Forum Statistika dan Komputasi : *Indonesian Journal of Statistics*
ISSN : 0853-8115
Vol. 18 No.1, April 2013, p: 1-10

available online at:
journal.ipb.ac.id/index.php/statistika

Survival Analysis Of Customer In Postpaid Telecommunication Industry

Authors :

Doni Suhartono[1], Asep Saefuddin[2], I Made Sumertajaya[2]

[1]DEKA Research, Jakarta, Indonesia
[2]Department of Statistics, Bogor Agricultural University, Indonesia
E-mail : [1] *doni.pro@gmail.com*

ABSTRACT

Currently, the business competition in mobile telecommunication industry among providers in Indonesia is tighter and it has given rise to a phenomenon of customer defection which has serious consequences for the business performance. In the current circumstances, customers are faced numerous options to be selected that probably cause them at risk to get churn. Therefore, it becomes one of the challenges encountered by Division of Loyalty and Retention to makes the efforts of decreasing customer defection. So that it is important conducting a model of churn practically applied to predict tendency of customer churn and also recognizing the prognostic factors influence customer churn. Survival analysis modelling, such as Cox's proportional hazard model, was very successful in previous research, which investigatedthe relationship between survival time and possible prognostic factors. Based on the research, Cox's proportional hazard model of customer lifetime is effective to distinguish relative risk between churn customers and others, and also between which loyal customers and with other short time customers with their significant prognostic factors. Afterwards the simulation of the survival probability estimated over time with particular possible combination of the most significant characteristics affecting tendency of churn, are able to predict such information of lifetime to churn event and compare the survival performance of one another. Finally, the results of this research is able to yield simple, helpful and applicable results as the principle of taking decission for optimizing their customer retention and/or treatment resources in their customer retention efforts for the company.

Key words : Churn, Cox's proportional hazard model, customer retention, survival analysis and telecommunication industry.

INTRODUCTION

Background

In the telecommunications industry, customers are able to choose among multiple service providers and actively exercise their rights of switching from one service provider to another. In this fiercely competitive market, customers demand tailored products and better services at less prices, while service providers constantly focus on acquisitions as their business goals. However, there is a fact that the telecommunications industry experiences an average of 30-35 percent annual churn rate and it costs 5-10 times more to recruit a new customer than to retain an existing one, now customer retention has become even more important than customer acquisition (Lu 2002).

With the telecommunications market getting more and more mature, telecommunication companies do not satisfy themselves with predicting customer churn; they instead start viewing customers in terms of customer lifetime. Not only do the telecommunications companies distinguish between which customers stay and which ones churn, they also distinguish between which customers stay longer and which ones stay shorter. Modelling of customer lifetime is therefore developed to satisfy telecommunication companies that need to evaluate their customer lifetime.

Conventional statistical methods such as; Logistic Regression, Decision Tree, Artificial Neural Network, Naïve Bayessian Data Mining , Hierarchical neuro-fuzzy systems, Lift Curve, etc. are very successful in predicting customer survival/churn. There were related works that have already been performed in the similar field. Setiabudi (2009) and Oktarina (2010) also adopted the steps of the analyst by modeling the potential churner with binary logistic regression. However, these methods could not predict when will customers churn, or how long will the customers stayand identify prognostic factor related to survival time.

Based on the present challanges, survival analysis wasinitialy designed to handle censored cases in survival data. Proportional Hazard model or well-known by Cox Regression is able to investigate the relationship between survival time and possible variables (Lee 1992). Typically, not only do we predict the timing of customer churn, we also analyze how covariates impact the occurrence and the timing of customer churn.

Widyaningsih (2002) also stated that survival analysis modelling by Cox Regression is better than Logistic Regression. Logistic Regression ussualy conducts data elimination for incomplete data (censored cases) therefore the result of this analysis is less appropriate with the actual cisrcumtances and can not handle the censored data. Therefore it is an efficient and powerful tool to predict or to describe customer survival/churn.

Objectives

This research is aimed to

1. Identify the prognostic factors affecting customer churn related to survival time.
2. Estimate the relative risk of customer churn.
3. Estimate the survival probability of customer with their significant characteristics.
4. Predict the lifetime of churn customers .

LITERATURE REVIEW

Churn

In the telecommunications industry, the broad definition of churn is the action that a customer's telecommunications service is canceled. This includes both service-provider initiated churn and customer initiated churn (Lus 2001).

Churn means attrition or degradation of the number of subscribers. Yang and Chiu (2006) mentioned that there are three terms of churn, those are:

1. Involuntary churn: provider stops the service
2. Unavoidable churn: customer died or migrate to unserviceable areas.
3. Voluntary churn: customer changes the service into more prospective providers or resigning from them definitely.

Unavoidable and voluntary churn are classified into one group of our attention since it is hardly distinguished.

Survival Analysis

Survival analysis is the phrase used to describe the analysis of data that correspond to time from a well-defined time origin until the occurence of some particular event or end point. If the end point is death of an individual, the resulting data are literally survival time.

Survival time measures the time to a certain event such as failure, success, death, life, response, relapse, the development of a given disease, parole, divorce or churn. The distrbution of survival time is usually described or characterized by three functions :

1. *Survival function*

This function denoted by S(t) or 1-F(T<t), is defined as the probabilty that an individual survives longer than t.

2. *Probability Density Function*

This function is defined as the limit of the probability that an individual fails in the short interval t to t+Δt per unit width Δt, or simply the probability of death in small interval per unit time.

3. *Hazard Function*

The hazard function h(t) of survival time T gives the conditional failure rate. This is defined as the probability of death during a very small time interval, assuming that the individual has survived to the beginning of the interval, or as the limit of the probability that an individual dead in a very short interval, t to t+Δt per unit time, given that the individual has survived to time t.

$$h(t) = \frac{f(t)}{\{1 - F(T < t)\}}$$

These three functions are mathematically equivalent, if one of them is given then the other two can be derived (Lee 1992).

Censored Data

Censored data means that the observations are only partially known or with incomplete information. There are three different types of censoring possible: right censoring, left censoring and interval censoring. The data is categorized into right censoring if one or more individuals are only known the lower limit of time. Whereas the left censoring is when the interest event already occured to the individual before the individual enters into the research period and it is less common more than the right censoring. And the other one, interval censoring, is the case when the data is divided into intervals and recorded periodically (Klein and Moeschberger 1997).

This research focused only on right censoring. When an observation is right censored it means that the information is incomplete because the subject did not have an churn event during the time that the subject was part of the study.

It is important to understand the difference between calendar time and time in the study. It is very common for subjects to enter the study continuously throughout the length of the study. This situation is reflected in the Figure 1a where we can see the staggered entry of four subjects. The subjects in red were censored and the subjects in blue experienced an event. It would appear that subject 3 dropped out after only a short time (ex : died) and that subject 4 did not experience an event by the time the study ended but if the study had gone on longer (had more funding) we would have know the time when this subject would have experienced an event. Thus, in calendar time both the entry and the exit time of the subjects are staggered and can occur at any time throughout the course of the study.

Study time indicates only with the length of time or lifetime that the subjects were a part of the study. Thus, as described in Figure 1b, every subject start at study time zero and have ending points corresponding to the entire lifetime that they subscribed in the study, in other words, until they experienced an event or were censored.

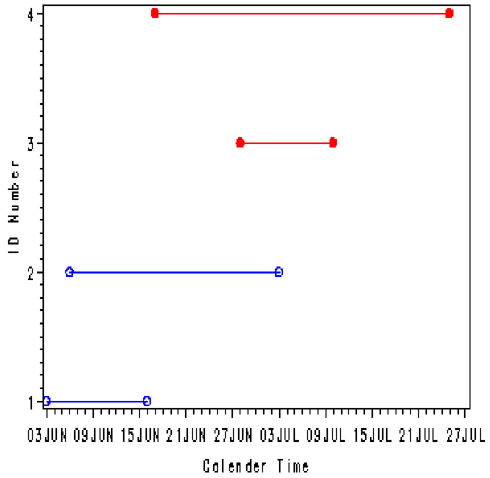

Figure 1a Right censoring

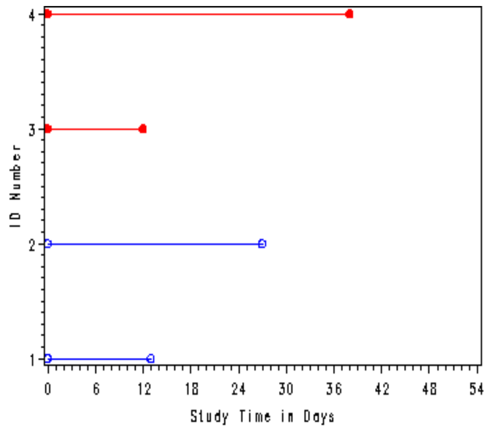

Figure 1b Right censoring

Kaplan-Meier estimate of The Survival Function

Collet (2003) stated that to determine the Kaplan-Meier estimate of the survival function from a sample of censored survival data, a series of time intervals is formed. However, each of these intervals is constructed to be such that one death time is contained in interval, and this death time is taken to occur at the start of interval.

Since there are n_j individuals who are alive just before $t_{(j)}$ and d_j dead at $t_{(j)}$, the probability that an individual dies during the interval from $t_{(j)} - \delta$ to $t_{(j)}$ is estimated by d_j/n_j. The corresponding estimated probability of survival through that interval is then $(n_j - d_j)/n_j$.

Therefore the Kaplan-Meier estimate of the survival function is given by

$$\hat{S}(t) = \prod_{j=1}^{k} \left(\frac{n_j - d_j}{n_j} \right)$$

for $t_{(k)} \leq t \leq t_{(k+1)}$, k = 1, 2,…, r, with S(t)=1 for $t \leq t_{(1)}$ and where $t_{(k+1)}$ is taken to be ∞. Strictly speaking, if the largest observation is censored survival time, t* , say S(t) is undefined for t > t*. on the other hand if the largest observed survival time, $t_{(r)}$ is an uncensored observation, $n_r = d_r$ and so S(t) is zero for $t \geq t_{(r)}$. A plot of Kaplan-Meier estimate of the survivor function is a step function, in which the estimated survival probabilities are constant between adjacent death times and decrease at each death time.

Cox Regression Model

Cox's proportional hazard regression or simply Cox Regression is a method for modeling time-to-event data in the presence of censored cases. Cox Regression will handle the censored cases correctly, and it will provide estimated coefficients for each of the covariates, allowing us to assess the impact of multiple covariates in the same model. We can also use Cox regression to estimate the effect of continous covariates.

The proportional hazard model does not make any assumptions about the nature or shape of the underlying survival distribution. The model only assumes that the hazard of death at any given time for an individual in one group is proportional to the hazard at that time for a similar individual in the other group. If the hazard functions are proportional, the survival function for two groups of survival data do not cross one another (Collet 2003).

Every observation in survival analysis can be written as (t_j, w_j, X_j), where j = 1, 2, 3,..., n which n is the number of observations, $t_j \in (0,\infty)$ is the time that a customer survives, and w_i equal to 1 if the customer is churn at t_j, and equal to 0 if the customer is censored at t_j. X_j is covariate of j-th customer where

$$X = \left[X_{j1}, X_{j2}, X_{j3}, \ldots, X_{jp}\right]$$

X_j is either dummy variable or continous variable.

The hazard function could be expressed as a product of hazard function which is dependent on time and a function of the covariates. The expression is following

$$\left[h(t,X) = h_0(t).G(X,\beta)\right]$$

h(t,X) and $h_0(t)$ are positive, so G(X;β) must be positive.
Cox proposed;

$$G(X;\beta) = \exp\left(\beta' X\right)$$

Thus the model becomes

$$h(t,X) = h_0(t).\exp\left(\beta' X\right)$$

Where
h(t,X) = the hazard of customer churn at time t with characteristic X.
$h_0(t)$ = customer hazard function at standard condition X = 0 (baseline hazard function).
β' = $[\beta_1\ \beta_2\ \beta_3\ \ldots\ \beta_p]$ is a regression coefficient vector or parameter vector.

Linear model for covariates effect is

$$\log \frac{h(t,X)}{h_0(t)} = \sum_{k=1}^{p} \beta_k X_k$$

(Cox and Oakes 1984).

The model is known as proportional hazard because the hazard function ratio from two customers with covariate vector X_1 and X_2, is independent of time t.

$$\frac{h(t,X_1)}{h(t,X_2)} = \exp(\beta'(X_1 - X_2))$$

The ratio above is also called relative hazard ratio and the ratio shows wether the risk of cuatomer churn increases or decreases according to certain treatments or conditions.

Checking The Proportional Hazard Assumption

Cantor (2003) sugested to inspect visually the curve between the natural logarithms of the negatives of natural logarithms of survival density function and the natural logarithms of the survival time for checking proportional hazard assumption separately for each covariates.

Let $S_0(t)$ and $S_1(t)$ be the survival functions for two values of a covariate. It is known that $S_1(t) = S_0(t)\ \exp(\beta)$. Taking the natural logarithms of the negatives of natural logarithms of both sides in the last survival function formula yields

$$\text{Log}[-\log S_1(t)] = \log(r) + \log[-\log S_0(t)]$$

According to above equation, the proportional hazards assumption implies that graphs of $\log[-\log(S_0(t)]$ and $\log[-\log(S_1(t)]$ are parallel or in the other hand visually if the curves seem to be vertically separated by a near constant amount. The same approach can be used for a categorical variable with more than two values.

Besides that, the researchers in University of Carolina Laboratory (2011) also recommended that beside looking at the survival function estimated by Kaplan-Meier curves, it is considered to test of equality across strata by Log-rank test which is a non parametric as supporting the satisfied assumption.

Paramater Estimator

Cox in Lee (1992) sugested a maximum likelihood procedure where the likelihood function is based on a conditional probability of failure. Suppose that $t_1<t_2<...<t_k$ are the k exact failure times. Let $R(t_i)$ be the risk set at time t_i. $R(t_i)$ consists of all individuals whose survival times are at least t_i. For the particular failure at time t_i, conditionally on the risk set $R(t_i)$, the probabilty that the failure is on the individual as observed is

$$\frac{\exp(\sum_{j=1}^{p}\beta_j X_{ji})}{\sum_{l\in R(t_i)}\exp(\sum_{j=1}^{p}\beta_j X_{ji})}$$

The probability product for each uncensored observation time forms a likelihood function that only depends on β. This function is called conditional partial likelihood function :

$$L_c(\beta)=\prod_{i=1}^{r}\left[\frac{\exp(\sum_{j=1}^{p}\beta_j X_{ji})}{\sum_{l\in R(t_i)}\exp(\sum_{j=1}^{p}\beta_j X_{ji})}\right]$$

The function does not depend on $h_0(t)$, because to estimate parameter in Cox regression model, it is not necessary to know $h_0(t)$ as long as data come from the same population.

The maximum likelihood estimates of the β-parameter in the proportional hazards model can be found by maximizing this log-likelihood function using numerical method. To ease the estimation of maximum likelihood estimator $L_c(\beta)$ then we used $Ln(L_c(\beta))$. And to maximize the value we have to derive $Ln(L_c(\beta))$ to β :

$$\frac{\partial}{\partial\beta} lnL_c(\beta)=0$$

The maximization is generally accomplished using the Newton Rhapson procedure.

Variable Contributing Testing

For testing the contribution of variable in univariate analysis a Wald test can be used, with the test statistic :

$$W=\left(\frac{\hat{\beta}}{SE\left(\hat{\beta}\right)}\right)^2$$

With SE($\hat{\beta}$) is standard error of parameter estimator. W is assumed to be normally distributed. This statistic is based on the asymptotic normality property of maximum likelihood estimates.

To test the joint variable significance in multivariate analysis, we use likelihood ratio test with the followng test statistic :

$$\chi^2=-2\left[lnL_{m-r}-lnL_m\right]$$

With L_m as likelihood with m variables in complete model and L_{m-r} as likelihood with r omitted variables from the model. The value of χ^2 at significant level α=0.05 that are χ^2 table with r degree of freedom, shows that those variables have significant effect at 0.05 significant level (Lee, 1992).

Estimation of Survival Function

Estimation of survival function in this Cox regression model use Breslow estimator. The individual survival function until t with covariate X is :

$$S(t,X)=S_0(t)^{\exp(\beta X)}$$

From the above equation shows that to estimate S(t,X), $S_0(t)$ must be estimated firstly, where $S_0(t)$ can be determined as follow :

$$\hat{S}_0(t_i)=\prod_{l,t_1<t_2}\left(1-\frac{d_i}{\sum_{j\varepsilon R_i}\exp\left(\beta^T X_j\right)}\right)$$

with d_i is the number of 'at risk' event at t_i (Collet 2003).

METHODOLOGY

Data Sources

This research involved 26590 customers which comprises from the existing customer in October, November, and December 2009 sampled randomly from the entire customer database existing in these three months. The lifetime equal to zero or it means the time point of a customer is first registered is defined *the origin of time* and their full lifetime until the last observation date

(December 31st, 2009) is *the termination time*. During this lifetime observation period, the timing of customer churn was recorded.

Explanatory Variables include demography data, such as; ages, occupation, marital status, vip, gender, education, region, and religion and feature usage of invoice for example: voice domestic, voice international, voice VoIP, voice 3G, Voice Intl. Roaming, SMS, MMS, GPRS, SMS Intl. Roaming, GPRS Intl Roaming, Data 3G (Oktarina 2009). The definition of this terms shows at Table 1.

A variable of lifetime is used to indicate the time that customer churn occurred, or for *censored cases*, the last time at which customers were observed, both measured from *the origin of time* (lifetime = 0).

Table 1 Description of Feature Usage varibles for analysis

Feature	Description
MSISDN	Mobile subscriber's ISDN, or commonly called mobile telephone number.
Tenure	Duration of subscription, calculate in day unit.
Invoice	Average of invoice in three months (segmentation) or invoice per month (churn). Invoice is sum of following variable.
Voice	
VoDO	Invoice for domestic voice, included local and long distance dialing.
VoIN	Invoice for international voice.
VoIP	Invoice for Voice over Internet Protocol (VoIP).
Vo3G	Invoice for voice 3G, i.e. 3G video call.
Value Added Service	
SMS	Invoice for sending short text message.
MMS	Invoice for sending multimedia message.
GPRS	Invoice for accessing or browsing internet via GPRS network.
Da3G	Invoice for accessing data or browsing internet via 3G network.
International Roaming	
RoVO	Interconnection fee when dialing to or from foreign countries.
RoSMS	Interconnection fee when sending SMS to or from foreign countries.
RoGPRS	Interconnection fee when accessing internet via GPRS.

And a response varible of STATUS is used to distinguish the *censored cases* from *observed cases* (churn event*)*. It is common to have STATUS = 1 for *observed cases* (churn) and STATUS = 0 for *censored cases* and in this study, the survival data are *right censored*.

Method

Methodologies of this research are summarized as follows:

Section I : Cox's Proportional Hazard Model Analysis

1. Preprocessing the data; which includes determining framework, cleaning data, classifying explanatory variables's values into categories.
2. Represent categories of each variable into dummy variables.
3. Conduct exploratory data and univariate analysis for each categorical variable to provide first insight into the shape of the survival function for each group and give an idea of whether or not the groups are proportional (checking the proportional hazard assumption).
4. Conduct Cox's proportional hazard model 1 of customer churns in full model to evaluate the significant influence factors.
5. Interpretation of the result.

Section II : Estimation of survival function

1. Conduct Cox's proportional hazard model of customer churns with Stepwise selection model to find the most influence factors.
2. Make possible characteristic's combinations as our interest based on the significant factors from the stepwise model.
3. Estimate the survival function by Breslow estimator for the characteristic's combination.
4. Describe the customer survival graph for the characteristic's combination of our intereset.
5. Interpretation of the result.

and for processing data, here the researcher use software SAS V9.1 and Ms. Excel 2007.

RESULT AND DISCUSSION

Exploratory Data

From a total of 26590 data, 25511 are identified as censored data (95.94%). It shows that the churn rate of customers is very low. However, we must keep close attention on the time by the time as the preventive effort of increassing churn rate, because there is a fact that in the telecommunications industry the costs of churn experient is many of times more to recruit a new customer than to retain an existing one. Therefore, the customer retention has become even more important than customer acquisition.

the log-rank test of equality across strata which is a non-parametric test.

From 21 explanatory variables, only 14 variables with the natural logarithms of negatives of the natural logarithms of survival function are visually approximately parallel and it is also supported by the test of equality (Log-rank test) that have p-value of 0.05 or less were statistically assumed significant proportional and considered having survival function approximately parallel among the strata in each categorical variable.

For Covariate GPRS shown at Figure 2, it describes that the natural logarithms of negatives of the natural logarithms of survival function and among categories are approximately parallel and it is also supported by the p-value of log-rank test less than 0,05. Beside that the Kaplan-Meier survival

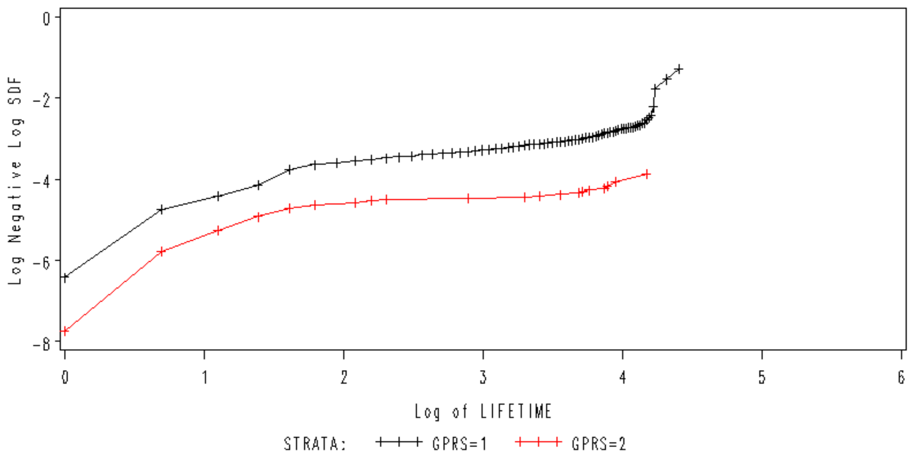

Figure 2 The graph of Log[-Log(S(t)] for each strata in covariate GPRS

Model Assumption

The assumption of proportional hazard model is tested by univariate analysis or exploratory for each of the categorical predictors. In survival analysis it is recommended to inspect visually the graph of the the natural logarithms of negatives of the natural logarithms of survival function or Kaplan-Meier curves for each the categorical predictor. This will provide first insight into the shape of the survival function for each group of a coavariate and give an idea of whether or not the groups are proportional (i.e. the survival functions are approximately parallel).

The researcher also considered the test of equality using across strata or level/groups in categorical variable to explore whether or not to include the predictor in the final model. For the categorical variables the researcher used

function estimation shown in Figure 3 lead us to conclude that customers who access and browse internet via GPRS network with invoice more than Rp 5000,00 (GPRS 2) have slightly longer survival time compared to others who have invoice less than Rp. 5000,00 (GPRS_1). Another potential variables are VIP, Age, Occupation, Marital, Education, Tenure, Invoice, VoDO, VoIN, VoIP, RoVO, Vo3G, SMS, MMS, and RoSMS. All these potential covariates were also included into the next modeling step.

Cox's Proportional Hazard Model Analysis

Cox's proportional hazard model was used to analyze the survival of customer until churn event at time t. The explanatory variables were categories of all variables, which included demography and feature usage data. These explanatory variables were related to the hazard function over time with the status of customer at current lifetime; churn (observed case, status = 1) and stay (censored case, status = 0). This analysis hopefully may give a prognostic twards the customer's tendency to survive subscribing or to churn.

The full Cox model is applied to find the most important factors affecting customer churn related to survival time. The model fit statistic of the model is presented in the table below. The value of -2 LOG L when the model without covariates is 21055.694, whereas for the model with covariates is 13973.443, so yielded X^2- value = 7082.2510 and p-value = <.0001 it shows that for full model H_0 is rejected and means at least one β_i is not equal to zero or significant in alpha 5% of explanatory variables.

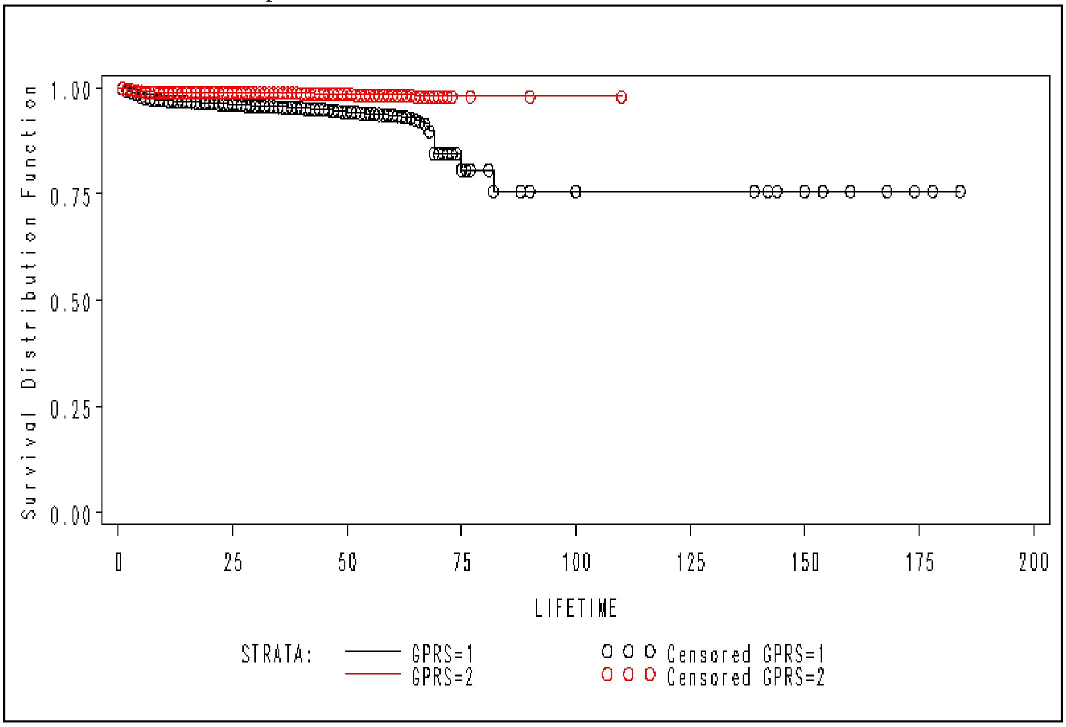

Figure 3 Kaplan-Meier graph of survival density function : GPRS

Table 2 Model Fit Statistics and Testing Global Null Hypothesis: β=0

Criterion		Wtihout Covariates	With Covariates
-2 LOG L		21055.694	13973.443
Test	Df	χ^2	p-value
Likelihood Ratio	33	7082.2510	<.0001

The contribution of each covariate which is significant include 21 dummy variables. The coeficient β-value that have positive value indicates that the reference variable (dummy variable = 0) has better survival, and vice versa. The customer registered as the non-VIP customer has worse survival than VIP customer (β<0) and the status of VIP is significant affecting customer churn (p-value < 0.05). The risk of the VIP customer getting churn is 0.262 times relatively lower than the non - VIP customer.

The average of invoices categories in three months (October, November, and December 2009) which have a significant influence are the customer whose invoice are category II (between Rp 25,000 and Rp50,000) and category IV (more than Rp 150,000). For customer who has an invoice at category II the survival probability is better than the reference variable (i.e. customer that have an invoice at category I or less than Rp 25,000) and the relative risk to get churn is 0.688 times lower than customer who has an invoice at category I. In contrast, the customer who has an invoice at category IV, their survival probability is worse than the reference variable (variable of invoice at category I) and their relative risk is 3.654 times greater than customers at category I (less than Rp. 25,000). Other remaining categories of this explanatory variable (category III), eventhought is not significant, the hazard ratio can also be seen. However the value has a relatively large standard error so it indicates that the parameter estiate is not accurate enough.

The unsignificant explanatory variables are caused relatively by the amount of few data and percentage of censored data is relatively high. In Cox regression analysis, the comparisson of failure percentage for categorical covariate are done, so it needs to have enough large cases in order to ensure that between characteristic oneand another has the difference of failure probability.

Estimation of Survival Function

The survival function, S(t), Bresslow estimator, is estimated with the various of characteristic combinations. Its combinations is conducted by the covariates which are considered as the most significant affecting customer churns. By stepwise selection (α_{stay}=5%, α_{remove}=5%), there were 23 significant dummy variables.

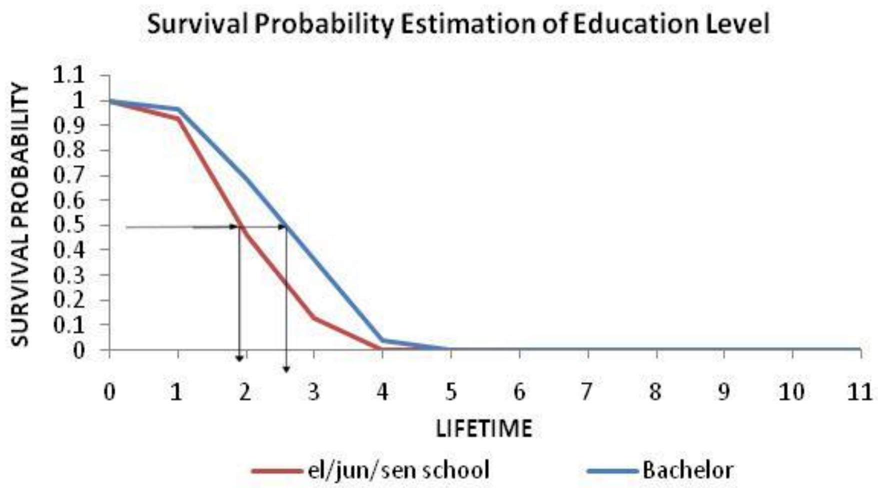

Figure 4 Survival function estimation (Bresslow estimator)

From these 23 dummy variables, the researcher can make possible characteristiccombinations considered as the simulation of interest characteristics of customer to show the survival probability of customer between one another.

$$S(t, Bachelor) = S_0(t)^{\exp(-0.70645)}$$

$$S(t, El - Sen.School) = S_0(t)^{\exp(0)}$$

In the simulation of customer in covariate education_3 (Bachelor) and another covariate's value is cateris paribus or equal to zero (refference variable), we can see at Figure 3 that customer who has education until level of Bachelor, it has the survival value longer than the customer who has education until Elementary, Junior and Senior high school as the refference variable. And we can also predict that half of customer that have these charaterisitcs has a tendency of geting churn at about lifetime 2 and 3 months respectively for Bachelor and El-Sen School. In the same case, we can make another simulation as our interest of different characteristic's combianations then look at the survival performance.

CONCLUSION AND RECOMENDATION

Conclusion

The Cox's Proportional Hazard model of customer lifetime was effective to distinguish relative risk between churn customers and others, and also distinguish between which customers are loyal (stay longer) and which ones stay shorter with their significant prognostic factors. Using all categorical covariates comprises both in feature usages and demography data performed by stepwise selection method in modelling, finally it resulted the best model with the most affected covariates. Afterwards we can conduct a combination of customer characteristics as our interest based on the last significant covariates were obtained and also estimated the survival probability of each the interest combination over time. Afterwards we are able to predict such information of time to churn event.

Finally, the results of this research is able to obtaon a simple, helpful and applicable result for optimizing their customer retention and/or treatment resources in their customer retention efforts for the company.

Recomendation

For the next researcher, it is suggested to include time-dependent variables and interactions between one independent variables into model. It aims to know the effect of possible interactions with time and also between one covariate to each others. Else, building model of lifetime to churn in each single provided segment would be also more practically applied in the bussiness expertise. Afterwards, it will also be interesting too analyze the tendency of reactivation since particular customer churn to the time they have registered back again. It is also interesting to enroll other factors such as BBM (Blackberry Messenger) as one of the feature usages services as covariates.

Another recommendations for the provider company especially that are; analysis will be more applicable if it is conducted 'before –after'any program retention, so that we can easily evaluate and compare specify program developmentn regarding the results.

REFERENCES

Cantor AB. 2003. SAS® *Survival Analysis Techniques for Medical Research, Second Edition*. Cary, NC : SAS Institute Inc.

Cox DR, Oakes D. 1984. *Analysis of Survival Data*. Chapman and Hall,London.

Collet D. 2003. *Modelling Survival Data In Medical Research, 2nd Ed.* Chapman and Hall, London.

Klein JP, Moeschberger ML. 1997. *Survival Analysis : Techniques for censored and Truncated Data*. Springer-Verlag, New York.

Lee ET. 1992. *Statistical Methods for Survival Data Analysis*, 2nd ed. Jhon Wiley and Sons, Canada.

Lu J. 2001. Predicting Customer Churn in the Telecommunications Industry — An Application of Survival Analysis Modeling Using SAS, *Proceeding of Data Mining Conference of DiaMondSUGI 27, Overland Park, 2001, pp114.*

Oktarina SD. 2010. *Understanding Customer Insight by Profiling, Segmentation, and Churn Analisys in Mobile Telecommunication.* Bogor : Departemnt of Statistics, Bogor Agricultural University.

[UCLA researcher] University of Carolina Laboratory. Statistical Computing Seminars Survival Analysis with SAS. http://www.ats.ucla.edu/stat/sas/seminars/sas_survival/default.html [June 1st, 2011]

Widyaningsih Y. 2002. *Penerapan Analisis Regresi Logistik dan Analisis Ketahanan pada Data Masa Laktasi Wanita Indonesia* [postgraduate thesis]. Bogor : Postgraduate Program, Bogor Agricultural University.

Yang LS, Chiu C. 2006. Knowledge discovery on customer churn predivtion. *Procedings of the 10th WSEAS International Conference on Applied Mathematics, Dallas, Texas, USA, November 1-3, 2006, pp523-528.*