A Method of QoE Management in Online Shopping Web Services with TCP Variables

Daisuke Yamauchi, Yoshihiro Ito and Toru Tahara
Department of Computer Science and Engineering, Graduate School of Engineering
Nagoya Institute of Technology, Nagoya, Japan

Abstract—**For service level agreement over company networks or community ones, this paper proposes a simple method of QoE evaluation with objective TCP variables in an online shopping Web service for SLM. It also confirms the effectiveness by experiment. In this experiment, the authors measure 31 variables concerned with TCP performance and assess user's satisfaction as a QoE parameter for three actual shopping Web services. Then, to clarify relationship between the user's satisfaction and the objective TCP variables, they utilize multivariate statistics: the principal component analysis and the multiple regression analysis. The experimental results show that the user-satisfaction can be easily estimated with three or four objective TCP variables in the online shopping service.**

***Index Terms* — SLM, QoE, QoS, TCP variables, user satisfaction, online shopping service, multivariate statistics**

I. INTRODUCTION

We can currently utilize various Web services over the Internet. The services have been including the traditional and social ones, such as educational services, administrative ones and so on. (For instance, see [1] and [2], respectively.) In other words, Web services have been indispensable for our life in these days.

The more Web services have played an important role in our society, the higher quality of them is required. Indeed the quality can be evaluated from many points of view, for many services, such as online shopping services, it is important to assess the user-centric quality like user satisfaction, namely Quality of Experience (QoE). This reason is that a Web service that gives the higher QoE can attract the more customers.

Since QoE is decided by a user, it generally takes the high cost to manage how much QoE of a Web service is provided by the service. Thus, it is required to get a good estimate of QoE from some related objective variables, which we can obtain without any users' effort; moreover the variables must be measured easily.

For some Web service, many factors can affect its QoE. One of the most influential factors is how Web pages are designed [3][4]. However, since any Web service is provided via communication between a Web server and a Web client, quality deterioration of the communication also degrades the QoE. Especially, in SLA (Service Level Agreement) among company networks or community ones, its network service provider has to consider the QoE deterioration caused by the communication degradation in the Web service. ITU-T G.1010 [5] and G.1030 [6] refer to QoE for a Web service, but they are not necessarily appropriate to current complicated Web services. Thus, ITU-T SG 12 has just standardized recommendation for QoE of Web-browsing as G.1031 [7].

On the other hand, HTTP [8], which is the application layer protocol of Web services, employs TCP [9] as its transport layer protocol. This means that TCP performance can become one of serious factors that affect user-satisfaction of a Web service. Thus, it is feasible to treat variables concerning TCP performance as the objective ones mentioned above.

Many TCP implementations have been developed [10][11][12]. In these great studies, we can see that these implementations can give good TCP performance under a variety of environments. However, many functions provided by the implementations, such as congestion control, make behavior of TCP very complex. Consequently, there are many variables that are closely connected with TCP performance. For example, we can imagine a mean of throughput or that of response time to be one of such variables. Indeed, in many Web services, they are a good index of estimation of QoE but they are not always enough for some Web service.

In this paper, we study a simple estimation of QoE from TCP variables in online Web shopping services by experiment. In our experiment, we treat a user-satisfaction for actual Web services as a QoE parameter and assess it; then we measure many variables related to TCP. In order to clarify the relationship between the TCP variables and the user-satisfaction, we utilize the multivariate statistics: the multiple regression analysis and the principal component analysis [13].

The rest of this paper is organized as follows. Section II describes user satisfaction of Web usability. We discuss evaluation of TCP performance in Section III. Sections IV and V show our experiment and its results, respectively. Finally, in Section VI, we conclude our paper.

II. QOE PARAMETERS OF WEB SERVICES

In [3] and [4], (user-)satisfaction is one aspect of Web usability and not always represented by a single scale. However, [3] and [4] do not define a concrete scale and a methodology for its evaluation.

For the sake of simplicity, this paper deals with only one scale indicating how users are satisfied with a service. To measure the scale, we use the rating scale method [14]. It is one of the psychophysical methods. In the rating scale method, subjects (users) classify objects for evaluation into five or seven categories. From a result of the rating scale method, we can calculate MOS (Mean Opinion Score), which is often used for assessing quality of various media, such as audio [15] or video [16]. For the first step of our research, this paper also makes use of MOS as a scale for assessment of a degree of user's satisfaction in a Web service and calls the scale "user-satisfaction" for convenience's sake.

III. TCP VARIABLES

Since current TCP implementations have many functions, behavior of TCP has been very complex. We therefore need to consider TCP performance with various variables. Reference [17] defines extended performance statistics for TCP. Thus, among the variables defined in [17], we choose 29 ones that concern with TCP data transmission. In order to measure them, this paper utilizes the Web100 software [18]. The Web100 software implements instruments in the Linux TCP/IP stack; it is distributed in two pieces: a kernel patch adding the instruments, and a suite of libraries and tools for accessing the kernel instrumentation. Table I shows names and short descriptions of the 29 variables we selected. For further details of them, see [17][18]. In addition to the above 29 variables, we also take care of two general variables: an average of TCP throughput and that of TCP response time.

IV. EXPERIMENTS

A. Experimental environment

Figure 1 shows our experimental environment. A user (subject) gets an actual Web service with a Web client PC that is connected with the Internet via a network emulator and a Web proxy server. The Web100 software is implemented in the Web client so that the 29+2 variables shown in the previous section can be measured. We can add delay and packet loss between the Web client PC and the proxy server by the network emulator. The specification of the Web client PC and that of the network emulator are shown in Table II and Table III, respectively. The Web browser on the Web client PC is Firefox [19]. This paper uses Dummynet [20] as our network emulator. In this experiment, we set additional round trip delay and packet loss; the fixed round trip delay is 150 msec and fixed packet loss ratio is 10%.

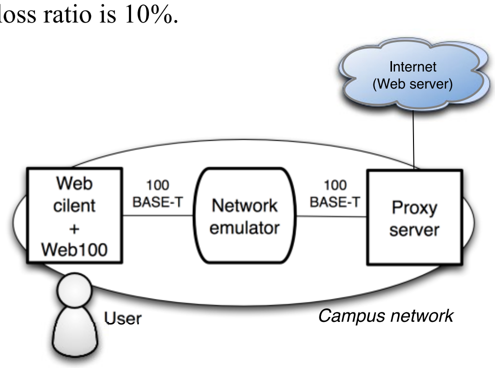

Figure 1. Experimental environment.

TABLE I. VARIABLES FOR TCP PERFORMANCE EVALUATION

Name of variable	Short description
SegsIn	The total number of segments received.
DataSegsIn	The number of segments received containing a positive length data segment.
DataOctesIn	The number of octets contained in received data segments including retransmitted data.
SegsOut	The total number of segments sent.
DataSegsOut	The number of segments sent containing a positive length data segment.
DataOctetsOut	The number of octets of data contained in transmitted segments, including retransmitted data.
SampleRTT	The most recent raw round trip time measurement used in calculation of the RTO (retransmission timer).
SmoothedRTT	The smoothed round trip time used in calculation of the RTO. See [21].
RTTVar	The round trip time variation used in calculation of the RTO. See [21].
MaxRTT	The maximum sampled round trip time.
MinRTT	The minimum sampled round trip time.
SumRTT	The sum of all sampled round trip times.
CountRTT	The number of round trip time samples included in SumRTT.
SegsRetrans	The number of segments transmitted containing at least some retransmitted data.
OctesRetrans	The number of octets retransmitted.
CurRTO	The current value of the retransmit timer RTO.
MaxRTO	The maximum value of the retransmit timer RTO.
MinRTO	The minimum value of the retransmit timer RTO.
CongAvoid	The number of times the congestion window has been increased by the Congestion Avoidance algorithm.
CurCwnd	The current congestion window in octets.
MaxSsCwnd	The maximum congestion window used during Slow Start in octets.
DupAcksIn	The number of duplicate ACKs received.
DupAcksOut	The number of duplicate ACKs sent.
SlowStart	The number of times the congestion window has been increased by the Slow Start algorithm.
CurSsThresh	The current slow start threshold in octets.
MaxSsthresh	The maximum slow start threshold, excluding the initial value.
MinSsthresh	The minimum slow start threshold.
SacksRcvd	The number of SACK options received.
SackBlocksRcvd	The number of SACK blocks received (within SACK options [22]).

In this experiment, we assigned the users into two groups: Group A and Group B. We don't add any delay and any packet loss by the network emulator for the users assigned Group A while we add additional delay and packet losses for users in Group B. The users are male and female and their ages were 20s; the number of users is 32. Both Group A and Group B have different 16 users.

TABLE II. SPECIFICATION OF THE WEB CLIENT PC

Model	Dell Vostro Desktop 220S
OS	Linux 2.6.27-43vl5
CPU	Intel Celeron 2.00GHz
Memory	2 GByte
Network Interface	100 BASE-TX
Monitor	19 inch TFT

TABLE III. SPECIFICATION OF THE NETWORK EMULATOR

Model	Dell Vostro Desktop 200S
OS	FreeBSD 7.0
CPU	Intel Core2 Duo E7200 2.53GHz
Memory	2 GByte
Network Interface	100 BASE-TX
Model	Dell Vostro Desktop 200S

B. Web service and users' task

This paper considers three actual online Web shopping sites they are very popular in Japan.

As an experimental task, each user searches some commodities specified by the experimenter, and puts them into his/her cart every Web site.

C. Assessment of a user satisfaction

As mentioned in Section II, we adopt the rating scale method to assess a scale that indicates how much users are satisfied with a Web service. In the rating scale method, we utilize five categories of impairment for given Web services: "imperceptible", "perceptible, but not annoying", "slightly annoying", "annoying", and "very annoying". The categories have a score, 5, 4, 3, 2, and 1, respectively. We calculate an average of category scores voted by all the users and treat the average as a user-satisfaction.

V. EXPERIMENTAL RESULTS

A. An estimate of a user-satisfaction from two general TCP variable

First of all, let us estimate a user-satisfaction from the measurement of the two general variables, i.e., average TCP throughput and average TCP response time in the services. For the estimation, we use the multiple regression analysis [15]. Namely, we consider the two variables and the user satisfaction as predictor variables and a criterion variable, respectively. In the multiple regression analysis, if we have n predictor variables $(\alpha_1,\cdots,\alpha_n)$, an estimate of criterion variable $\hat{S}$ can be calculated by

$$\hat{S} = \beta_0 + \beta_1\alpha_1 + \cdots + \beta_n\alpha_n \quad (1)$$

where $\alpha_i (1 < i < n)$ and β_i are the i-th predictor variable and its partial regression coefficient, respectively; β_0 is an intercept.

From the result of the multiple regression analysis, the obtained regression line is not statistically significant under the alpha level of $p < .05$. This means that, in this service, its user satisfaction cannot be predicted by only the two general variables.

B. TCP performance

In a similar way to the previous subsection, we try to estimate the user satisfaction from the 29 variables described in Section III and the 2 general variables, that is, an average TCP throughput and an average TCP response time. Among the 31 variables, however, some variables can correlate with each other. Then, to prevent multi-colinearity caused by the correlation among the variables, we perform *the principal component analysis (PCA)* [15] for the variables before the multiple regression analysis to select necessary variables as predictor ones.

The result of PCA shows that the cumulative contribution rate of the first four principal components becomes 89 % but that of the first five principal components is 94 %. Since the fifth principal component does not contribute to increase of the cumulative contribution rate very much, we handle the first four principal components. We show the eigenvectors for the first four principal components in Table IV.

From Table IV, we find that the first principal component is connected with some variables, e.g. the total number of received segments, the number of times the congestion window has been increased by the Slow Start algorithm, and so on. On the other hand, the second principal component correlates with the number of retransmitted segments, the sum of all sampled round trip times. Similarly, the third principal component and the fourth one seem to correlate with some scales each other.

TABLE IV. EIGENVECTORS OF THE FIRST FOUR PRINCIPAL COMPONENT

Variable	1st principal component	2nd principal component	3rd principal component	4th principal component
SegsIn	-0.823	-0.414	-0.313	0.180
DataSegsIn	-0.822	-0.420	-0.309	0.176
DataOctesIn	-0.822	-0.402	-0.325	0.187
SegsOut	-0.809	-0.505	-0.245	0.121
DataSegsOut	-0.809	-0.505	-0.245	0.121
DataOctetsOut	-0.807	-0.412	-0.307	0.183
SampleRTT	-0.020	-0.831	0.452	-0.259
SmoothedRTT	-0.160	-0.755	0.471	-0.307
RTTVar	0.713	-0.310	-0.616	-0.034
MaxRTT	0.732	-0.259	-0.621	0.001
MinRTT	0.730	-0.215	-0.637	0.016
SumRTT	-0.118	-0.192	0.074	0.610
CountRTT	0.518	-0.767	-0.189	-0.259
SegsRetrans	0.106	-0.934	-0.004	-0.281
OctesRetrans	-0.822	-0.434	-0.293	0.163
CurRTO	0.731	-0.221	-0.635	0.014
MaxRTO	0.729	-0.238	-0.631	0.006
MinRTO	0.713	-0.434	-0.293	0.163
CongAvoid	-0.834	-0.326	-0.365	0.144
CurCwnd	-0.013	-0.811	0.526	-0.189
MaxSsCwnd	-0.557	0.353	-0.326	-0.574
DupAcksIn	-0.716	-0.011	-0.273	-0.553
DupAcksOut	-0.564	0.511	-0.472	-0.363
SlowStart	0.446	-0.644	0.516	0.199
CurSsThresh	-0.555	0.515	-0.482	-0.360
MaxSsthresh	0.117	0.877	-0.019	0.218
MinSsthresh	-0.195	-0.771	-0.429	-0.339
SacksRcvd	0.371	-0.102	-0.146	0.340
SackBlocksRcvd	-0.118	-0.192	0.074	0.610

Here, we consider the obtained principal component scores and the user satisfaction as predictor variables and a criterion one, respectively; we then perform the multiple regression analysis. Equation (2) shows the obtained regression line.

$$\hat{U} = -0.126Z_1 + 0.129Z_2 - 0.259Z_3 - 0.151Z_4 \qquad (2)$$

In Eq. (2), $\hat{U}$ is an estimate of the user satisfaction and $Z_i(i=1,2,3,4)$ means the i-th principal component score. The multiple correlation coefficient adjusted for degree of freedom of Eq. (2) is 0.728. In Eq. (2), every partial regression coefficient is statistically significant but the intercept is not.

From the result of the multiple regression analysis, we see that the user-satisfaction can be nearly predicted by the first four principal component scores. However, to obtain the principal component scores requires all the measured values for each 31 variable. From a view of service level management (SLM), we should decrease the number of TCP variables for measurement as many as we can. Then, instead of the principal component scores, we try to use some variables among the above 31 ones as predictor variables as follows.

First, we classify the above 31 variables into four groups according to Table IV. Then, we pick up one variable among each group and perform the multiple regression analysis for all combinations of predictor variables. Among all the combinations of predictor variables, we consider some combinations that give high multiple correlation coefficient adjusted for degree of freedom.

From the result of the multiple regression analysis, we got two combinations. Equations (3) and (4) show the obtained equations. The multiple correlation coefficients adjusted for degree of freedom of Eq. (3) and that of Eq. (4) become 0.739 and 0.731, respectively.

$$\begin{aligned}\hat{U} = &-2.091 + 7.690\times10^{-4}\,CurCwnd \\ &+ 6.069\times10^{-2}\,SlowStart - 2.758\times10^{-4}\,SumRTT \\ &+ 1.367\times10^{-4}\,RTTVar\end{aligned} \qquad (3)$$

$$\begin{aligned}\hat{U} = &-1.477 + 6.853\times10^{-4}\,CurCwnd \\ &+ 4.949\times10^{-2}\,SlowStart \\ &- 1.193\times10^{-4}\,CongAvoid\end{aligned} \qquad (4)$$

Moreover, the result of the statistical tests for Eq. (3) and those for Eq. (4) are shown in Tables V and VI, respectively.

TABLE V. RESULT OF MULTIPLE REGRESSION ANALYSIS
(THE MULTIPLE CORRELATION COEFFICIENT ADJUSTED FOR DEGREE OF FREEDOM IS 0.739.)

Variable	Partial regression coeff.	Standard error	*t* value	Probability
CurCwnd	7.690e-04	1.208e-04	6.362	< 0.001
SlowStart	6.069e-02	1.309e-02	4.633	< 0.001
SumRTT	-2.758e-04	7.687e-05	3.588	< 0.001
RTTVar	1.367e-04	6.291e-05	2.173	3.246e-02
Intercept	-2.091	7.273e-01	2.875	5.060e-03

TABLE VI. RESULT OF MULTIPLE REGRESSION ANALYSIS
(THE MULTIPLE CORRELATION COEFFICIENT ADJUSTED FOR DEGREE OF FREEDOM IS 0.731.)

Variable	Partial regression coeff.	Standard error	*t* value	Probability
CurCwnd	6.853e-04	1.243e-04	5.509	< 0.001
SlowStart	4.942e-02	1.232e-02	4.009	< 0.001
CongAvoid	-1.193e-04	3.515e-02	3.395	1.030e-03
Intercept	-1.477	7.401e-01	1.996	4.894e-02

In Tables V and VI, we deleted predictor variables whose partial regression coefficient is statistically insignificant. Therefore, in order to predict the user satisfaction, three or four scales, such as ones shown in Tables V and VI, are required. As a result, we see that TCP variables concerning with congestion control are suitable for estimation of the user-satisfaction. Note that these results do not indicate that the other scales have no contribution to the user satisfaction. The result of PCA mentions that the variables shown in Eqs. (3) and (4) correlate with some variables in the 31 scales.

Although the conclusion of this experimental study cannot be applied to all the existing Web services, as a final point, we should note that it can be possible to apply this approach to estimation of some QoE parameters in other Web services.

VI. CONCLUSIONS

This paper studied an estimation of QoE from TCP variables in online shopping Web services by experiment. We measured 31 variables related to TCP performance by using Web100 software and assessed a user-satisfaction with the rating-scale method. Then, we calculated regression lines that relate TCP variables to the user-satisfaction by means of the general multivariate statistics: the principal component analysis and the multiple regression analysis. As a result, in this experiment, we confirmed that the user-satisfaction in the online shopping Web services can be estimated the three or four objective TCP variables.

We have many issues to be discussed as our future work. First, we considered an online shopping Web service but we can tackle other Web services. Since the effect of TCP performance on a QoE depends on contents of a service and there are innumerable Web services exist, we also try to categorize Web services at first. Second, we would like to utilize our experimental results to control service level of the Web service.

ACKNOWLEDGMENT

The authors thank the subjects who cooperated with our experiment.

REFERENCES

[1] Web-Based Training Information Center. Available: http://www.wbtic.com/

[2] Japanese Government Internet TV. Available: http://nettv.gov-online.go.jp/eng/index.html

[3] J. Nielsen, Designing Web Usability. Berkley CA: New Riders Publishing, 2000.

[4] J. Nielsen and H. Loranger, Prioritizing Web Usability. Berkley CA: New Riders Publishing, 2006.

[5] ITU-T: ITU-T G.1010, End-user multimedia QoS categories, ITU (2001).

[6] ITU-T: ITU-T G.1030, Estimating end-to-end performance in IP networks for data applications, ITU (2005).

[7] ITU-T: ITU-T G.1031, QoE factors in web-browsing, ITU (2014).

[8] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1," IETF RFC 2616, Jun. 1999.

[9] J. Postel, "Transmission Control Protocol," IETF RFC 793, Sept. 1981.

[10] M. Allman, V. Paxson, and E. Blanton, "TCP congestion control," IETF RFC 5681, Sept. 2009.

[11] L. Brakmo and L. Peterson, "TCP Vegas: End to end congestion avoidance on a global Internet." IEEE Journal on Selected Areas in Communication, vol. 13, no. 8, pp. 1465–1480, Oct. 1995.

[12] M. Gerla, M. Y. Sanadidi, R. Wang, A. Zanella, C. Casetti, and S. Mascolo, "TCP Westwood: congestion window control using bandwidth estimation", in Conf. Rec. IEEE Globecom 2001, vol. 3, pp 1698-1702, San Antonio, Nov. 2001.

[13] L. G. Grimm and P. R. Yarnold, Reading and understanding multivariate statistics. Washington DC: American Psychological Association, 1995.

[14] J. P. Guilford, Psychometric methods. New York NY: McGraw-Hill, 1954.

[15] ITU-T Rec. P.800, "Method for subjective determination of transmission quality," Aug. 1996.

[16] ITU-R Rec. BT.500-12, "Methodology for the subjective assessment of the quality of television pictures," Sept. 2009.

[17] M. Mathis, J. Heffner, and R. Raghunarayan, "TCP extended statistics MIB," IETF RFC 4898, May 2007.

[18] The Web100 project. Available: http://www.web100.org/

[19] Firefox Web browser. Available: http://www.mozilla.org/firefox/

[20] L. Rizzo. Dummynet. Available: http://info.iet.unipi.it/luigi/dummynet/

[21] V. Paxson and M. Allman, "Computing TCP's retransmission timer," IETF RFC 2988, Nov. 2000.

[22] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, "TCP selective acknowledgment options", IETF RFC 2018, Oct. 1996.