International Journal of Smart Home
Vol. 10, No. 6 (2016), pp. 247-260
http://dx.doi.org/10.14257/ijsh.2016.10.6.24

Research on Mobile Internet User Behavior Preference Based on Support Vector Machine

Ye Yuan[1], Haigang Sheng[1], and Haohao Shi[1]

[1] *Chongqing University of Posts and Telecommunications Chongqing, 400065, China*
914867023@qq.com, 1533455810@qq.com, 1103745171@qq.com

Abstract

With the rapid development of the mobile Internet, mobile Internet users grow rapidly and the preference is more diverse, and the user's behavior is also showing a new feature. In this study, we use qualitative research and quantitative research, combined with empirical data and simulation analysis and the combination of traditional statistical analysis and modern data mining methods, to study the behavior preference of mobile Internet. Besides, we use the discrete choice model to build the binaryLogit model using SAS tools, and analyze the selected variables through data processing. Based on the significant influence factors, the behavior preference of mobile Internet users is classified by using support vector machine. The C-SVM binary classification machine is used as kernel function, and the parameters of the model and kernel function are optimized by cross validation. In this research, we introduce the support vector machine theory into the research field of the mobile Internet user behavior effectively, solve the problem of small and medium-sized samples effectively, and provide new ideas and methods for the research of mobile Internet user behavior.

Keywords*:Mobile Internet; User Behavior Preferences; Discrete Choice Models; Support Vector Machines*

1. Introduction

In recent years, the rapid development of mobile Internet has become one of the most remarkablephenomena of world's information and communication areas. According to "The 36th China Internet Development Statistics Report" from China Network Information Center (CNNIC) statistics as at the end of June 2015, the scale of China reached 668 million Internet users, of which mobile phone users scale up to 594 million, representing an increase of 3679 million by the end of 2014, Internet users the proportion of people using mobile Internet increased to 88.9%. Mobile Internet has become a new driving force development of the Internet.On the one hand, the development of mobile Internet promotes the popularization of the Internet in China.Especially thatprovides the possibility of using the Internet for the people and areas that affected by network, terminal and other restrictions cannot access the Internet.On the other hand, mobile Internet promotes the new growth of the Internet economy.The innovation boom based on mobile Internet provides a new business model and development space for the traditional Internet business.Compared to users in the traditional telecommunications market, the user'spreference of mobile Internet environment is more diverse, and user’sbehavior also presents new features, which bring great opportunities and developments to the telecommunications industry, and give the intensified traditional telecommunications market a new development space. The innovation brought by mobile Internet makes people pay more attention to the user's behavioral preferences, and then meets the needs of users, create more wealth.

With the development of computer database technology, people use some calculation

ISSN: 1975-4094 IJSH

means analyze consumer' behavior, data mining, data analysis and others dominate the market.Among them the Support Vector Machine approach is increasingly applied to many fields and has been the major concern of many fields. Support Vector Machine in face detection, validation and recognition, text/handwriting recognition, image processing has achieved a lot of research achievements, and also has made some significant achievements in the field of telecommunications.

2. Related Research Status

2.1 Research on Mobile Internet Users

So far the research for mobile Internet users is mainly from the user's perspective to study the mobile Internet user'sattitude, characteristics and behavior, but in general the relevant research literature is still relatively scarce. In feature-based user attitudes and demographics, the method by Okazaki(2006) uses a two-stepclustering method to divide mobile Internet users in Japan into four categories, while he found among young people who have a good economic condition, core components of mobile Internet users are the office clerk and student, and a little proportion is highly educated experts or freelancers.Lee(2010) and others surveysseven areasincluding Korea, Hong Kong, Japan, Denmark, Australia, Taiwan and Greece based on time perception, uncertainty avoidance, context and individualism.Through the measurements of personal and cultural identity related to mobile Internet,they came to a conclusion that over the region, while the users have the same idea at the time perception and personal uncertainty avoidance, but not in the context of and individualism.Through comparison of different genders, Okazaki and Romero (2010)find that the Japanese people, who are both mobile Internet and PC Internet users is more than a third of the total number, in which mainly are employed women, and men only take up a large share in the PC Internet users. Seongil(2009) further researches the attitude of Korea's mobile Internet users , and found differences in the type of mobile Internet services (mainly including personal, communication, information and entertainment), age differences and differences in the form of service delivery(such as pictures or text) will lead to the user's different attitudes.

2.2 Research on Preference of Mobile Internet User Behavior

Current research on user behavior preferences is widely, but most from the aspect of impact on the user's preference by marketing strategies study brand selection inertia, brand preference and brand loyalty of users selection.Guan(2005) found that when the user has certain preference in attribute of commodities, their degree of preference for commodities can be predicted.Butler(2008)found when the user's preference about product attributes is not sure, you can take advantage of multi-attribute preference models and fuzzy set theory to build preference prediction model. Zenebe(2010)holds that the user's preferences for goods or services are based on specific attributes of goods or services, such as brand, price, appearance, etc. Hu Changping(2008)choose the influence on personalized information services by user preferences as a starting point.Through the analysis of the formation of preferences and preferences for personalized information and services under network environment, he studied the features of individual user's behavior and accessed to information under the guidance of preference, and proposed the personalized information service improvement strategy. Shang Huiying(2009) established a dynamic game-model of consumer and enterprise on product factors, which analyzes the constitutions of expected product for the user and impact factors in the long-term situation.

At present in some emerging research, through the data mining method to study the behavior of users gradually become the trend. On the basis of analyzing the structure and the acquisition of the user behavior in the library, Wang Wei (2012)established the user

behavior model of the library by using the data mining technology. Huang Wenbin (2015) based on that domestic telecom operators were randomly sampled from ten thousand mobile users a week of log records, extracted 14 kinds of characteristic index in this data from consumption ability, call volume, network requests and the amount of displacement. And then, by using K-means clustering method, they divided the Mobile user into 4 types including regular telephone calls, random access, home saving and random high consumption. Wang Renwu, Yuan Yi (2011) by cleaning extracting and loading of huge amount of Web log, built user behavior data warehouse. Besides, in combination with the user access path probability matrix model, they made data mining and wanted to realize user behavior monitoring intelligently, and provide timely and high-quality information service to users.

2.3 Review

For mobile Internet users, life services business as well as leisure and recreational activities have the same importance. In addition, the users change towards the direction of diversification, and mobile Internet users are more inclined to focus on the class of business communication and information services business, which is different from the traditional users who pay more attention to the media and the game.Compared in preferences, the mobile Internet users is better than the traditional telecom in species. Further, use of the Internet user's time is more fragmented, the information obtained is fragmented, so the research on mobile Internet user behavior preferences is necessary.

3. Discrete Choice Model and Support Vector Machine Theory

3.1 Discrete Choice Model

Discrete Choice Model, also known as Qualitative Response Models, is represented by the difference exists between the set of options for continuous variables and discrete variables. Mainly there are four discrete choice models: Logit model, GEV model, Probit model, Mixed Logit model.

In reality, many factors affect the choice of decision-makers.When this effect arises to a certain extent, a choice behavior will generate, otherwise it will not happen. Utility refers to a measure of their needs met for consumers through other consumer or service. The random utility theory refers to the customer in accordance with the principle of utility maximization behavior to make decisions, that is, choose products or services of the maximum effectiveness. American economist Professor Daniel McFadden proves the probability of a product from a collection of many different products for a consumer can be expressed by a function of all product features.

3.2 Support Vector Machine

Support vector machine is a new machine learning method based on statistical theory.Through non-linear kernel function, the input sample space is mapped into a high dimensional linear feature space, with the ability to handle non-linear regression problems. This methodovercame the long artificial neural network training time, generalization ability, easy to fall into local minima and other shortcomings, and improved the generalization ability of the model.

4. Preference of Mobile Internet User Behavior Based on Discrete Choice Model

4.1 Subdivision of Mobile Internet Users

4.1.1 Dimensional Analysis: The premise of the user segment is to select dimensions of user segmentation, and this dimension is determined according to the actual target of number mining projects. In the process of data mining, data determines whether we can succeed. Not all data can predict the ideal result, because data quality directly determines whether the model is to be successful,and only the data which meet the quality requirements will lead to the ultimate success model. In determining the complete project indicators, the index system of the mobile Internet is very important,for it directly affects the project index, so establishing a user index system directly affect whether data mining system can be succeed.

Dimensions include dimensional elements and metrics. Dimensional elements correspond to the analytic perspectives, while metrics is concerned with specific indexes.Indexes can be discrete data allowing only a limited value, and can also be continuous data that value be unlimited.

Dimension is a certain angle of data observation, that is, a class of attributes when consider an issue.Attributes set to form a dimension, such as the time dimension, the geographic dimension, etc.Assume that for a mobile Internet user, there are some factors that can affect the behavior of his use of the mobile Internet such as age, gender, occupation, etc. These are all dimensions.

Dimension has its own inherent attributes, such as hierarchical structure, sorting, computing logic which is very usefulfor decision support. In this study, the user's attributes are divided into three kinds including natural attributes, behavior attributes and perceived attributes.

4.1.2 Data Collection

In the data collection process, we use the combination of survey data and objective data from carrier systems, so that we can both understand the true feeling of the users and can analyze user’s real preference scientifically.

In this study, the survey data acquisition is firstly by telephone survey. We randomly selected mobile number to conduct research, and ultimately get 300 success samples,access to basic services indicators for those users using being from the system background. Next, based on expert interviews and the interview outline determined by literature, we conduct telephone interview to 300 users about behavior attributes and sensible attributes.

Ultimately, this research obtains 300 valid survey samples and made statistical analysis of the main indicators of the sample. In the respondents, in the view of the post distribution the unit and department heads,self-employed occupy the highest proportion 22.3%; in the view of age the proportion of 20-19 year-old age make up to 48.4%.

4.2 The Main Factors in the Mobile Internet User Behavior Preference

As can be seen from the data collection,there are many mobile Internet user's preferences, and we want to find out the most important factor. Since the data collected are mostly discrete data, so we choose to use discrete choice model to analyze.

4.2.1 Data Processing: The survey deals with the user's choice by means of whether or not. For example, for online shopping (mobile payment, group-buying, download coupons, ordering hotels and tickets, etc.),one of the use of mobile Internet users, we changed into whether the users do online shopping (mobile payment, group-buying, download coupons, ordering hotels and tickets, etc.), if yes labeled as 1, otherwise labeled 0, thus a new data matrix willbe formed.

For data such as the month ARPU value, the month total flow, month WLAN duration of users, we do normalized to avoid the effects brought about by the different dimensions, the specific formula is as follows:

$$X' = \frac{X - minvalue}{maxvalue - minvalue}$$

Among them, X' is the raw data,X is normalized data,minValue is the minimum value of the X field, andmaxvalue is the maximum one.

We label whether people are willing to increase the cost of mobile Internet as the dependent variable y, the other 39 variables as from x1 to x39 to fulfill the definition of variables.Such as month ARPU value is labeled as x1, user's gender as x39.

Assume that mobile Internet users are willing to pay for mobile Internet probability is p_i, $1-p_i$ represents the probability of mobile Internet users do not want to pay for the mobile Internet.By establishing basic binary Logit model, the mobile Internet user's willing to pay for mobile Internet model can be expressed as follows:

$$log\frac{p_i}{1-p_i} = \alpha + \beta X_i = \alpha + \beta x_1 + \beta x_2 + \beta x_3 + \cdots + \beta x_{39} + \varepsilon_i$$

Among αis a constant, β is positive number when positive correlation,is negative when negative correlation, $\varepsilon_i$is random variable.

4.2.2 Experimental Analysis: In this study, by means of SAS programming, in the process of obtaining results, the SAS system selection is a binary Logit model, optimization method used is Fisher scoring method, and the number of observations is 300. Depending on the operating results of the SAS System, in 300 mobile Internet users those willing to increase the cost of mobile Internet are only 32. And it can generate maximum likelihood estimates of the analysis table based on the SAS System, whereby the logistic regression equation is:

$$\begin{aligned}\log\frac{p_i}{1-p_i} = &-0.00052x_1 + 0.000213x_2 - 0.00212x_3 + 2.3432x_4 - 0.3670x_5 \\ &+ 0.0636x_6 + 0.1964x_7 - 0.0441x_8 + 1.5033x_9 + 1.5658x_{10} \\ &+ 0.7412x_{11} - 13.7082x_{12} + 1.1241x_{13} + 2.1537x_{14} + 3.6836x_{15} \\ &+ 1.8223x_{16} - 0.401x_{17} + 0.7268x_{18} - 0.00710x_{19} + 0.3303x_{20} \\ &+ 1.2918x_{21} + 0.6395x_{22} + 1.0437x_{23} - 0.149x_{24} - 0.1492x_{25} \\ &- 3.2841x_{26} + 1.3267x_{27} + 1.0956x_{28} - 0.472x_{29} + 0.00255x_{30} \\ &- 0.2345x_{31} + 1.3932x_{32} + \beta x_{33} - 1.0178x_{33} - 0.5497x_{34} - 0.2326x_{35} \\ &- 13.4136x_{36} - 0.0416x_{37} - 0.0241x_{38} - 0.5493x_{39} - 3.0296\end{aligned}$$

This formula can be used to calculate the probability of mobile Internet user's willing to increase the cost of mobile Internet. This can be completed by putting the value of factors which affect user's willing to pay into this equation. As can be seen in the 39 factors, there are many positive factors. These three factors x_4, $x_{14}$and $x_{15}$are the most significant ones, and coefficients are 2.3432, 2.1537 and 3.6836.Among the negative factors, x_{12} (-13.7082), x_{26} (-3.2841), x_{36} (-13.413) are the most significant factors.

In the P value test, the main influencing factors (x_4 = 0.0231 <0.1, x_{14} = 0.0779 <0.1,

x_{15} = 0.0322 <0.1, x_{12} = 0.0681 <0.1, x_{26} = 0.0165 <0.1, x_{36} = 0.0766 <0.1) P-value are all less than 0.1, so they all pass through the *p* value test and the impact is significant.

At the same time SAS also generated the forecasting probability which fits well with operating results and observations obtained,as to the percentage of same section is 85.4%, much higher than 70%, indicating a good predictive capabilities.

5. Preference of Mobile Internet User Behavior Based on Support Vector Machine

5.1 Data Processing

Through these studies, we have identified the six significant factors affecting mobile Internet user's willing to increase the cost of mobile Internet: three positive correlation factors, three negative.This article choose these six dimensions as main dimensions in support vector machine, shorten the process for future research and predictions, save survey costs and improve the success rate of research. These six dimensions are marked with six new variables to replace the original 39 variables: (1)think that the Internet speed is more important, denoted by *A*; (2) think that the use CMCC is trouble, denoted by *B*; (3) think CMCC charge is expensive, denoted by *C*; (4) think that the mobile phone data plan is enough, denoted by *D*; (5) the use of mobile Internet business for social networking sites (happy, everyone, school, etc.), denoted as E; (6) the use of mobile Internet for the diverse needs, recorded as F.

5.2 Model Construction

5.2.1 Model Selection: The experiment in this article is based on the SVM software package -LibSVM, which is developed by the Dr. Chih-Jen Lin who is from National Taiwan University, and LibSVM is a simple, fast and effective software program about SVM pattern recognition and regression. LibSVM software which supports C-SVM classification is the most basic classification of two- element, the specific form is as follows:

C-SVM classification which is supported by LibSVM is the most basic binary classification;the concrete form is as follows:

(1)Set known training set;

$T = \{(x_1 \cdot y_1), (x_2 \cdot y_2) \cdots, (x_i \cdot y_i) \in (X \times Y)^m\}$

the $x_i \in X = R^n, y_i \in Y = \{1, -1\}, i = 1,2, \cdots, m;\ x_i\ is\ eigenvector.$

(2)Select the appropriate kernel function K and the appropriate parameters C to construct and solve the Optimization Problems:

$$min \frac{1}{2} \sum_{i=1}^{j} \sum_{j=1}^{m} a_i\, a_j y_i y_j K(x_i \cdot x_j) - \sum_{j=1}^{m} a_j$$

$$s.t. \sum_{i=1}^{m} a_i y_i = 0$$

In this formula$0 \le a_i \le C, i = 1,2, \cdots, m$,the optimal solution is $a^* = (a_1{}^*, \ldots, a_m{}^*)^T$;

(3)From the a^*, choose $a_j{}^*$ which is less than *C* as a normal classification and Calculate:

$b^* = y_j - \sum_{i=1}^{m} a_i{}^* y_i K(x_i \cdot x_j)$

(4)The constructed decision function is:

$$f(x) = sgn\left(\sum_{i=1}^{m} a_i{}^* y_i K(x \cdot x_i) + b^*\right)$$

The $K(x, x_i)$ is kernel function,C-SVM is the classification method we should use in

this study.

In the construction of the decision function, it is necessary to determine the parameter *C* of the kernel function and the model. When we select the kernel function, the parameters of the kernel function also need to determine.

5.2.2 Kernel Functions Selection: Learning algorithm depends on the selection of parameters, and the parameters control the prediction accuracy, so the selection of parameters is very important. The selection of kernel functions is related to the field of application. The kernel functions we often use including Linear Kernel, Polynomial Kernel, Radial Basis Function (RBF) and Sigmoid Kernel. In this article, we select RBF as the kernel function, it is:

$K(x, x_i) = Exp(-\gamma \|x - x_i\|^2), \gamma > 0$

The main reasons are as follows:

1. The linear kernel is a special case of the RBF, and if we will use the RBF, it is not necessary to consider the using of linear kernel.

2. Using RBF on the calculation is simpler than Polynomial Kernel, and the parameters which need to determine in polynomial kernel are more than RBF.

3. The Sigmoid may be invalid in some parameters, and the accuracy is not higher than that of the RBF.

5.2.3 Parameter Selection: In this article, the parameters of the experiment are divided into two types: the C of model parameter and the γ of kernel function parameter. The specific instructions are as follows:

1.The *C* of the model parameter is the penalty parameter, and the greater the *C* is, the greater is the penalty of the error classification. In the SVM, the *C* of model parameters is used to adjust the proportion between the complexity of model structure and empirical risk. The size of the adjustment factor *C* determines the share of the experience risk in the adjustment risk. With the increase of *C*, the proportion of empirical risk in the real risk gradually increased. When *C* tends to infinity, the structural risk completely becomes the experience risk. And the structural risk minimization also degenerates to the empirical risk minimization.

2.The parameter γ of kernel function: The parameters of kernel function usually reflect the distribution of the input values in the training data, and the parameter γ of RBF reflects the range or the distribution of the receptive field.

In order to get the optimal value of the parameters, and make the prediction results of the training samples are optimized,this article uses Cross-Validating,through the cross-validation of training samples and the gradual approximation of the parameters of the model, to make the prediction results are optimal. The purpose of using cross-validation is to get as much effective information as possible from the limited learning data; from multiple directions to start learning samples, to avoid getting into local minimum point, and to avoid over fitting problems to a certain extent. On the basis of single cross-validation, the different (C, γ) values were repeatedly tested, and the optimal (C, γ) parameters were obtained.

5.3 Experimental Process

5.3.1 Experimental Analysis: First, we have to select the training set and test set, and extract the data to carry on the pretreatment. Then we select the best parameters c and g with cross-validation. Next, we use the training set to train the SVM, and finally use the model to predict the test set. The algorithm flow is as following graph:

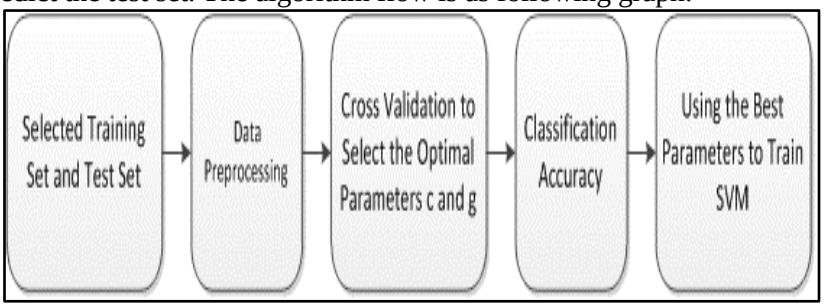
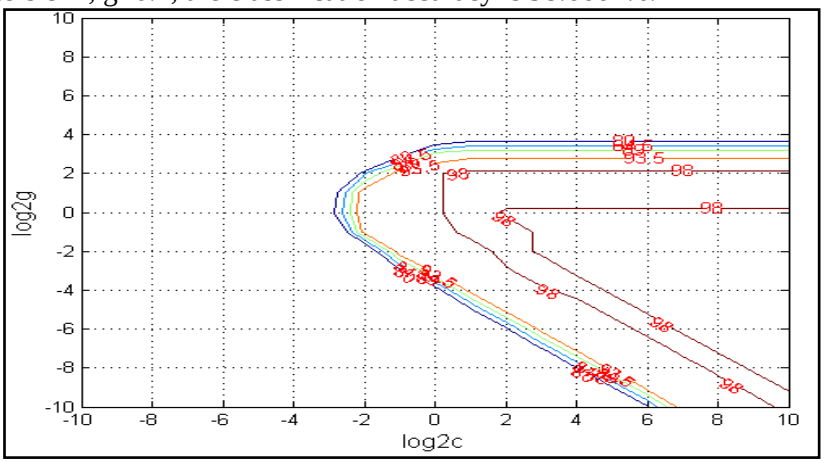


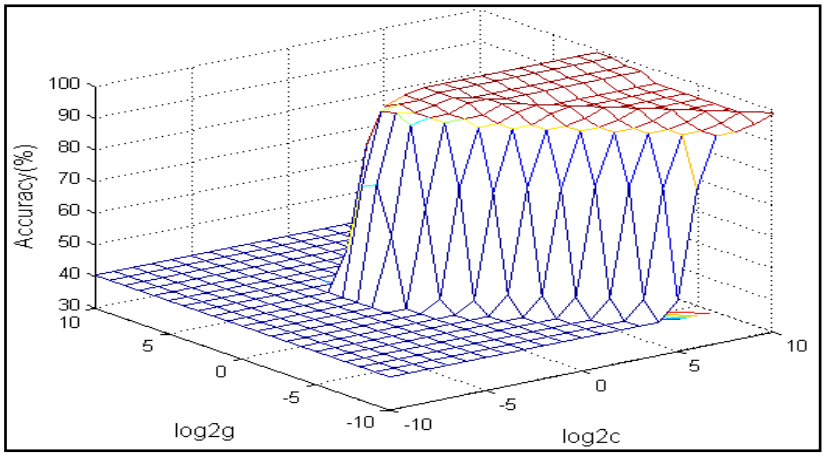
Figure 1. Flow Chart of the Model

Select training set and test set. The 32 (class label is 1) users of the first kind that are willing to increase the cost of mobile Internet access, and the 268 (class label is 2) users of second kind that are reluctant to increase the cost of mobile Internet, are divided into two groups. Make a part as the training set, the other part as the test set. We choose 16 users from 32 users as the training set, 134 users from 268 users as the training set. In this article, we choose the RBF as the kernel function and determine the optimal parameters c (parameter c of model) and g (parameter γ of kernel function) by using the method of parameter optimization.

Operating function SVMcgForClass can realize the parameters optimization algorithm, and we get the selection result map about relevant parameters. The results show that when the parameters c=2, g=0.1, the classification accuracy is 98.6667%.



（a）Contour Map of Parameter Selection Result



（b）3D View of Parameter Selection Result

Figure 2. Graph about Parameter Selection Result

5.3.2 Experimental Comparison:

1. Classification accuracy comparison of different kernel functions

For different kernel functions in SVM, this study further compares the accuracy which is of the prediction of classification of the test set. Parameters are chosen for c=2, g=0.1, such as table: Parameter options

Table 1. Comparison of Different Kernel Functions

Kernel Functions	Classification Accuracy	Parameter options of Svmtrain
Radial Basis Function	98.6667%(148/150)	c=2,g=0.1
linear	95.3333%(143/150)	c=2,g=0.1
sigmoid	58.6667%(88/150)	c=2,g=0.1
polynomial	98.6667%(148/150)	c=2,g=0.1

Through the comparison of the above table, we can see that when the radial basis function is used as the kernel function, the final classification accuracy is the highest.

2. Classification accuracy comparison of different parameters.

In this study, the parameters of c and g were randomly generated by *rand_c* and *rand_g*, and SVM were trained by the parameters of this group. Finally, through prediction, the classification accuracy is obtained.

Compare the classification accuracy obtained by the parameters optimization. The results are as follows:

Table 2. Comparison of Different Parameters

Running times	Random parameters c	Random parameters g	Classification accuracy
1	36.6754	89.8734	44.6667%
2	60.4562	5.2753	91.3333%
3	3.8756	79.2386	45.3333%
4	77.5109	21.8952	54.6667%

As can be seen from the table, the parameters c and g cannot guarantee the accuracy of the final classification. Sometimes it can choose the appropriate parameters by experience, but it cannot guarantee that this is the best parameters, even though it can make the correct classification rate reached the highest. But the parameters which are obtained by parameter optimization algorithm in this research can ensure that the parameters of the

algorithm can guarantee the best in a sense and the classification accuracy is also higher.

5.3.3 Analysis about the Importance of Forecast Variables: In this article, we use IBM SPSS Modeler to predict the importance of variables and analysis. The RBF is used to build the model. The established model is as follows. In the process of modeling, the parameters optimization data that are obtained from MATLAB in the upper section, are *c=2*, *g=0.1*, which are used to edit the parameters of RBF.

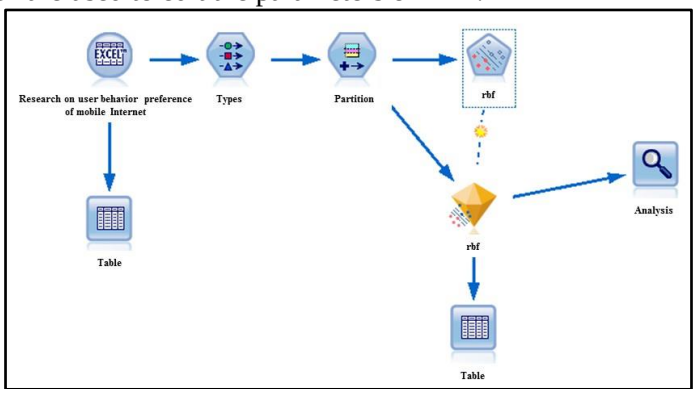


Figure 3. Modeling Flow chart

Finally, the analysis of the importance of the predictive variables can be obtained by analyzing options, as shown below. From the graph, we can see that the option *D* ("Think mobile data plans is enough") is the most important for the classification and prediction, and relatively the option *F* ("The demand of surfing the Internet through mobile phones is diversified") is the least important.

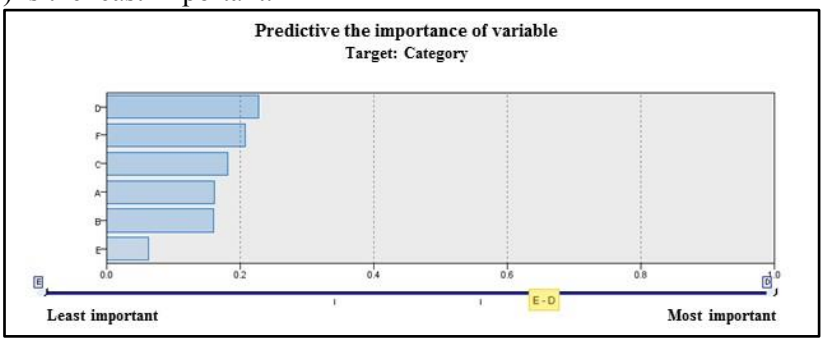


Figure 4. The Importance Predictive Variable

6. Conclusions and Suggestions

6.1 The Main Conclusions of This Study

On the one hand, this paper studies the behavior preferences of mobile Internet users based on the discrete choice model. Firstly, the user is subdivided, and the main variables of the study are selected to study through the analysis of the user's natural attributes, behavior attributes and perceived attributes. Then we obtain the corresponding data through the telephone survey and operator background system, and do the basic statistical analysis of these data. Combined with the discrete choice mode, we use the SAS tool to build the binary Logit model. After data processing, the selection of the variables is analyzed. The significantly influencing factors are found out, and the steps of concrete algorithm and the empirical analysis are given. Experimental analysis shows that the majority of mobile Internet users are male users who in the 20-29 years old。And most of them are person in charge of the units or departments, individual boss, and staffs of the

enterprises and institutions or organ. The users who are willing to increase the cost of mobile Internet users pay more attention to the speed of the phone's Internet, and have higher requirements on the use of the convenience and cost of WLAN. While students and low income groups are more inclined to use social networking sites or mainstreammobile phone business, and generally do not want to increase the cost of mobile internet. And we found that if the mobile phone traffic packages to meet the needs of users, users generally do not want to continue to increase the cost.

On the other hand, based on that the main factors influencing the increase of mobile Internet cost are established by using the discrete choice model, this article use SVM to study the behavior preference of mobile Internet users, and do classification and prediction on the mobile Internet users.Then this article use the LIBSVM tool, select the C-SVM classification algorithm and the RBF and find the optimal parameters c and γ by cross validation. After SVM training and prediction, the accuracy of the classification is very ideal. This shows that the method based on SVM is very suitable for the research on the behavior preference of mobile Internet users.Furthermore, by comparing different kernel functions and parameters, the classification accuracy of the method that we choose is the highest. Finally, we use IBM SPSS Modeler to analyze the importance of predictor variable of each variable. The results show that "Think mobile data plans is enough" is the most important variable in classification and prediction.

6.2 Suggestions on the Development of Mobile Internet

Through the above research conclusions, this study gives the following suggestions for the development of mobile Internet.

1.Make reasonable tariff standards

In the case of network base is developed, terminal function is strong, content/application is more abundant, tariff has become a key factor to stimulate or impede the development of mobile Internet business.For international mainstream operators, the support and reserves of network, terminal and content/application are already available. Therefore, they in order to further promote the development of mobile Internet, push out tariff package scheme that based on monthly traffic.Hutchison Whampoa Limited pushed out two kinds of mobile Internet tariffs: 2.5 pounds/month 10M, 5 pounds/month 1GB. Vodafone's tariff packages are more flexible. They not only have the monthly charged package, but also the daily charged tariff package: 5 pounds/month 500M, 20 pounds/month package 5GB, 50 pence/day 25M. From mobile Internet tariff of foreign operators, their tariffs are very flexible way. They not only have monthly charged tariff package and daily charged tariff package, but also have tariff package of the different business and the tariff package of the different terminal. Flexible and diverse tariff packages adapt to the requirements of different users, such as the use of different levels of users, the use of different characteristics of users, etc. Therefore, when the operators make tariff policies, they need to take into account the relationship between the amount of income and the quantity of the traffic, and maintain a proper balance between the tariff adjustment and the increase in traffic flow.

2.Emphasis on network construction, improve network speed

According to the results of previous studies, the user's perception of mobile Internet speed significantly affects the intention of increasing the cost of the mobile phone. Network, especially the two-way high-speed network is the basis of the development of mobile multimedia applications and high bandwidth applications. Therefore, foreign operators attach great importance to the continued escalation of the mobile network. Currently, most foreign WCDMA networks have been upgraded to HSUPA, providing downlink speeds of up to 21.6Mbps and 5.76Mbps uplink rate; and C network operators such as Verizon, KDDI, SKT and other operators have to upgrade the network to EV-DO Rev.A. The improve of network speed enables it to carry a series of mobile multimedia

applications and download services based on high bandwidth and high speed, and further promote the development of mobile Internet business.

On December 4, 2013, the Ministry of Industry and Information Technology issued operating licenses of 'LTE/TD-LTE' to China Mobile Communications Corporation, China Telecom Group Corporation and China United Telecommunications Co. Ltd. And when condition is mature, the MIIT will issue operating licenses of FDD-LTE according to the application. Launched the 4G will help the rapid development of China's mobile Internet.

3. Develop More Mobile Applications

From the results of empirical research, the mobile application business is also the key factor. Operators can develop more interesting online business themselves. But operators do not have the core *R&D* capabilities in the content. So the operators can take a strategic alliance strategy. And they can cooperate with content providers to launch mobile Internet business with high entertaining. For this reason, operators move the contents that have been popular in fixed Internet to the mobile Internet. Such web content on the fixed Internet has a number of loyal customers, and therefore it provides the possibility for these people to become mobile Internet users. And the content of the fixed Internet is very attractive. So moving it to the mobile Internet will inevitably lead to a new group of customers, which can improve the current value and potential value of mobile Internet users. And this can also upgrade the business that have been popular in the mobile, such as the interface is more humane, content personalization operation is more intelligent. In short, this is more entertaining after upgrading. Operators can reevaluate some of the low usage rate of business or content and in order to maintain and improve the current value and potential value of the user, they can abandon or improve this. They can also develop innovative mobile Internet business. Innovation is always the driving force of an enterprise, so that they can develop their own or joint the content providers to develop new mobile Internet business to make the mobile Internet attract more mobile Internet users.

6.3 Prospects for Further Study

Due to the limitation of time and research ability, there are still many problems need to be further studied and improved in this paper. In the future research, we can further study the following aspects.

1. In this paper, the method of cross validation is also a rough parameter search algorithm. How to select the parameters of a fast and effective way is still a problem to be solved in the future.

2. Because of the hidden nature of information, as well as some qualitative index cannot be accurately quantified and other series of problems, the selection of indicators cannot completely reflect the needs of users. Therefore, in the construction of the user index system is also required to do further work.

3. SVM is a machine learning method based on small sample data. In practical applications, most of the data sets are very large. How to improve the application performance of SVM in large data sets is a very practical research direction.

4. Different stages of the development of mobile Internet naturally have different characteristics. With the continuous development of mobile Internet business, in the following research, the research on the behavior preference of mobile Internet users can expand the sample range and sample size at different levels to make the data sources more abundant and more representative.

Acknowledgement

This work was supported by Doctoral Fund of Chongqing University of Posts and Telecommunications) (K2014-13), Social science planning project of Chongqing city in 2015 (2015YBGL111), 2015 National Social Science Foundation of China West Project (15XGL024), and the Chongqing municipal education commission project of humanities and social science, research on accelerate the path of the triple play in Chongqing and policy (16SKGH058)

References

[1] Okazaki,S. What do we know about mobile internet adopters? A cluster analysis[J]. Information&Management.**(2006)**, 43(2):127-134.
[2] Lee, I., et al. Measurement development for cultural characteristics of mobile internet users at the individual level. Computers in Human Behavior[M]. **(2010)**,In Press, Corrected Proof.
[3] Okaxaki, S. &Roinero, J. Online media rivalry: a latent class model for mobile and pc internet users[J]. Online Information Review.**(2010)**, 34(1):98-124.
[4] Seongil,L. Mobile internet services from consumers' Perspectives[J] International Journal of Human-Computer Interaction. **(2009)**, 25(5):390-413.
[5] Guan, S. U., Chan, T. K.,Zhu,F. M. Evolutionary intelligent agent's fore-commerce: Generic Preference detection with feature analysis[J]. Electronic Commerce Research and Application, **(2005)**,4(4); 377-394.
[6] Butler, J. C., Dyer, J. S., Jia, J. M.,et al. Enabling e-transactions withmulti-attribute preference models [J], European Journal of Operational Research,**(2008)**,186(2):748-765.
[7] Zenebe, A., Zhou, L., Norcio, A. F. User preference discovery using fuzzy models[J]. Fuzzy Sets and Systems,**(2010)**,161(23): 3044-3063.
[8] HU Changping, SHAO Qigan, Sun Gaoling. Analysis of user's preference and behavior in personalized information service [J]. Information Studies:Theory & Application.**(2008)**(01):4-6.
[9] SHANG Huiying, Xi Gang. Dynamic analysis of the impact of marketing tools on consumer preferences [J]. Statistics and Decision. **(2009)**(10):187-188.
[10] WANG Wei. Study on Library Users Behavi or Analysis and Preferences Based on Data Mining[J]. Information Science,**(2012)** (3):391-394.
[11] Huang Wenbin;Xu Shanchuan;Ma Long;Wang Jun. Analysis of Mobile User Behaviors with Telecommunication Data[J]. New Technology of Library and Information Service,**(2015)** (5):80-87.
[12] WANG Ren-wu,YUAN Yi. An Empirical Study of User Behavior analysis Based on Web Log Mining [J]. Library Tribune, **(2011)**, 31(4):100-102.

Authors

Ye Yuan, is an instructor of Management Science in School of Economics and Management at Chongqing University of Posts and Telecommunications. His main research interests are next generation network, consumer behaviour and management of science.

HaigangSheng,is a graduate student in Chongqing University of Posts and Telecommunications, research direction for the management of science and Engineering.

HaohaoShi, is a undergraduate student in Chongqing University of Posts and Telecommunications, research direction for marketing management and consumerbehavior.