

View Point



Churn Prediction

Approach to retain Profitable Customers

Sivaprakasam S.R.

Abstract

For Communication Service Providers (CSPs), cost of acquiring new customers is significantly higher than retaining existing customers. It is vital for CSPs to identify profitable customers who are likely to discontinue services with them in the near future. This identification process has to be very accurate as customer retention efforts will also incur significant cost.

This paper discusses all these issues in detail with a special focus on present market trends, predicting the churn model, and application of this model on real-time data. This paper is the outcome of lengthy discussions as well as analysis of real time statistics.

Executive Summary

Due to increase in the competitive intensity within the telecom industry, the subjects of 'customer acquisition, churn, retention, loyalty' have gained special focus. With an increase in number of players in this industry, customers now have the option of choosing among various Communication Service Providers (CSPs). Attractive promotional campaigns as well as lucrative offerings play a major part in luring a customer. From a CSP's perspective, the major challenge lies in identifying and retaining the 'profitable' subscribers by offering them attractive service packages, proactive customer service and high quality Value Added Services (VAS).

Telecommunication Market Trend

Every year all major CSPs incur huge costs due to customer churn. When customer decides to discontinue with the existing CSP, the service provider loses the acquisition cost as well as the potential revenue. Service providers have to keep an eye on churn trends in the market in order to take proactive steps for customer retention.

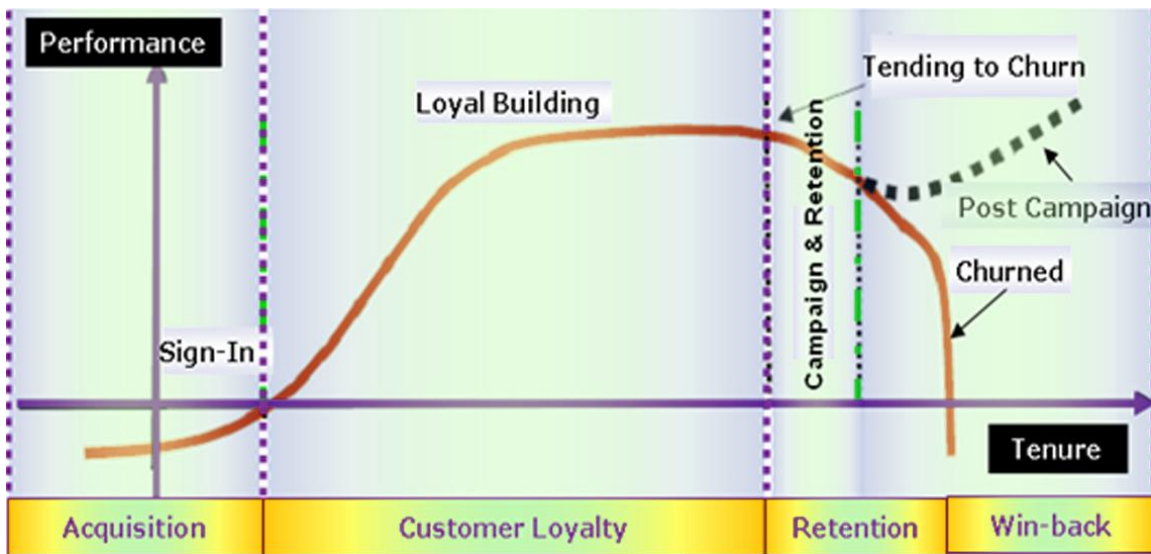
Based on previous engagements with a US wireless service provider, the annual churn rate is approximately twenty seven percent i.e. two percent of the subscriber base dissociates with the service provider every month. Typically, CSPs have to spend about \$400 to \$500 for acquiring a new subscriber. And, to make this new subscriber connection a profitable one, the service provider should generate a minimum bill of \$50 per month for a period of 10 months.

Surprisingly, in-depth churn analysis of a US wireless carrier revealed that twenty percent of their yearly revenue could have been saved by retaining five percent of the subscribers who chose to dissociate themselves with this CSP. An interesting fact is that nearly one third of US wireless subscribers chose a different service provider at the maturity of their 1 year tenure.

To address these issues, service providers have been taking keen interest in retaining existing subscribers. This can be done by investing in the 'best subscribers' i.e. tailoring the service offerings/ packages as per the value they contribute to the carrier.

Figure 1 clearly illustrates the phase where subscribers tend to churn. During this phase, the CSPs need an effective campaign for customer retention. The performance index of customer life cycle is found to be peak when customer shows loyalty towards the service provider.

Figure 1: Performance Index of Customer Life Cycle



In today's highly competitive environment, CSPs have designed attractive incentives for acquiring new customers.

Recently, a US based service provider revealed, in a monthly report on 'customer churn', the top reasons for customer attrition. The main causes for customer dissociating with the service provider can be classified in the following categories - payment defaults, quality of service, lack of service coverage area, disputes regarding billing as well as service disputes.

Top 15 reasons for customer churn are presented in Figure 2.

Figure 2: Customer Churn Analysis

Churn Reason Description	No. of Churns	Churn %
Non-Pay Terminations	2,331	32.35%
Can't / Won't Provide Service	1,809	25.10%
Portability Termination	1,240	17.21%
Rates / Better Offer by Competitor	420	5.83%
Relocating to another Service Area	330	4.08%
14 Days or Less Termination	229	3.18%
Subscription Fraud	195	2.71%
Dissatisfied with Cell	161	2.43%
Termination within 48 Hours	146	2.23%
Coverage	76	1.05%
Equipment Reasons	52	0.73%
Mobile Broken	38	0.56%
Customer Expired	35	0.50%
Termination by reason of hurricane disaster	31	0.45%
Employee / Agent / Demo / Testing Churns	29	0.42%
Number of Churns in Top 15 Reasons:	7,122	98.83%
Number of Churns in Left Over Reasons:	84	1.17%
Total Number of Churns:	7,206	100%

To address these areas of concern, service providers have to focus towards making their operations more effective and having a campaign aimed at customer retention.

It is apparent from the above diagram that the top two reasons are initiated by the CSP and thus, can be called 'voluntary churns'. All the other 'churns' are initiated by the customer and can be called 'involuntary churns'.

After analyzing this report, the CSP addressed the voluntary churn (which comprised half of the total churn) by revamping their billing process. Here, the CSP took proactive steps for collecting monthly bill and did multiple follow-ups to make sure that there is not much outstanding amount.

Simultaneously, the CSP also did a feasibility analysis pertaining to wide coverage for better service. Under involuntary churns, mobile number portability was the top reason for customers switching their service provider. The CSP took corrective action by implementing the necessary service assurance workflow amendments for smooth portability.

From this churn analysis, it is obvious that any CSP can retain their high value customers for improving the organization's profitability.

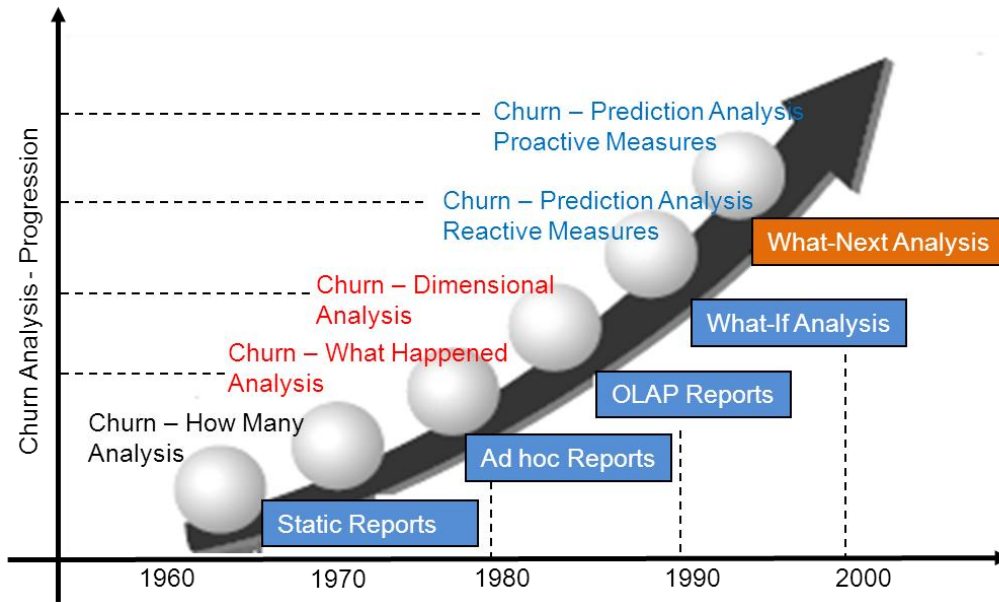
Churn Prediction: Is it vital?

CSPs usually have a database on customer related information like their orders, provisioning details, usage, billing, service desk, network fault management, etc.

Over the past four decades, service providers have been performing extensive analysis on customer churn. Sophisticated tools like historical analysis, trend analysis, dimensional analysis and What-if analysis are used for this analysis. The latest tool being used is the what-if analysis. Based on the recent historical data, this tool has the ability to list customers who are likely to change service providers.

Figure 3 illustrates how the 'customer churn' analysis technique has evolved over a period of time. To conduct such analysis, a service provider would need a model that has the ability to not only segregate subscribers in various segments but also calculate the probability of churn associated with these subscribers. This model needs to consider various pieces of information including customer demographics, call usage patterns, payment behavior, Quality of Service (QoS), customer service feedback, etc. This model, when applied for all subscribers, will give the service provider an indication of customer's inclination for churn as well as the percentage of churn risk.

Figure 3: Trend of Churn Prediction Analysis



Subsequent sections will explain the data mining to perform customer churn analysis.

Mining: How is it done?

Data mining involves extensive analysis of recent historical data to predict customer churn trends. Detecting the churn pattern enables better understanding of a particular line of business. This, in turn, helps the organization in optimizing the business processes and thus, improving operational efficiency.

Data mining process uses statistical algorithms for processing large volumes of data; the outcome of this process is a predictive model. This predictive model can be used for decision making in key business areas like Product Sales, Customer Churn, Weather forecasting, etc. Ideally, for this model to work, CSPs would require large data sets for better predictions and lower estimation errors/ variance.

Figure 4: Typical Data Mining Analysis

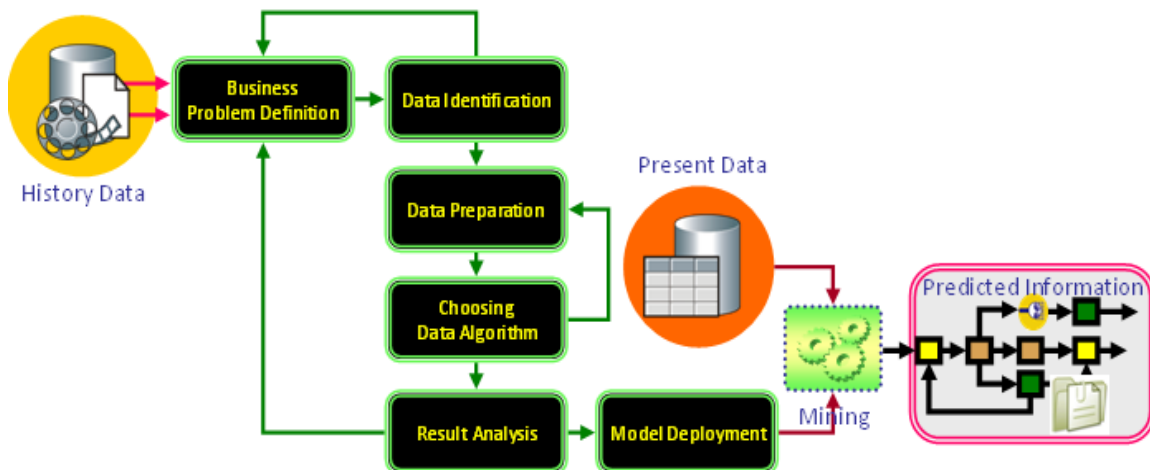


Figure 4 depicts the different stages involved in the data mining process. As a result of mining process, the fine-tuned model is prepared based on the historical data and analyzed against the present data. This analysis is then used to predict the pattern for better decision making. The following subsections will explain the process taken care in each stage of mining process.

Data Identification: This data is identified based on the extensive discussions with business users, key decision makers and data stewards to predict the 'customer churn' behavior. The accuracy of the output is proportional to the number of attributes involved in the modeling. For example, the following information can be included in this analysis: Customer, subscriber, products, services, price plans, activation, billing, payment method, call usage, order registrations, service activation, product catalog, product category, age, gender, income range, QoS, education, occupation, etc.

Data Preparation: Data should be cleansed and transformed¹ form as dimensions, facts and demographics, then loaded into data warehouse.

Data Reduction: For downsizing of data, the technique of Data reduction or pruning is used. This step is extremely vital because the data source may or may not be specific to the data mining process. Accuracy and efficiency of queries degrade as the number of dimensions increase. This hampers the speed of learning and predictive accuracy. For example, dimensions like education, hobbies of customer can be excluded from the scope of analysis as they have minimal effect for predicting churn related behavior.

Mining Algorithm Identification: Data mining tools are bundled with industry standard statistical algorithms. It is analyst's responsibility to identify a suitable mining algorithm for building an accurate churn model pertinent to a specific business area. Each algorithm generates at least one result; results are bound to vary with different algorithms.

The classification algorithm is usually applied for churn prediction. This algorithm assigns cases to different classes and is used to predict the 'discrete' variables. This algorithm is supported by Naïve Bayes, Decision trees as well as Neural Network techniques.

Analyze the Results: The churn model generated in the previous step is tested against various predictor² values of training records. This validates the model using sample data which is subjected to further analysis. By comparing the actual results (obtained from the training sample) with the predicted result (from the mining model), the predictor values are changed gradually to meet the actual results.

Deploy the Model: This stage deals with implementation of the mining process; thereby resulting in a list of 'profitable' customers who are most likely to churn. These models are a mathematical representation of data which provides not only the customer characteristics but also their behavior.

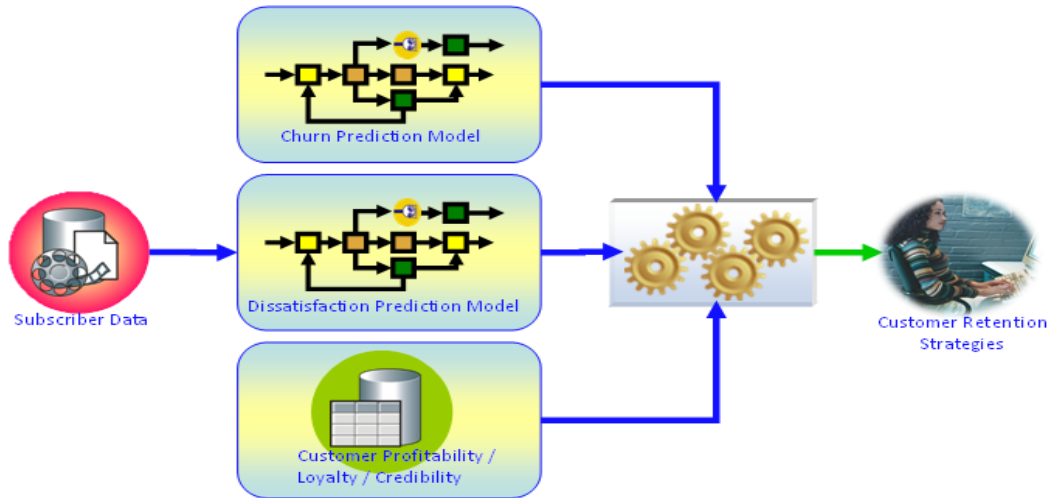
¹ **Transformation:** It is the process of standardizing, formatting, cleansing, removing duplicates and inconsistencies, aggregating, augmenting and make sure the data is concise.

² **Predictor:** Field in a database that could be used to build a predictive model to predict the values in another field. aka Variable, Independent variable, Dimension, or Feature.

Churn Model: Way to Implement

Commercially off-the-shelf (COTS) data mining tools that merely predict the obvious churns are available in the market. For e.g., predicting a subscriber may churn, who has not paid bills in over three months.

Figure 5: Prediction Model & Retention Strategy



Customized data mining solutions are necessary to provide more accurate results to meet the exact business needs of a CSP. The list of subscribers predicted with high risk of churn has to be screened for their credibility, profitability, loyalty, satisfaction index, etc. Based on this basic analysis, appropriate campaign aimed at customer retention should be designed. The process of retaining the high value customers is depicted in the figure 5 above. As per the diagram, the subscriber data is fed against the past churn prediction model, dissatisfaction prediction model, and customer loyalty, profitability, and creditability data to recommend the retention strategies for high value customers, Depending upon the campaign's efficiency and incentives offered, value of Confusion Matrix will vary. An efficient campaign combined with excellent incentives will result in 'low' Confusion Matrix. This means, the predicted and actual values of both 'possible churns' and 'possible non-churns' will be very close. This concept is illustrated in Figure 6.

Figure 6: Confusion Matrix on Churn Prediction

	Predicted to Churn	Predicted to Continue
Actual Churn	A	B
Not Churn	C	D
-----	Threshold	Observation
A	Churn as per prediction	Need to Tune Further Campaign Efficiency
B	Predicted to continue but Churned	
C	Predicted to churn but continuing	
D	Predicted to continue	
Confusion Matrix		Accuracy = (A + D) / (A + B + C + D)

Conclusion

In the telecom industry, customer satisfaction is a vital factor for customer retention. Owing to this fact, the customer services have made rapid strides in both quality as well as quantity. In order to retain 'profitable' customers, the predicted churn model has to be periodically revisited with varying predictor variables and differing training sets. This will help the CSP obtain accurate results.

Churn prediction is a cross-functional activity which involves marketing, sales, finance, service provisioning, customer service, etc. Carriers need integrated effort between all these business areas for improving long-term profitability. Churn prediction exercise also improves other business processes including

- Automating decisions in evaluating credit scores during new subscriber activation process.
- Renegotiating more favorable contracts with commercial customers by comparing initial contracts with actual usage patterns.
- Evaluating the effectiveness of marketing campaigns, right service for the subscriber at right first time

About the Author

Sivaprakasam S.R.

Sivaprakasam S.R. is a Principal Architect and mentors the Database and Business Intelligence track in the Communication, Media, and Entertainment business unit. His interests include Enterprise Data Modeling, Enterprise Data Integration, Enterprise Data Warehousing, Enterprise Data Mining, Enterprise Data Quality Management and Semantic Data Integration. He can be reached at sivaprakasam_s@infosys.com.

References:

A Data mining glossary. (n.d.). Retrieved from <http://www.thearling.com/glossary.htm>

Data mining. (n.d.). Retrieved from http://en.wikipedia.org/wiki/Data_mining



Infosys among the world's top 50 most respected companies

Reputation Institute's Global Reputation Pulse 2009 ranked Infosys among the world's top 50 most respected companies.



About Infosys

Many of the world's most successful organizations rely on Infosys to deliver measurable business value. Infosys provides business consulting, technology, engineering and outsourcing services to help clients in over 30 countries build tomorrow's enterprise.

For more information about Infosys (NASDAQ:INFY), visit www.infosys.com.

Global presence

Americas	Asia Pacific	Europe	Middle East and Africa
Brazil	Australia	Belgium	Kingdom of Saudi Arabia
Nova Lima	Brisbane	Brussels	Riyadh
Canada	Melbourne	Czech Republic	Mauritius
Calgary	Perth	Brno	Reduit
Montreal	Sydney	Prague	UAE
Toronto	China	Denmark	Dubai
Mexico	Shanghai	Copenhagen	Sharjah
Monterrey	Hangzhou	Finland	
United States	Hong Kong	Helsinki	
Atlanta	Central	France	
Bellevue	India	Paris	
Bentonville	Bangalore	Germany	
Bridgewater	Bhubaneshwar	Frankfurt	
Charlotte	Chandigarh	Stuttgart	
Detroit	Chennai	Walldorf	
Fremont	Gurgaon	Ireland	
Hartford	Hyderabad	Dublin	
Houston	Jaipur	Italy	
Lake Forest	Mangalore	Milano	
Lisle	Mumbai	Norway	
New York	Mysore	Oslo	
Phoenix	New Delhi	Poland	
Plano	Pune	Lodz	
Quincy	Thiruvananthapuram	Spain	
Reston	Japan	Madrid	
	Tokyo	Sweden	
	New Zealand	Stockholm	
	Wellington	Switzerland	
	Philippines	Geneva	
	Metro Manila	Zurich	
	Singapore	The Netherlands	
		Amsterdam	
		United Kingdom (UK)	
		London	

For more information, contact askus@infosys.com

www.infosys.com

© 2011 Infosys Technologies Limited, Bangalore, India. Infosys believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of the trademarks and product names of other companies mentioned in this document.