover why they are dissatisfied.

The problem of this analysis is that the volume of data to be analyzed is large, and normally the data bank of these companies is very extensive, which makes quick and efficient access to the information and its consolidation more difficult.

Along with the problem of the volume of data, there is also the problem of the lack of knowledge, theoretically as well as technically, on the part of the companies. The majority of the businesses do not possess the basic knowledge in data exploration and statistic analysis, prerequisites important in foreseeing the *churn*. Once the variables are discovered, the companies can take the best action to avoid the churn.

Another point is that many of these customers are customers who have the potential to spend a great deal with the company. Thus, these are valuable customers that the company is losing and that probably will never return. A detailed analysis of the churn permits the businesses to determine those customers which present behavior that indicates that could turn into churners and that these churners are customers who deserve special attention. This information contributes in helping the companies to better their programs that deal with the loyalty of customers and reduce the costs of marketing, especially since there has been a fall in the growth of the market and greater competition.

With the heating up of competition, the telephone operatives search in the CRM solutions, ways to fatten up the customers' wallets and augment their incomes.

With the opening up and deregulation of the telecommunication companies, worldwide, this is one of the sectors that experience most, the phenomenon of the churn. Other sectors such as economics learned a long time ago how to administer the churn. Banks and the administrators of credit cards are two well-known examples.

The great migration to the free Internet companies is a recent example of this facility. Many people use a web-mail to choose and change Internet companies whenever they want, and as the behavior of people and corporations is different, the suppliers will draw up specific strategies to deal with this. They will have to reach higher levels of segmentation and differentiation in order to obtain their objective of loyalty. Customers should prepare themselves to be the targets of marketing plans that are very bold.

These days the word is relationship, or in other words, gets to know your users and have alternative solutions for them. In the past, attending the customer could be summarized simply as welcoming him. This type of initiative, whose objective is to perfect the channels of attendance, represents only one of the elements that compose CRM politics. The first race of the telephone companies involved the centers of attendance based on the demands of ANATEL (govern agency for telephony regulation), and because of this, people started to associate CRM only to call center. Initiating a project touching on the channels of attendance is great, but if there is no notion of the whole, it will become inefficient.

5 The wavelet transform

When the basis function needs to expand a signal, it uses a time-frequency resolution, which is the main idea in the Fourier transform [4]. For example, the function used in the analysis presents infinite precision when we talk about



frequency localization. However, it is very poor in terms of time because of its endless extension.

The wavelet can be defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right); \quad a, b \in \mathbf{R} \neq 0 \quad (1)$$

Where ψ is called the mother wavelet and has time frequency localization, ψ is generally a band pass filter.

So, as written by Master in his book [5] we can say: “the scale is proportional to the duration of the basis function used in signal dilation”. Mathematically, this is:

I. Daubechies’ [4] presented the discrete form of a and b :

$$\bullet \quad a = a_0^m \quad (2)$$

$$\bullet \quad b = nb_a a_o^m \quad (3)$$

The signal representation can be considered as different levels of resolution (multiresolution). This idea is frequently used in image/video processing where general image versions with low resolution are stored at different levels of resolution and the algorithm is responsible for composing the picture. As a practical example, we can site the image database in WWW (world wide web), where initially the low-resolution versions are transmitted with a delay. Later details are transmitted only if it is really necessary.

The concept of multi-resolution has been implemented by digital filter banks, Hi-pass, and Low-pass. Each sub-band will give the best spectral resolution for each spectral region. This process can be repeated infinitely because of the complexity of the implementation

6 RBF Neural Networks

The functional networks of the RBF (Radial Basis Function) are supervised in this way as MLP networks, and both are considered as universal stochastic approximation, possessing algorithm backpropagation, yet they still differ in regards to their architecture and characteristics.

The construction of a RBF, involves three different layers. The first layer, the entry layer, is formed by sensory units, or nodes, which connect the network to its environment. The second layer is the only hidden layer in the network. It is the second layer, which is responsible for the non-linear application of the entry space to the hidden space, which almost always possesses a high dimensionality. The third layer is the exit layer and is completely linear. The third layer is responsible for the response for the network to the pattern inserted in the entry layer. According to Cover (*apud* [5]) “... a problem of classification of patterns

laid out in a space of high dimensionality has a greater probability of being linearly separated than in a space of low dimensionality”, thus the reason and mathematical justification for a strategy of non-linear transformation to follow a linear transformation. One other very important point in the use of a RBF is that the fact that you are working with a large dimension in the hidden layer and this allows a better approximation for mapping the delicate entry-exit. According to Mhaskar (*apud* [5]) the greater the dimension of the hidden space, the more precise will be the approximation of the entry pattern to the classification. One cannot forget that the Cover Theorem, which treats the separability of patterns, rigorously defines the RBF types of networks mathematically. The Cover Theorem states that for a RBF to realize a complex task of the classification of patterns, the problem is to transform non-linearly, from a high dimensional space, until the final approximation of the entry pattern.

7 The experiment

The experiment aimed at classifying/predicting customer who presents the inclination and/or tendency to escape from the database of a cell phone company. The population is composed of a determined segment of customer. These customers represent the profile of an artificial person, possessing 4 cell phone lines on an average, and are considered heavy users, that is, customers who retain a high index of minutes each month.

The experiment begins with the ideal search for a sample. For this, several rehearsals were made with the technique of the bootstrap sampling, which searched among chosen variables, the ideal average, or in other words, that which reflected the same parameters of the population. Considered as the values of the parameters for the sampling [8], were the confidence interval of 95.5% and 2% of sampling error. These values were distributed proportionally for both extracts. In this case, the probability of finding one active inside a group was 93% against 7% chance of finding an inactive. The rigidity of the choice of the sample error is due to the problem of eliminating outliers and/or fraudulent behaviors from the base.

The choice of variables was at the suggestion of a senior specialist who possessed a high level of knowledge about the market and more than 15 years of experience in the area of telephony. The population for this segment is more than 10,000 customers, being however, for the means of calculating samples, an infinite population.

The variables had by definition the following description:

V1: Additional in roaming – The client pays when receiving the call; *V2*: International call in roaming; *V3*: From cell phone to home phone; *V4*: From cell phone to cell phone same operator; *V5*: From cell phone to cell phone with other operators; *V6*: From cell phone to home phone using long distance (DDD); *V7*: From cell phone to another cell phone same operator in roaming; *V8*: From cell phone to cell phone cooperative; *V9*: From cell phone to cell phone with other operators in roaming; *V10*: Diverse services 0800, 0300, information

The size of the sampling was from 817 registers (CI 95.5% and error 1.89%). After data cleaning the number of registers fell to 722 registers (CI 95.5%, error 2.0%).

The strategy of choosing from the registers was to choose 12 complete months, in this way taking a photograph of the status, active or inactive, in the 13th month. In this way, the mass of data was chosen between August 2004, and July 2005, and taking the status, active or inactive in the month of August 2005.

For this study, it was thought for the good of all to take out all the outlier's data, especially because it dealt with an extremely unbalanced problem. The proportion between active and inactive customer is very large, as can be observed. In this way, the right to remove data outside of the pattern was respected, following statistical techniques. But, the rule was not the same for both extracts. The extract of active customer was penalized more. The cut that was considered was to remove registers, which were above 2 deviations from the pattern, in relation to the maximum values. For this extract the monthly averages were considered. Consequently, there could exist an outlier for the month of September 2004 and no other outlier for the other months. Exactly in this way, this register was removed. The criteria for removing a register, signifies that there exists an outlier at least considered being the extract of active customer. As regards the minimum values it was decided to remove 15% of the registers from the 1st quartile. This decision was arrived at due to the behavior of the data. The minimum values were better "behaved" (homogenous), than the superior values.

8 Results

The criteria in use for the evaluations of the active customer, is in function of the inactive extract. In other words, what was chosen for the inactive customer will be used for the active ones as well.

The networks performance and the results corresponding to db 2 and db 3 were grouped in Tables 1 to 6.

Table 1: Characteristics of networks with filtration db 2.

	Profile	Train Error	Select Error	Test Error	Training Members
1	RBF 115:115-14-1:1	0,1199	0,1248	0,1337	KM,KN,PI
2	RBF 115:115-9-1:1	0,1091	0,1236	0,1706	KM,KN,PI
3	RBF 115:115-12-1:1	0,1034	0,1203	0,1585	KM,KN,PI
4	RBF 115:115-18-1:1	0,1266	0,1117	0,1562	KM,KN,PI
5	RBF 115:115-19-1:1	0,1365	0,1026	0,1003	KM,KN,PI
6	RBF 115:115-15-1:1	0,1571	0,0951	0,1594	KM,KN,PI
7	RBF 115:115-19-1:1	0,1449	0,0865	0,0827	KM,KN,PI

Table 2: Hits filtration db 2 (Inactive).

RBF1	RBF2	RBF3	RBF4	RBF5	RBF6	RBF7	Hits	CNT
62	60	65	71	72	74	76	76	82
							Hits	0,9268

Table 3: Hits filtration db 2 (actives).

RBF1	RBF2	RBF3	RBF4	RBF5	RBF6	RBF7	Hits	CNT
641	643	643	639	641	641	631	640	643
							Hits	0,9953

Table 4: Characteristics of networks with filtration db 3.

	Profile	Train Error	Select Error	Test Error	Training Members
1	RBF 57:57-48-1:1	0,1992	0,2608	0,2946	KM,KN,PI
2	RBF 57:57-62-1:1	0,2418	0,2549	0,2778	SS,KN,PI
3	RBF 57:57-58-1:1	0,2229	0,2482	0,2958	SS,KN,PI
4	RBF 57:57-61-1:1	0,2243	0,2420	0,2608	SS,KN,PI
5	RBF 57:57-63-1:1	0,2582	0,2405	0,2752	SS,KN,PI
6	RBF 57:57-61-1:1	0,2243	0,2420	0,2608	SS,KN,PI
7	RBF 57:57-63-1:1	0,2582	0,2405	0,2752	SS,KN,PI

Table 5: Hits filtration db 3 (Inactive).

RBF1	RBF2	RBF3	RBF4	RBF5	RBF6	RBF7	Hits	CNT
58	57	54	60	58	51	50	58	82
							Hits	0,7073

Table 6: Hits filtration db 3 (active).

RBF1	RBF2	RBF3	RBF4	RBF5	RBF6	RBF7	Hits	CNT
602	609	605	603	607	599	594	611	643
							Hits	0,9502

9 Conclusion

In response to the filtration, it is shown that filter db 2, which has the best resolution in the search for transients, responded according to expectations. Even



though it is not possible, with the naked eye, to see these transients, they were detected by the filtration, which gave an excellent input generating a neural network, which responded very well to the inactive customer. Without doubt the best hit rate was presented by db 2 pre-processing and looking for the best neural network result. To achieve the better market model it is necessary the use of an ensemble of models or the use of cross validation techniques.

The concepts utilized for the development of this research are in consonance with the results of several other studies. The scientific methodology used for the study was shown to be efficient in relation to the organization and cognitive gathering of facts for production and development of a research basis.

As a suggestion for future study in this area we can extend the scope of this work, using the concept information entropy. This concept, within the information theory has extreme importance to the statistic control and manipulation of probabilistic density functions. Such a concept is shown to be more powerful each day, to aid in the analysis and validation of decision makers in the international market.

References

- [1] Pepper, Don; Rogers Martha – One have one manager: Real – World lessons in CRM – The New York: Currency/Doubleday (1999).
- [2] Kotler, P.; Armstrong, G. – Marketing Principles - & ed. Prentice Hall, 1997.
- [3] Mattison Rob – Telecon churn Management: The golden opportunity – APG Publishing.
- [4] Meyer, Yves – Wavelets and operators – Cambridge University Press 1992.
- [5] Master, Timothy – Signal and image processing with neural networks – Wiley 1994.
- [6] Haykin, Simon – Neural Networks: A comprehensive foundation – Macmillan 1991, p. 284.
- [7] Haykin, Simon – Neural Networks: A comprehensive foundation – Macmillan 1991, p. 285.
- [8] Kish, Leslie – Survey Sampling – Wiley inter science 1995.