URSI AP-RASC 2019, New Delhi, India, 09 - 15 March 2019

A Competitive Hedonic Consumption Estimation for IoT Service Distribution

Deepsubhra Guha Roy * (1), Bipasha Mahato (1), and Debashis De(1)(2)
(1) Centre of Mobile Cloud Computing, Department of Computer Science and Engineering, Maulana Abul Kalam Azad University of Technology, WB, India
(2) Department of Physics, University of Western Australia, 35 StirlingHwy, Crawley WA 6009, Australia

Abstract

Cost estimation of IoT service subscription is directly associated with the quality of experience, upload/ download speed of a subscription link. This article consider a dynamic Adaptive Bit Rate (ABR) streaming based wireless communication service to estimate hedonic consumption of bandwidth distribution for IoT service subscription. Hedonic price model has exploited jointly with Semi-Log estimation and Lagrange multiplier tests. Judgment of the total estimation has concluded that the value of internet information is totally dependent upon the user- willingness to pay and elasticity of individual service variable. Based on the result of hedonic price estimation it has established that ISPs can charge on the basis of IoT service usage over wireless connectivity, instead of offering all the services at fixed cost.

1. Introduction

In our country it is quite tough to correctly define the internet service usage motive where the total number of broadband subscriber's number is 276.52 million, with a 5.82 percent monthly growth rate, till the end of march 2017(source data published by TRAI). There are five service providers holding 87.48 percent of total market share. Market is too uncertain and to choose a broadband provider is more depending upon the speed than the reliability. We select a comparatively small but actively using IoT service area, for survey with 310 different data sets to collect day on basis information to support our formulas first. Then we follow the SLA (Service Level Agreement) to choose the hedonic specification.

2. System model

Four probable function specifications are there in the Hedonic model, describe following:

A. Linear specification: Illuminating and dependent variables both go through the regression equation with linear form.

$$P_H = \beta_0 + \sum_{n=1}^{N} \beta_n x_n + \varepsilon \,; \tag{1}$$

P_H is the property value, obtain from β_n (n=1,2,..,N): unit price adjustment of the nth characteristic (x_n)of a particular good. ε is a random error term vector.

B. Semilog specification: Either the dependent variable or the explanatory variable is in log form. And the other one is in linear form into a regression equation.

$$\ln(P_H) = \ln(\beta_0) + \sum_{n=1}^{N} \beta_n x_n + \varepsilon \,; \tag{2}$$

P_H is the property value, obtain from β_n (n=1,2, ..,N): specifies the rate of the price at which the cost value raises at a assured point, ε is a random error term vector.

C. Log–log specification: Together the dependent variables and explanatory are in logarithm form.

$$\ln(P_H) = \ln(\beta_0) + \sum_{n=1}^{N} \beta_n \ln(x_n) + \varepsilon \,; \tag{3}$$

P_H is the property value, obtain from β_n (n=1,2,.,N): increasing value of P_H, in percent at a certain marginal point when the nth characteristic x_n upgrades by 1 %, ε is a random error term vector.

D. Box–Cox transform: Before entering the regression form, the problem can itself identify its best individual transformed form shown in equation (4).

$$P_H[\lambda] = \beta_0 + \sum_{n=1}^{N} \beta_n x_n^{(\lambda_n)} + \varepsilon \,; \tag{4}$$

Box- Cox conditions applying over equation (4) are:

$$P_H[\lambda] = \frac{P_H^{(\lambda)} - 1}{\lambda}, \lambda \neq 0; \tag{5.1}$$

$$= \ln(P_H), \theta = 0; \tag{5.2}$$

λ stands for Box-Cox coefficient.

Downlink bandwidth allocation:

Consider, total bandwidth allocation is ≤ 1 by a broadband supplier at time slot t. Then fractional bandwidth $p_n(t)$, allocated by the base station to an individual user can be denoted by:

$$p_n(t) \in [0,1], \forall n, t \,;$$

$$0 \leq \sum_{n=1}^{N} p_n(t) \leq 1, \forall t \,;$$

$$C_n(t) = B * n_k(t) * \log_2(\frac{1+Pgn(t)}{P_n(t) * B * P_0}, \forall n,t \quad (6)$$

Total available bandwidth in the network is B. Channel gain is defined by $g_n(t)$ at the time slot t, during the downlink data flow towards the user from the base station. In the presence of constant transmit power P, the power spectral density of the base station is denoted by P_0, with additive white Gaussian noise.

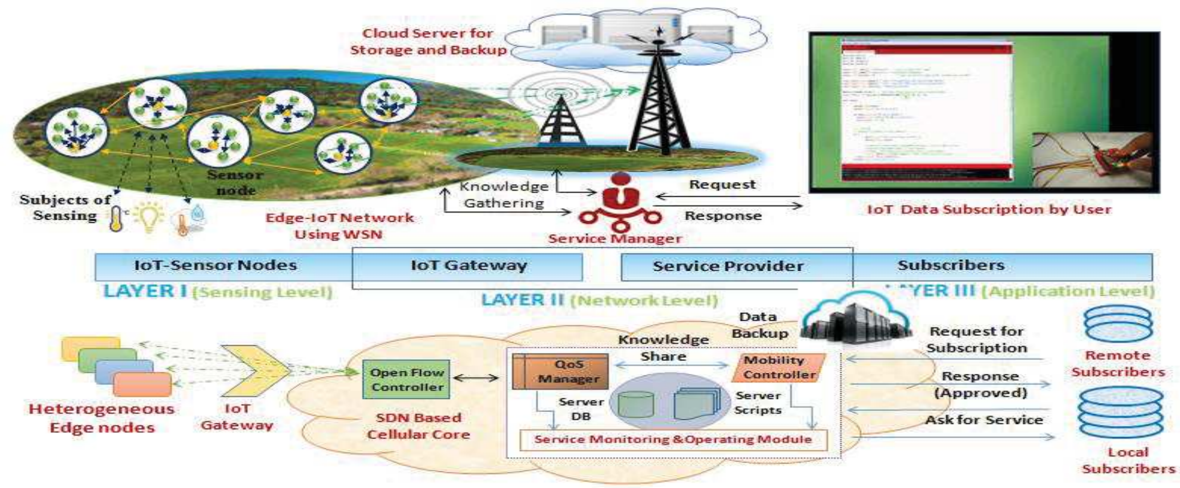

Figure 1: Three Layered Architecture of Edge-IoT Service Selection Method

3. Performance analysis

A. Data gathering: We consider a spatially dynamic average process (SDA) for $W:(N \times N)$ spatial matrix. In simple way we assume that the consumption matrix contains $(N \times N)$ elements, representing estimated likelihood of the subscriber's demand. Specifically, R numbers of component relations indicating to M different regions, observed near about T time period are individually can be represented as:

$$x_{it} = \gamma_i W x_{it} + Y_{it}\beta_i + \varepsilon_{it} \; ; \quad (7)$$

Where, i = 1,2, …, R and t = 1, 2, …, T;

B. QoS measurement motivations: Application or web service layer is the most top and near to subscribers' layer. In this layer some service parameters has considered to build the proposed model are following:

1. *Cost of service* $C(\beta)$: Service cost generally denotes the service providing cost or service subscribing cost. Per unit service cost is usually constant or fixed which is offered by the end service provider.
2. *Service time* $I(\beta)$: The request compilation time with acceptance of probable output from the time when the query was submitted to the network by a user is considered as the service time. $I(\beta)$ varies over some network related factors, mostly upon network resource allocation and process capability etc. The inspected service time can be defined as: $E(I(\beta)) = (\pi\rho_k)/((1-\pi)^2 \varepsilon + \frac{1}{\mu})$, against k numbers of IoT service nodes for $N/N/k$ queue model.
3. *Service Load* $L(\beta)$: Load of the network is measured through the usage of provided IoT service by various IoT consumers at an instant of time. It is measured by $L(\beta) = \varepsilon\mu$,where ε is the arrival request-to-response rate and μ is the providing service rate.
4. *Service reliability* $\mathrm{Re}(\beta)$: Probable failure to complete a requested service by the IoT service provider on the edge node is denoted as reliability, which is modelled as $\mathrm{Re}(\beta)$, a random variable, measured by $\sum_{k=1}^{N} \frac{x_i}{n}$ and D.
5. *Service reputation* $Rp(\beta)$: Reputation of an IoT node deadly depends upon the edge-level trust factor. It increases with the better quality of experienced IoT service at the edge-node.

C. Data filtering: Spearman rank correlation coefficient (SRCE) has applied to filter out the effective 310 set of samples data, among the collected data set of 52 samples. We consider SRCE, a neighborhood formation method to select each data based on their individual current status rather than the outcome of the whole scenario while choosing relevant data to create a homogeneous data set.

$$SRC(f_a, f_b) = \frac{\sum_i^N (Rank_{ai} - \overline{Rank_a})(Rank_{bi} - \overline{Rank_b})}{\sqrt{\sum_i^N (Rank_{ai} - \overline{Rank_a})^2}\sqrt{\sum_i^N (Rank_{bi} - \overline{Rank_b})^2}} \quad (8)$$

Two data matrixes are a and b. Ranks are provided to a and b according to their effective data base and survey results. Considering x_{it} a $(N \times 1)$ vector matrix, where Y_{it} holds $(N \times k_i)$ exogenous variables, β_i is a $(k_i \times 1)$ parameter vector and ε_{it} holds N number of instabilities or disturbances, typically a $(N \times 1)$ vector. With the allowance of some general spatial lags for the other endogenous variables, we also refer the spatial lag of x_{it} as $\bar{x}_{it}(=Wx_{it})$. Then the equation (7) becomes:

$$x_{it} = \gamma_i \bar{x}_{it} + Y_{it}\beta_i + \varepsilon_{it} \quad \textbf{(9)}$$

In particular, we consider only the instabilities due to spatially dynamic average process (SDA). Therefore ε_{it} can be represented as: $\varepsilon_{it} = \lambda_i W u_{it} + u_{it}$ **(10)**

Where the error component u_{it} can be calculated from:

$$u_{it} = \mu_i + v_{it} \quad \textbf{(11)}$$

For total time period T , we can get from equation (9)

$$x_i = \gamma_i (E_T \otimes W)x_i + Y_i \beta_i + \varepsilon_i \quad \textbf{(12)}$$

Where $\varepsilon_i = \lambda_i (E_T \otimes W)u_i + u_i$ **(13.1)**

With $u_i = (\gamma_T \otimes E_N)\mu_i + v_i$ **(13.2)**

Where $\mu_i = (\mu_{11},..., \mu_{iN})'$, $v_i = (v_{i11},..., v_{iNT})'$ and γ_T is a vector representation of $(T \times 1)$ according to Anselin, Le Gallo and Jayet (2008).Therefore, from equation (8) we can say:

$$x = (\Gamma \otimes E_T \otimes W)x + Y\beta + \varepsilon \quad \textbf{(14.1)}$$

And $\varepsilon = (\wedge \otimes E_T \otimes W)B\varepsilon + B\varepsilon$; **(14.2)**

Where $Ax = Y\beta + \varepsilon$ with $A_i = E_M - \gamma_i W$ **(15.1)**

$$B_i = (E_N + \lambda_i W)^{-1} = L_i^{-1}; \quad \textbf{(15.2)}$$

A and B matrixes are given belowand variance-covariance matrix ε is referred from [8]

$$A = \begin{pmatrix} E_T \otimes A_1 & & \\ & \cdot & \\ & & \cdot \\ & & & E_T \otimes A_R \end{pmatrix} \text{ and } B = \begin{pmatrix} E_T \otimes B_1 & & \\ & \cdot & \\ & & \cdot \\ & & & E_T \otimes B_R \end{pmatrix};$$

Therefore the log-likelihood function designed for SDA of x observation is possibly proportion to the following.

$$\ell \propto -\frac{1}{2}\sigma|\Omega_u| + \ln|B| + \ln|A| - \frac{1}{2}a'a;$$

Where Ω_u , an $(M \times M)$ error matrix, is denoted by

$$\Omega_u = \begin{pmatrix} \omega_{u1}^2 & \omega_{u12} & \cdot & \omega_{u1R} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \omega_{uR1} & \omega_{uR2} & \cdot & \omega_{uR}^2 \end{pmatrix}$$

And according the joint standard normal distribution,

$$a'a = (A_x + Y_\beta)'\Omega\varepsilon^{-1}(Ax - Y_\beta) = \varepsilon'\Omega\varepsilon^{-1}\varepsilon; \textbf{(16)}$$

To analysis the results we need to make all the collected data equally effected. Therefore we have applied two rules discussed later and the results turn out exclusively shown in table 1.

Rule 1: Leased line capacities vary slightly varies with different region. As an example, a L_1 procession in an inhabitant has a ability of x_1 Mbps while L_2 line in another part of the city has a capacity of x_2 Mbps. With the purpose to match bit rates, the cost of the L_1 procession is accustomed rising proportionately while evaluating the two services. In this case, the price distribution has applied on the base of equations (17.1), (17.2) and (17.3).

$$L_1(x_1 Mbps) = P^{(L1)} \quad \textbf{(17.1)}$$

$$L_2(x_2 Mbps) = P^{(L2)} \quad \textbf{(17.2)}$$

$$L_1(Adjusted \quad to \quad x_2\ Mbps) = P^{(L1 \quad Adjusted)}$$

$$= P^{(L1)} * (\frac{x_2}{x_1}); \quad \textbf{(17.3)}$$

Rule 2: In a situation, when d_1 Mbps service is not responsive, d_2 Mbps is used as a proxy. In these cases the price of the d_2 Mbps is attuned downward to imitate the oblique per-Mbps penalty of d_1 Mbps. Equation (18.1) gives an example of estimate subscription value due to applying this condition.

$$P^{(L1 \quad Adjusted)} = P^{(L2)} * (\frac{d_1}{d_2}) = P_{subs}; \quad \textbf{(18.1)}$$

Suppose internet usage duration is T seconds, and it ends at t seconds before the duration of one unit usage whenever it comes to a point to give the service a number of N users at a time slot. A developed empirical function has considered as: w

$$T_{\%} = \frac{\log(10 * N^{L1})}{\log(10 * t)}; \quad \textbf{(18.2)}$$

Where t is the total time (in minutes) duration and N is the services volume, nominated user number to provide service at a time slot.

Maximum Likelihood Estimation: Previously designed log-likelihood function for SDA is also can be represented following:

$\ell \propto -\frac{1}{2}(A_x - Y_\beta)'B'(\sum_\mu^{-1} \otimes E_M)B(A_x - Y_\beta)$ This is established with the help of the results from Magnus (1982), followed by Baltagi (1980) respectively.

4. Numerical studies

Hedonic prising function: The hedonic prising function is

$$P_H = f(S,U,V); \quad \textbf{(19)}$$

S, U and V individually stand for Services, Usage and Conditions of internet subscription respectively. These are the variables which are consisting of some variables each. Say, service(S) quality depends upon number of family members, residential location, upload/download speed, whereas usage (U) of internet depends upon some purpose of subscription like web browsing, email checking, social site accessing, game downloading (heavy or light in nature is also under observation). Other conditions like contract with provider, registrations are under condition variable V. We also have gone through some detailed study. Like how many users are male or female, age group of the users, or most of the time the data upload/ download speed of the internet connection etc. Therefore our estimation follows the semi log function derives from equation (2) and equation (19) is shown in equation (20).

$$\ln(P_H) = \ln(P_{subs}) + \sum \beta_i S_i + \sum \beta_j U_j + \sum \beta_k V_k + \varepsilon \qquad \textbf{(20)}$$

We calculate possible mean and standard deviation of each variable. In our survey users were 49.9% from metro city and rest 50.1% were from outside of the city. 14.1% users having 2 mbps internet speed, where 13.4% with 5 mbps upload/download speed, 52.2% enjoying 12 mbps data rate and rest 20.3% entertaining by 45 mbps of data rate. With the help of mandatory registration process most of the users 82.9% are registered internet subscribers. Table 1 shows the variable dependencies after regression co-efficient analysis of sample matrix $(N \times N)$. Sample matrix analysis showing how much bandwidth is consumed by the sample users, however it does not showing that how much they are interested to give extra price for individual usage. We suggest calculating willing to pay factor to fill this gap which is further concluded with elasticity measurement.

Table 1: Regression analysis table for $(N \times N)$ sample matrix

Variable	$C(\beta)$	$I(\beta)$	$L(\beta)$	$\mathrm{Re}(\beta)$	$Rp(\beta)$
Co-efficient	-24.43	0.465	0.207	0.436	-0.218
R-value	0.683	0.821	0.398	0.151	0.752
Sig. P	0.000	0.000	0.0001	0.031	0.000

Degree of freedom (df) = Total sample -2=31-2=29.

Willing to pay estimation: Benefit valuation of non-traded or consumable products are analysed in the term of economic assessment with the help of willing to pay (WTP) estimation. Here we have done this with the help of equation (20).

$$W = y + \alpha_1 X + \alpha_2 P_H + \varepsilon \qquad \textbf{(21)}$$

Here, WTP price W depends upon y : yes or no response; X is the vector among the variable set of (S, U, V) reflecting internet usage characteristics. P_H is the estimated bid price and ε is an error term calculated by the equation (14.2). Mean WTP has derived from the expression:

$(\sum(\alpha_1 * X^a)/\alpha_2) * -1)$; Where X^a is the mean value of the variable $X \in \{S, U, V\}$ in our study it turns into:

$y + (\sum(\alpha_1 * X^a)/\alpha_2) * -1)$;

Result analysis shows that users are willing to pay positively for some instance like basic service price (1.13) and speed of accessing, and some of them are unlike to pay for variables like service delay (-2.435), load etc. Meanwhile the overall WTP factor is nearly 27.66, which supports the providers to invest their money.

Elasticity measurement: To measure elasticity of a variable we choose a simple way explain into the equation (22).

$$E = \frac{\%\ \ change\ \ in\ \ P_H}{\%\ \ change\ \ in\ \ \mathrm{var}ible} \qquad \textbf{(22)}$$

$$= \frac{Coefficient\ \ of\ \ \mathrm{var}iable}{\overline{P}_H} * \overline{Value}_{Var}$$

Here E stands for elasticity.

Table 2: Subscriber's average willingness-to-pay

Variable	Co-efficient	Elasticity (%)	WTP (Rupees)
$C(\beta)$	-24.43	1.826	1.13
$I(\beta)$	0.465	-0.001	-2.435
$L(\beta)$	0.207	0.011	11.764

Table 2 shows that the average price elasticity of offer price by the providers is 1.826, means users are willing to pay extra 1.13 rupees in Indian currency while one percent increases of internet usage, keeping others variable unchanged. Others results shows elasticity and willing to pay values over changing one percent of internet usage while others variables are remain constant. Note that price elasticity is the estimation over changing one service variable; which does not stand for the demand elasticity.

5. Conclusion

Current study is a small hop to achieve a pre judgment of country person's preference over a small community study. This paper concludes that the users are willing to pay for some of the variables like web browsing, e-shopping etc, while they are not will to being extra charged for some variables like e-mail etc. Considering the hedonic consumption and with the help of regression analysis we are able to give the direction to the ISP, to charge for specific services .Therefore we are hoping to make the study exhaustive with the estimation of heterogeneous population likelihood in our future work.

References

1. Guo, Yashuang, et al. "Dynamic Quality Adaptation and Bandwidth Allocation for Adaptive Streaming over Time-Varying Wireless Networks." IEEE Transactions on Wireless Communications (2017).

2. Monson, Matt. "Valuation using hedonic pricing models." Cornell Real Estate Review 7.1 (2009): 10.

3. Xiao, Yang. "Hedonic Housing Price Theory Review." Urban Morphology and Housing Market. Springer Singapore, 2017. 11-40.

4. Baltagi, Badi H., and Georges Bresson. "Maximum likelihood estimation and Lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in Paris." Journal of Urban Economics 69.1 (2011): 24-42.

5. Wu, Caesar, et al. "Hedonic Pricing of Cloud Computing Services." IEEE Transactions on Cloud Computing (2018).