Demand Forecasting on The Mobile Communication Service Market Using Diffusion Models and Growth Curve Models: A Case Study

Sangmun Shin[1]*, Yonghee Lee[2], Kyungjin Park[1],Yongsun Choi[1], Charles Kim[3],
[1]Department of Systems Management & Engineering, Inje University
607 Obang-Dong, Gimhae, Gyungnam 621-749, South Korea
[2]Department of Industrial Management Engineering, Dong-A University
840 Handang-Dong, Saha, Busan 604-714, South Korea
[3]School of Computer Engineering, Inje University
607 Obang-Dong, Gimhae, Gyungnam 621-749, South Korea

Abstract: - For the efficient management of radio resources, a scientific/systematic management system using an engineering approach based on cost-benefit analysis needs to be prepared. In forecasting market competition and demand for various services provided by frequency band, it should be considered necessary to enhance the reliability of forecasting data using the most rational forecast techniques. Many researchers have investigated models of forecasting the demand for new products. For the middle or long term demand forecasting, diffusion models and growth curve models are frequently used. In this paper, we demonstrate how to estimate reasonable and highly reliable forecasting data about demand of subscribers for service at the mobile communication market by applying diffusion and growth curve models.

Key-Words: - Diffusion model, Growth curve model, Demand forecasting, Mobile communication

1 Introduction

Continuous improvement of mobile communication technologies has become widely recognized by industry as critical in maintaining a competitive advantage in the marketplace. It is also recognized that the mobile communication market in South Korea have fast grown since 2000. For example, the number of mobile phone subscribers have reached 40 million in South Korea in December 2006, which represented that the mobile phone subscription rate exceeded 82% of the population. In addition, South Korea may recently lead the mobile communication market in the world by providing the new commercialization service of HSDPA (3.5G frequency band).

The existing frequency band amounted to a saturation state for extended frequency demands after the 1990's, and economic scarcity of a frequency was highlighted. Developed nations faced with an excess demand in a specific frequency band and other nations have made political efforts to settle such a frequency scarcity problem. In South Korea, the lack of frequency also has already surfaced for some bands and the economic weight of industries related to radio wave and an equipment manufacturing is rapidly increasing. Responding to these trends actively, it is necessary to establish an advanced radio resource management system. For an efficient management of radio resources, a scientific/ systematic management system using an engineering approach based on a cost-benefit analysis needs to be prepared. However, previous researches reported in literature for the management system of the radio wave may not consider the demand forecasting for various services by frequency bands although it is essentially performed for building the management system. In order to perform market competition and demand forecast for various services provided by frequency bands, the reliability of forecasting results is one of the most important parts in the development of forecasting methods to manage frequency resources effectively.

Many researchers have investigated demand-forecasting models for new products. For the middle or long term demand forecasting, diffusion and growth curve models are frequently used. The Bass difussion model [1] in particular has been used extensively to investigate diffusion patterns and demand forecasts since it has a relatively high power of demand forecasting despite its simple structure.

Several modifications of the Bass model were introduced in the 1980s and beyond to reflect market complexities. These modified models incorporate concepts of product replacement and repeat purchases, substitution between generations, competition among products, marketing mix variables. A classification of the demand forecasting methods reported in the literature associated with the information and communication is shown in Table 1.

For the middle or long term demand forecasting, diffusion models and growth curve models are frequently used. In this paper, we demonstrate how to estimate reasonable and highly reliable forecasting data about demand of subscribers for service at the mobile communication market by applying diffusion and growth curve models.

2 Literature review

The initial work of the diffusion model probably began with Fourt and Woodlock [2] who considered the demand of the so-called "innovators" who adopt the product independently and not influenced by former adopters. Mansfield [3] considered the demand of people who are influenced by adopters, the so-called "imitators". In some situations, Bass [1] presented a model to consider both "innovators", representing the external influence in the adoption process and "imitators", reflecting the internal influence of previous adopters on those who have not yet adopted the product.

While the basic bass model can forecast the demand for a new product, it may not necessarily identify concepts of product replacement and repeat purchases, substitution between generations, competition among products, marketing mix variables, and so forth, in the diffusion process. Several variations of the Bass model were introduced in the 1980s and beyond to reflect market complexities and incorporate these concepts.

Some authors have made the saturation level (market potential) of a function of price, advertising or some other measure of market activity. An early example of this was given by Mahajan and Peterson [4] who used US housing starts to parameterize the market potential for washing machines. The impact of price on market potential has been studied by [4], [5], [6], and [7]. Robinson and Lakhani [8] were includec a price impact in the Bass model by reformulating the Bass hazard function. Horsky and Simon [9] examined the effect of advertising on the probability of adoption by modifying the hazard function. Thomson and Teng [10] proposed a model incorporating the effect of price impact and the effect of advertising on the probability of adoption. Generalizing an approach by Thomson and Teng [10], Simon and Sebastian [11] found that linking the imitation coefficient to advertising was the most effective way of using this information to a model associated with the diffusion of telephones. Kamakura and Balasubramanian [12] proposed a general model, which allowed price effects in the probability of adoption and in market potential.

[Table 1. A classification of the demand forecasting methods in the information and communication]

Life-cycle	Time series data	Explanatory variables	Forecasting model
Yes	Yes	Yes	Choice-based diffusion model, Advanced bass diffusion model
Yes	Yes	No	Bass diffusion model, Logistic growth curve model
Yes	No	Yes	
Yes	No	No	Analogy
No	Yes	Yes	Regression analysis
No	Yes	No	Box-Jenkins Time Series
No	No	Yes	
No	No	No	Delphi method, Logit choice model, Analytic Hierarchy Program(AHP)

Technology trend analysis model	Diffusion model	Choice model
Fisher-Pry model	Bass model	Binary choice model
Gompertz Model	Logistic model	Logit model

Source: KISTI (Korea Institute of Science and Technology Information)

The model is a generalization of the probability density function of Bass model. Bass et al. [13] developed a generalized Bass model (GBM) which allowed the incorporation of marketing mix variables in the modeling of new product diffusion.

Mahajan et al. [14] and Kamakura and Balasubramanian [15] incorporated repetitive or overlapping purchase effects of products and services into the Bass model to address the limitation that Bass model may not necessarily identify a repetitive or overlapping purchase of products and services in the diffusion process. Hahn et. al. [16] presented the repeated purchase diffusion model which considered cancellation of subscribers. Norton and Bass [17] proposed an adaptation of the Bass model that considered different generations of a technology. Mahajan and Muller [18] extended the Norton-Bass model to allow adopters of early generations to skip generations, for example, an adopter of the first generation could replace it with third generation technology [18],[19].

3. Background

The South Korea initially started a cellular service in the Seoul and the suburbs of the capital in May, 1984 and became a country to get a great power in the world's mobile communication market by commercializing first commercial CDMA and PCS cellular service in the world in October 1996 and 1997, respectively. Then, SK Telecom introduced commercial 'CDMA 2000 1x' service in Seoul and Inchon starting in October 2000 for the first time in the world. SK Telecom and KTF commercialized the launch of commercial synchronized IMT-2000 service (CDMA2000 1xEV-DO) with 800MHz bandwidth in Seoul in February, 2002. SK Telecom proved Korea as the leader of third-generation (3G) mobile network technologies with the kick-off of the world's first commercial HSDPA (High Speed Downlink Packet Access) service. OFDM or MIMO techniques as technical formulus of 4G services are discussed and a frequency band will be settled. If South Korea lead to 4G markets using OFDM or MIMO techniques, it will become the best leader of mobile network technology in the world.

The number of the mobile communication service subscribers reached 210 million in 2005 and exceeded 246 million in 2006 for reason of increasing in the number of subscribers in immature markets such as China and India. A number of mobile phone subscribers in South Korea reached 2,658 in 1984, 2 million when CDMA service was lunched in 1996, 30 million in 2002, exceeded 40 million recently, and reached 82% mobile phone users in March 2006.

4. Methodology

4.1 Bass-Diffusion-Model

The Bass model is one of the more well-known and widely used models of first-purchase demand. It is a model of the timing of adoption of an innovation and will be central to subsequent developments. The Bass model has a behavioral rationale that is consistent with studies in the social science literature on the adoption and diffusion of innovations and in based on a simple premise about the hazard function (the conditional probability that an adoption will occur at time t given that an adoption has not yet occurred). Thus, if f_t is defined as the probability of adoption at time t, or the fraction of the ultimate potential which adopts the innovation at time t and F_t is the fraction of the ultimate potential which has adopted by time t, the fundamental premise is that the likelihood of adoption at time t given that one has not yet occurred is [1]

$$f_t / [1 - F_{t-1}] = p + qF_{t-1} \quad (1)$$

If m is defined as the total number of initial purchases or as the market potential during the period of study, the number of adopters at time t is expressed as $n_t = mf_t$. Equation (1) can be written as:

$$n_t = dN_t / dt = [m - N_{t-1}][p + q\, N_{t-1}/m] \quad (2)$$

where nt is sales in period t, $N_{t-1} = \sum_{t=0}^{t-1}$ is cumulative sales through period t-1 (equal to the number of adopters of the product), m is market potential, equal to the total number of potential adopters, p and q are parameters.

Equation (2) can be rewritten as

$$n_t = p(m - N_{t-1}) + q\, N_{t-1}/m(m - N_{t-1}). \quad (3)$$

The first term in (3) represents the demand of the so-called "innovators", i.e. people who adopt the product independently and are not influenced by former adopters. The second term can be interpreted as the demand of people who are influenced by adopters, the so-called "imitators". Accordingly p and q are referred to as "innovation" and "imitation" coefficient respectively. A further elaboration on the behavioral substantiation of the model can be found in Bass [1]. The importance of innovators will be greater at first but will diminish monotonically with time, while the imitation effect will increase with time. Equation (3) leads to the differential equation:

$$n_t = pm + (q - p) N_{t-1} - q/m[N_{t-1}]^2. \quad (4)$$

If N(0) = 0 the solution to equation (4) is

$$N_t = m\left[1-\exp\left(-bt\right)\right]\Big/\left[1+a\exp\left(-bt\right)\right] \quad (5)$$

and the density function of time to adoption will be

$$n_t = m\left(b^2/p\right)\exp\left(-bt\right)\Big/\left[1+a\exp\left(-bt\right)\right]^2 \quad (6)$$

where $a = q/p$ and $b = p+q$. It is easy to differentiate n to find the time of the peak in n or the inflection point of N to be $t^* = (1/b)\ln(a)$ [18].

Mathematically the first term in (2) is an exponential model (Fourt and Woodlock [2]), whereas the second term represents a logistic model (Mansfield [3]). The logistic model produces a symmetric life cycle curve. Superimposing an exponential model with $\alpha > 0$ yields a positively skewed life cycle (fast growth, slow decline). If a life cycle is negatively skewed (slow growth, fast decline) α is bound to be negative. This case is not covered by Bass' behavioral theory. On the other hand we observe negatively skewed life cycles in the real world. This aspect is important for our later discussion. For a further elaboration on these issues we refer the reader to the excellent paper by Mahajan and Muller [18] and to the reader edited by Mahajan et al. [14]. We can estimate the number of subscribers at the mobile communication market via the parameters p, q, and m estimated in equation (6).

3.2 Logistic regression

Logistic regression model is nonlinear regression model for two important cases where the response outcomes are discrete and the error terms are not normally distributed and used when the response variable is qualitative with two possible outcomes, such as financial status (sound status, headed toward insolvency) or blood pressure status (high blood pressure, not high blood pressure).

Simple linear regression model is

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad Y_i = 0,\ 1$$

(7)

where the outcome Y_i is binary taking on the value of either 0 or 1. The expected response $E\{Y_i\}$ has a special meaning in this case. Since $E\{\varepsilon_i\}=0$ expectation of equation(7) is

$$E\{Y_i\} = \beta_0 + \beta_1 X_i.$$

(8)

Y_i is considered to be a Bernoulli random variable for which we can state $P(Y_i=1)= \pi_i$ or $P(Y_i=0) = 1- \pi_i$ when Y_i is 1 or 0, respectively. Thus, πi is the probability that Y_i =1, and 1- π_i is the probability that Y_i=0. By the definition of expected value of a random variable, Equation (8) can be rewritten as

$$E\{Y_i\} = 1\left(\pi_i\right) + 0\left(1-\pi_i\right) = \pi_i. \quad (9)$$

Equation (8) and (9) leads to following equation:

$$E\left\{Y_i\right\} = \beta_o + \beta_1 X_i = \pi_i. \quad (10)$$

The mean response $E\{\ Y_i\ \}=\beta_0+\beta_1X_i$ as given by the response function is therefore simply the probability that Y_i=1 when the level of the predictor variable is X_i. This interpretation of the mean response applies whether the response function is a simple linear one, as here, or a complex multiple regression one. The mean response, when the outcome variable is a 0, 1 indicator variable, always represents the probability that Y=1 for the given levels of the predictor variables.

Both theoretical and empirical considerations suggest that when the response variable is binary, the shape of the response function will frequently be curvilinear. Note that these response functions are shaped either as a titled S or as a reverse titled S, and that they are approximately linear except at the ends. These response functions are called logistic response functions and of the form:

$$E\{Y_i\} = \pi_i = \frac{\exp(\beta_o + \beta_1 X_i)}{1+\exp(\beta_o + \beta_1 X_i)} \quad (11)$$

where, Y_i are independent Bernoulli random variables with expected values $E\{\ Y_i\ \}=\pi_i$.

Interesting property of a logistic response function is that it can be linearized easily. If $E\{\ Y\}$ is expressed as π, the mean response is a probability when the response variable is a 0, 1 indicator variable. Transformed equation is

$$\pi' = \ln\left(\frac{\pi}{1+\pi}\right). \quad (12)$$

The logit mean response can be retransformed ln to linear function using $E\{Y\}$ as follows:

$$\pi' = \beta_0 + \beta_1 X\ . \quad (13)$$

The transformation (ln) is called the logit transformation of the probability π. The ratio π/(1-π) in the logit transformation is called the odds. The transformed response function (β) is referred to as the logit response function, and π` is called the logit mean response.

Since each Yi observation is an ordinary Bernoulli random variable, its probability distribution can be represented as follows:

$$f_i(Y_i) = \pi_i^{Y_i}(1-\pi_i)^{1-Y_i} \quad Y_i = 0,\ 1; \quad i = 1,\ldots,n \quad (14)$$

where $f_i(Y_i)$ simply represents the probability that Y_i =1 or 0.

Since the Y_i observations are independent, their joint probability function is

$$g(Y_1,\ldots,Y_n) = \prod_{i=1}^{n} f_i(Y_i) = \prod_{i=1}^{n} \pi_i^{Y_i}(1-\pi_i)^{1-Y_i}\ . \quad (15)$$

It will be easier to find the maximum likelihood estimates by working with the logarithm of the joint probability function

$$\log_e g(Y_1,...,Y_n)=\log_e \prod_{i=1}^{n} \pi_i^{Y_i}(1-\pi_i)^{1-Y_i}$$
$$=\sum_{i=1}^{n}\left[Y_i \log_e\left(\frac{\pi_i}{1-\pi_i}\right)\right]+\sum_{i=1}^{n}\log_e(1-\pi_i). \quad (16)$$

Hence, (log) can be expressed as follows:

$$\log_e L(\beta_o+\beta_1)=\sum_{i=1}^{n} Y_i(\beta_o+\beta_1 x_i)-\sum_{i=1}^{n}\log_e[1+\exp(\beta_o+\beta_1 x_i)] \quad (17)$$

where $L(\beta_0, \beta_1)$ replaces $g(Y_1,\ldots,Y_n)$ to show explicitly that we now view this function as the likelihood function of the parameters to be estimated, given the sample observations. The maximum likelihood estimates of β_0 and β_1 in the simple logistic regression model are those values of β_0 and β_1 that maximize the log-likelihood function in (log).

4. A case study

In this Section, forecasting of the number of subscribers is conducted by using both Bass and Logistic models applying each 1% within the range of 85%~90% of the estimated population as the saturated value of the Bass and Logistic models.

Forecasting of the gross number of subscribers from 2006 to 2012 based on the past data associated with the gross number of subscribers is initially performed. The results show a stable increasing pattern compared to the number of subscribers for each service provider. Then, forecasting of the number of each service provider subscribers by dividing the gross number of subscribers into the average market share (03`~05`) of each service provider is conducted.

5. Conclusion

In this paper, a number of current methodologies for demand forecasting of the mobile communication market are investigated, and a classification of the demand forecasting methods reported in the literature associated with the information and communication is performed. In addition, we demonstrate how to estimate reasonable and highly reliable forecasting data about demand of subscribers for service at the mobile communication market by applying diffusion and growth curve models. We then conduct a case study for long-term demand forecasting of the mobile communication market using the two models. The results of both models provide a similar increasing pattern. In order to achieve better precision on forecast data, the consideration of business strategies and new services effects can be a possible further research issue.

References

[1] F.M. Bass, A new product growth model for consumer products. *Management Science*, Vol.15, 1969, pp.215– 227.

[2] L.A. Fourt and J. W. Woodlock, Early prediction of market success for new grocery products. *Journal of Marketing,* Vol.25, 1960, pp.31– 38.

[3] E. Mansfield, Technical change and the rate of imitation., *Econometrica,* Vol.29, 1961.

[4], V. Mahajan, and R. A. Peterson, Innovation diffusion in a dynamic potential adopter population. *Management Science*, Vol.15, 1978 , pp.1589– 1597.

[5] F.M. Bass and A.V. Bultez, A Note on Optimal Strategic Pricing of Technological Innovations, *Marketing Science*, Vol.1, 1982, pp.371-378..

[6] S. Kalish, A New Product Adoption Model with Price, Advertising, and Uncertainty. *Management Science,* Vol.31, No.12, 1985, pp.1569–1585.

[7] D. Horsky, A Diffusion Model Incorporating Product Benefits, Price, Income and Information, *Marketing Science*, Vol. 9 , 1990, pp.342-365.

[8] B. Robinson and C. Lakhani, Dynamic price models for new-product planning, *Management Science* , Vol.21, 1975, pp.1113–1122.

[9] Horsky and L.S Simon,A dvertising and the diffusion of new products. *Marketing Science,*Vol. 21, 1983, pp.18

[10] R. Thompson, J. Teng, Optimal pricing and advertising policies for new product oligopoly models. *Marketing Science,* Vol.2, 1984, pp.148–168,

[11] H. Simon, k. Sebastian, Diffusion and advertising: The German telephone Campaign, *Management Science*, Vol.33, 1987, pp.451-466

[12] W. Kamakura, S. Balasubramanian, Long-term view of the diffusion of durables: A study of role of price and adoption influence processes via test of nested models. *Int J Res Marketing* , Vol.5, 1988, pp.1–13

[13] F.M. Bass, T.V. Krishnan, D. Jain, Why the Bass model fits without decision variables. *Marketing Science*, Vol.13, 1994, pp.203–223.

[14] Mahajan, Vijay, Jerry Wind, and Subhash Sharma, An Approach to Repeat-Purchase Diffusion Analysis, *Educators' Conference Proceedings, Chicago, IL: American Marketing*

Association, No. 49, 1983, pp.442–446.

[15] W. Kamakura, S. Balasubramanian, Long-term forecasting with innovation diffusion models: The impact of replacement purchases. *Journal of Forecasting*, Vol.6, 1987, pp.1–19.

[16] M. Hahn, S. Park, L. Krishnamurthi, A. Zoltners, Analysis of new-product diffusion using a four-segment trial-repeat model. *Marketing Science,* Vol.13, 1994, pp.224–247.

[17] J. C. Fisher, and R. H. Pry, A simple substitution model of technological change. *Technological Forecasting and Social Change,* Vol. 3, 1971, pp.75– 88.

[18] Mahajan, V., and Muller, E. Innovation, diffusion and new product growth models. Journal of Marketing 43: 55– 58 (Fall 1979).

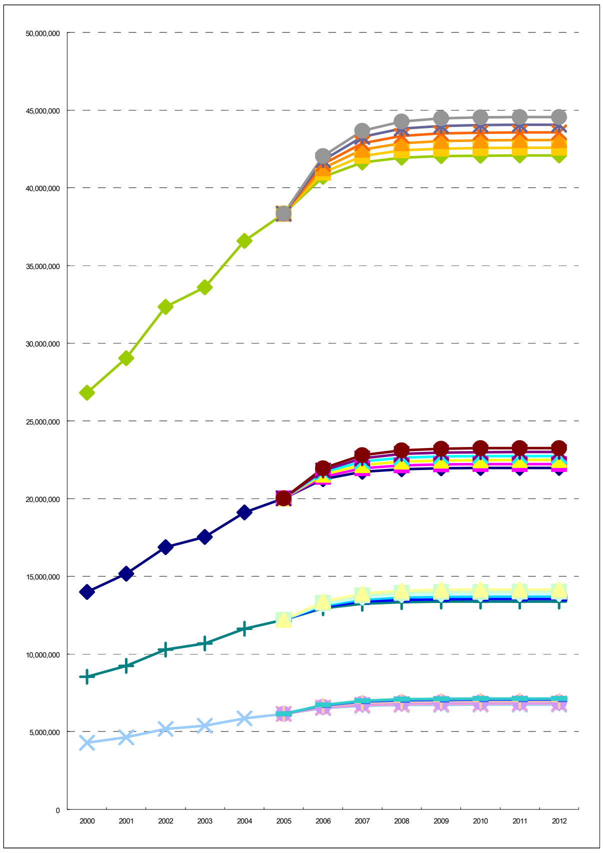

Fig. 1. The results of Bass model

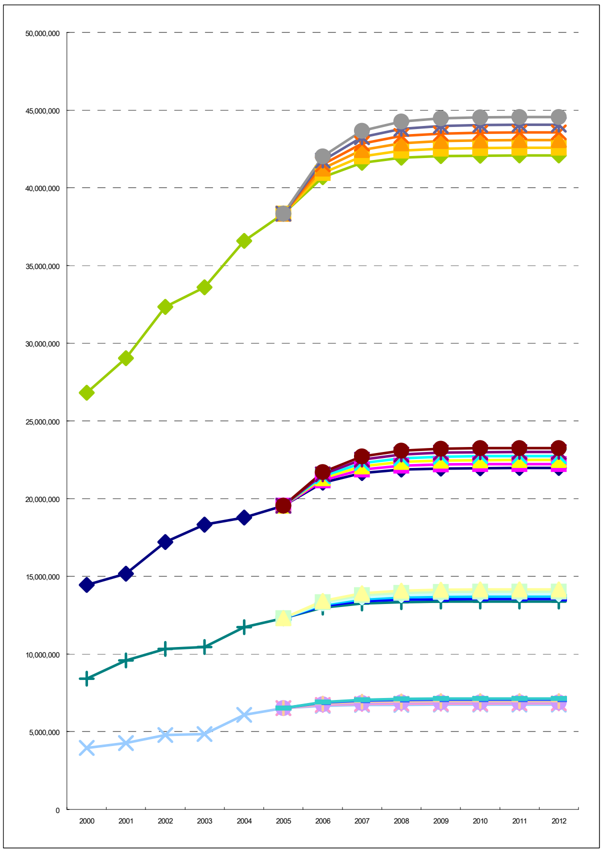

Fig. 2. The results of Logistic model