Journal of Universal Computer Science, vol. 18, no. 11 (2012), 1523-1543
submitted: 30/11/11, accepted: 15/5/12, appeared: 1/6/12 © J.UCS

Utilization-level and Serviceability of a Social Name-card Portal for QoS in a Cloud Social Networking Service

Yung Bok Kim
(Sejong University, Seoul, Korea
yungbkim@sejong.ac.kr)

Abstract: Various mobile-web services have expanded with the evolution of mobile Internet technologies and with the increasing variety of smart phones in the proliferating cloud computing environment. Web services for the integration of cloud applications/services can be evaluated at the utilization-level as well as the serviceability of the service on the server side in the management for cloud computing. In web activity, a web server can be a unified hub for web interactions as well as for the real-time estimation model of service-based parameters, i.e. utilization-level/serviceability for service monitoring. With the real-time estimation/analysis of the parameters in a web server for service-based contents delivery, the utilization-level and serviceability of a social-web name-card portal are presented. Empirical results are presented on the basis of the implementation of the real-time estimation scheme in a social-web name-card portal server for a cloud SNS.

Keywords: Social Web, Name card, QoS, Utilization level, Serviceability, Cloud, SNS
Categories: H.1, H.3, H.4, H.5

1 Introduction

A social-web information service with a unified name-based directory for efficient interaction as well as for the integrity of consistent name-based information was researched for a service-based name-card contents delivery. Social-web name-card contents (i.e. contents as a service) integrated with Facebook's 'Like' and Twitter's 'Tweet' API for mobile applications (for smart phone/TV/PC) were studied with a web information server in the cloud computing environment. As an example of a specific social-web application, a mobile-web information service was implemented with a social-web name-card directory as service-based contents for social networking. Social networks are "explicit representations of the relationships between individuals and groups in a community" [Lytras 07]. On the user's side, virtual identities, social relationships, and communities are increasingly established using the Internet [Trier 09]. Social networking is based on human socialization in a virtual environment, where one can chat with friends, show photo albums, see movies, listen to music and be a member of clubs and associations in daily life: socializing, talking, joining groups of interest and sharing with friends and family [Ferreira 10]. Ferreira studied the main advantages and implications of Web 2.0 by analyzing a social networking service (SNS).

A variety of social networking services (e.g. *Facebook*, *Twitter*, *Google+*) using smart phones such as the iPhone and Android phone have been proliferating rapidly around the world. A great deal of application software is downloadable from the iPhone App Store, and mobile-web services are also expected to be a major trend of the mobile service based on the emerging HTML5. Researched here is a social-web name-card portal service based on web standards including HTML5 for mobile-web social networking in the cloud computing environment.

The performance of a worldwide social-web service as a service platform is an important issue in providing a ubiquitous, reliable, and efficient social-web information services in the cloud computing environment. Client mobile devices for social-web information access have become very important for social networking interaction as well as for service-based contents delivery. The social-web name-card server, accessible ubiquitously with any Internet-capable device anytime and anywhere, plays a central role in social-web interaction for a social-web information service and real-time social-web activity analysis for cloud computing services. Our previous research [Kim 10a, 10b] focused on how the abstraction of accessibility and usability evolved more specifically and expanded to real implementation including the new concept of *serviceability* for service monitoring in the cloud computing service. We defined the *utilization-level* as a *level of current utilization* of a web server/service and tried to differentiate the time-variant utilization-level from the conventional usability. For the performance of social-web information access in a social-web name-card portal, several perspectives regarding the social-web server for social-web interaction were researched. Instead of a qualitative approach, estimation of quantitative parameters of the *utilization-level* and *serviceability* were researched for the real-time estimation/analysis in service-based name-card contents delivery.

We need to know the current status of utilization of a web server/service for service monitoring, and that is for resource management such as planning or resource allocation in service-based contents delivery. We defined a level of utilization of a web service/server based on measurable parameters representing the utilization such as web sessions (number, frequency, and duration) in a web server. Real-time estimation models of the utilization-level and serviceability in social-web activity for the service-based contents delivery with a social-web name-card portal server are proposed. Presenting the *utilization-level* and *serviceability*, this research implemented a social-web name-card portal site http://ktrip.net to find frequently used social contents such as social-web name-cards showing SNS URLs (e.g. Facebook ID, Twitter ID), lecture bulletin-boards, web-based phone numbers (of friends, relatives, or fellow alumni), and blog URLs, as application examples of a social-web name-card directory. The web service with service-based name-card contents (i.e. contents as a service) was integrated using the ‘Like’ or ‘Tweet’ button API with Facebook/Twitter SNS.

In the following sections, the background and related works are introduced. The utilization-level and serviceability of a web server/service for service monitoring are defined and presented. The real-time estimation model of utilization-level/ serviceability with social-web interactions for service-based contents delivery are presented. The implementation of real-time estimation/analysis with ‘global.asa’ in the http://ktrip.net server is presented. The empirical results (http://ktrip.net/display.asp) based on real-time estimation models with the implementation of a social-web name-card (as service-based contents) portal service http://ktrip.net are presented, then the utilization-level and serviceability with empirical data analysis are presented. Finally, this study concludes with a consideration of future work.

2 Background and Related Works

2.1 Utilization-level and Serviceability

Software technologies challenge not only the services but managerial perspectives and their management [Lytras 11], especially in this knowledge and information society. We have studied the *utilization-level* and *serviceability* as important factors for management of the cloud computing environment (e.g. SaaS services). The resource-allocation/billing schemes of cloud computing service providers are premature and usually based on a simple number of employees/client-computers rather than the actual utilization of the cloud computing service.

The conventional usability for various applications has been researched by many researchers. They have attempted to improve user interaction through the notion of context awareness by exploiting information relating to users, devices and environments. New approaches and related instruments are needed for meeting human requirements. In a typical interactive information retrieval evaluation, typical outcome measures were usability and performance with offline analysis. Usability measures were based on searchers' responses to questionnaire items or their interactions with the system, performance measures were based on the number of relevant document searchers found and the time it took them to do so [Kelly 09].

Some researchers [Seffah 06] studied usability in ISO standards, i.e. ISO 9241-11 (1998) and 9126-4 (2001), proposing that "Usability (focused on software) is generally a relative measure of whether a software product enables a particular set of users to achieve specified goals in a specified context of use. Usability can vary from one user to another and in general with the context of use, e.g. user profiles, task characteristics, hardware, software, and physical or organizational environments."

Weight-modeling of business to customer (B2C) system quality [Stefani 11] was studied with the external quality characteristics and sub-characteristics of ISO9126 software quality standard, which defines four quality sub-characteristics for *usability:* understandability, learnability, operability, and attractiveness. The usability is mainly for software quality, and it may be assessed and estimated qualitatively on the basis of offline analysis. The estimation scheme is not adequate for real-time analysis of a web information system (WIS) and not appropriate for service monitoring as well as for resource-allocation/billing in the cloud computing environment.

Sutcliffe and Angeli [Sutcliffe 05] assessed interaction styles, i.e. traditional menu-based and interactive metaphors, in the web user interface, and they found that the menu-based interaction style was superior for usability and information quality. The usability of mobile devices and their applications is a key factor for the success of mobile computing. An empirical study [Castro 11] demonstrated that users preferred a WIS with routes and a main menu including the beginning of each route, and perceived it to be more easily navigable. The use of routes in the navigation model improved the navigability of the WIS, particularly for subjects with less experience on the basis of offline data analysis with a SPSS statistics tool. For good navigability, we researched and implemented a social-web name-card directory service with numerous single-character alphabet/character TLDs (top level domain) and simple web pages (depth of two) for service-based contents delivery with a social-web name-card server.

Software usability is generally regarded as ensuring that software products are easy and effective to use from the user's perspective; diverse aspects that ensure the usability of a software product should be assessed during the different phases of its life cycle. The evaluation, with both a quantitative and qualitative analysis, was researched [Garcia 11] on the basis of a user-centric approach and the use of questionnaires. Today's ever-evolving online social networks (OSNs) (such as Facebook and Twitter), effective and usable access control framework, and the OSN access control can be built around the concept of user activity. Much of the recent literature on OSN access controls does not distinguish a session from a user [Park 11].

The aforementioned usability as a rather static characteristic has been researched by many researchers; however, we need a time-variant parameter representing current utilization of a web server/service for managerial-perspectives/management in cloud computing services. We need a more appropriate parameter representing the *current level of utilization* of the WIS for service monitoring. A social-web name-card directory service with *utilization-level* for social-web interaction activity in the ubiquitous Internet environment was implemented, especially for a cloud computing service. In our research, we found that we need different parameter/terminology from the conventional usability, because the conventional usability does not show properly the current level of time-variant utilization-level of WIS or SaaS service. Proposed here is the *utilization-level* and *serviceability* for the real-time quantitative estimation/analysis with a synthetic approach, rather than analysis with various qualitative factors, considering social-web interaction activity. Based on activity theory, web activity in the real-time ubiquitous computing and networking environment was studied for estimation of accessibility and usability [Kim 10a].

We defined the *utilization-level* as the level of utilization of a web service/server with measurable parameters such as session characteristics of the web service/server (e.g. number of sessions, session arrival rate and session duration). We also defined the *serviceability* as the service-capability of a web server/service with measurable parameters such as inter-arrival time/rate of sessions that can be serviced by the web server/service. We can estimate each utilization-level and compare it between different web services. Mobile social networking with Twitter and Facebook using smart phones (e.g. iPhone, Android phone, etc.) has been proliferating rapidly the world over. A common time metric was used to compare the time-variant *utilization-level/serviceability* between different/heterogeneous web services (e.g. SaaS service) including proliferating mobile-web services in the cloud compting environment.

In many current national SaaS projects and cloud computing services for small/medium-sized companies in Korea, the service monitoring, resource-allocation, and billing schemes of SaaS/(PaaS/IaaS) service providers are premature and based on simple number of employees/client-computers/(CPU/bandwidth) rather than actual utilization of the cloud service. The proposed scheme of *utilization-level* may be useful. *Utilization-level* and *serviceability* are not constant and they are stochastic random variables. *Utilization-level* has not been studied a great deal with a quantitative estimation scheme, as researchers tend instead toward some qualitative evaluation with the concept of usability as presented. The *utilization-level* and the *serviceability* of a social-web name-card portal for service-based contents delivery in a cloud SNS was researched quantitatively.

2.2 QoS/QoE in Social Networking Service

Many techniques and tools for validating web applications have been created, but none have attempted to leverage data gathered in the operation of web applications to assist with testing. The analysis of user-session techniques was suggested using captured user sessions [Elbaum 05]. Some researchers [Yang 09] studied the relationship between user behavior and user preference during web browsing on small screen devices to find users' interest blocks with offline statistical analysis.

In the next-generation network environment, networks are designed to be multiservice, supporting a wide range of premium services; new QoS policies are required that adapt the traditional QoS regulatory model to this emerging context. Some researchers [Ibarrola 11] presented a new approach to identify the key factors contributing to the development of a new QoS regulation for future networks, with a case study with the Spanish Internet service provider; it emphasized the user aspect and its crucial role in the ITU-T QoS framework. In services and cloud computing systems, managing a variety of QoS requirements is extremely difficult because numerous application developers are dynamically composing services over networks, and various providers with different techniques and policies are managing the services.

The QoS features of all services are tightly interrelated, and there are trade-offs among them, features like throughput and service delay rely on system resource allocation at the applications' runtime. Often, the same server hosts multiple services, which compete for the server's CPU time, memory, and network bandwidth [Yau 11].

For real-time analysis of web information services for SNS, we researched several parameters (e.g. utilization-level and serviceability) about social-web servers for quality of service (QoS) and quality of experience (QoE). QoE is a measure of user performance based on objectivity. A structured approach to defining and measuring QoE in relation to QoS was studied as a basis for deriving QoE guidelines for network operators, equipment manufacturers, and service providers [Brooks 10]. QoE ties together user perception, experience, and expectations to application and network performance, typically expressed by QoS parameters.

The utilization-level as well as the experienced serviceability (for service monitoring) as quantitatively measurable QoS/QoE parameters is proposed, respectively. In SNS such as Twitter (or Facebook) and many other local/global social networking services, the performance and QoE related to QoS are becoming increasingly important especially for real-time information networking with real-time registration and searching of service-based contents for social networking.

By personalizing service discovery, users can find the most appropriate services for their immediate situation; some researchers [Park 09] studied to help mobile phone users find appropriate services according to their preferences and contexts. The user interface proposed in this paper with a social-web name-card directory for service-based contents delivery must be convenient even for typing in the domain names with smart phones. Because the first step for web service with wired/mobile Internet (especially, with mobile Internet) is typing in the domain name of the targeted web site offering the requested information or service to get a main menu and right route for good navigability [Castro 11].

The scheme for multilingual domain names has been standardized world-wide by the Internet Engineering Task Force (IETF) and has been approved by the Internet Corporation for Assigned Names and Numbers (ICANN). The auto-conversion functionality (i.e. from multilingual domain name to Puny-code, or vice versa) for standardized multilingual domain name service has been embedded in web browsers as a built-in functionality, e.g. from the recent version of MS IE, Firefox, Safari, Opera, and Google Chrome. Abstractions about parameters, focused on the accessibility/usability with single-character multilingual domain names as single-character keys [Kim 10b], were researched to be transformed to functional solutions for implementation and operating service. Our model was refined more specifically for utilization-level and expanded to serviceability with simple implementation, especially for a cloud SNS. The *utilization-level* and *serviceability* of a social-web interaction server with smart-phone/TV/PC was researched, especially using a social-web name-card directory accessible with many simple (single-character) multilingual domain names, i.e. TLDs, related to search keyword(s), considering the navigability, the resource allocation and the billing scheme in the cloud computing environment.

3 Estimation Model of Utilization-level and Serviceability

3.1 Social-web Interaction in Social-web Activity

We considered the aforementioned usability, utilization-level and serviceability at managerial perspectives, and we focused on the real-time quantitative rather than qualitative metric for QoS in the cloud computing environment. The parameters, i.e. utilization-level and serviceability, are random variables, and should be estimated in real-time for real-time evaluations of QoS (quality of service) and QoE (quality of experience). Depending on the phase of the software life cycle in which they are applied, usability metrics can be classified into one of two major categories, testing and predictive [Seffah 06]; data from testing metrics are collected in order to measure the actual use of working service. We focused on the utilization-level (instead of a conventional usability) and serviceability of a web server/service for service monitoring.

Services that rely on mobile devices have proven popular and important in both developed and developing countries: one-third of Facebook users are mobile only, and half of Twitter use is mobile. Most of the mobile delay is caused by the connection's round-trip delay time [Skehin 11]. As the field of mobile sensing matures, researchers are increasingly encountering the limits of conventional approaches to activity recognition. The cooperative communities framework is contributing to a promising new alternative direction for activity recognition—one that studies how not just individual users, but the communities in which they live. This can be leveraged to better model human behavior. Lane et al. [Lane 11] presented how social networks for large-scale human behavior modeling can be exploited.

Time-based utilization-level is very relative and stochastically changing and is affected by other services/products because of the 24-hour time constraint in human activity. We attempted to find the utilization-level as a common parameter for real-time estimation and real-time comparison between different services in the cloud computing environment. We present a real-time estimation and analysis scheme for the utilization-level and serviceability of a social-web name-card (i.e. service-based contents) interaction server http://ktrip.net. A primitive concept was researched at the early stages of our research, focusing on the accessibility/usability with single-character multilingual domain names as single-character keys in a name-card directory service [Kim 10b], and further refinement and implementation have been researched for a cloud SNS, considering managerial perspectives and service-based contents delivery.

The session information is stored on the social-web server using the session identifier (session ID) generated as a result of the first request from the end user running a web browser. Session management is the process of keeping track of a user's activity across sessions of social-web interaction with the computer system. Real-time estimation/analysis based on a session (i.e. HTTP session) in a social-web server was the focus in our research for a cloud computing service. The important parameters of utilization-level and serviceability with social-web interactions from the service perspective are proposed here. They are random variables, and it is necessary to estimate them in a real-time way for the real-time analysis for a resource-allocation/billing scheme in the cloud computing environment.

There are several interactions in a session as shown in *Figure 1* below, which shows the social-web interaction sequence including accessing ktrip.net (I_1) with the iPhone Safari browser, reading a title/name-card list, reading the contents (I_2) after clicking a listed title in http://ktrip.net (mobile-web), then finishing the session (I_3) after closing the browser. On the http://ktrip.net server/site, we can estimate in real-time the interaction time of the social-web activity with variation depending upon network/server condition. The *utilization-level* and *serviceability* in social-web interaction service are presented. The *utilization-level* and *serviceability* are estimated from the usage frequency for simple implementation, with easily measurable parameters in a web server in the cloud computing environment. If a service is utilized more, then more sessions will be made; the number of created sessions per second, which we define as a usage frequency, is correlated to the utilization-level. The *utilization-level* based on the *session duration time* and the *usage frequency* will be discussed below.

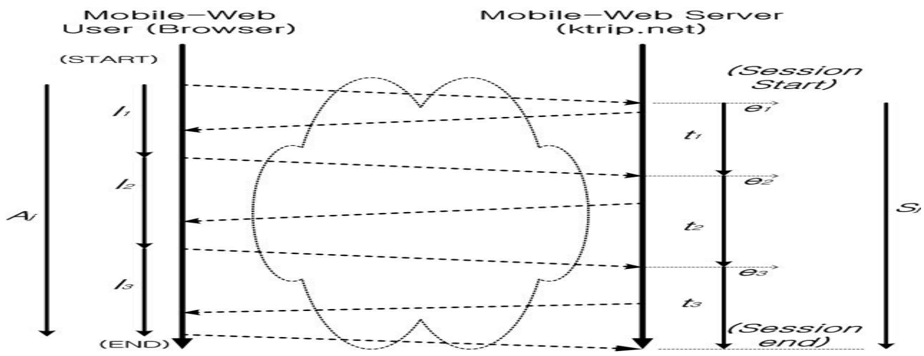

Figure 1: Social-web Interactions in a Session in Social-web Activity

3.2 Real-time Estimation Model of Utilization-level

Session duration time and utilization-level are not constant and they are stochastically changing random variables. The other services or activities are affecting the utilization-level, and new services or other activities may affect the session duration time including the utilization-level of a specific user or user groups (e.g. various communities). For example, the activity of a user with the Facebook service affects the activity with the Twitter service and others. This activity affects the utlizaton-level of a web server/service in the cloud computing service.

For real-time estimation, the session duration time can be estimated consistently on the basis of easy implementation in the programs (such as the ASP, JSP, or PHP) running on the social-web server, instead of estimation on the user's side. For the implementation, instead of estimation of each individual event ($\boldsymbol{e_1}$, $\boldsymbol{e_2}$, …) or each time interval ($\boldsymbol{t_1}$, $\boldsymbol{t_2}$, …) in *Figure 1*, whole session duration time is easy to estimate with the global.asa program. The global.asa program (e.g. with the ASP program in MS IIS server) is the recommended place to estimate in real-time with a simple implementation, and this is presented in the following sections.

$$\textit{Session Duration Time} = D_j \approx \sum_{i=1}^{n} t_i \approx S_j \ \textit{[sec]} \tag{1}$$

The right place for the real-time estimation/analysis is proposed considering the requirements and implementation based on the server program in the cloud computing service. For example, the *global.asa* ASP web server program is the right place for a simple implementation of the real-time estimation/analysis of session time and usage frequency (i.e. number of sessions within a time period).

$$\textbf{\textit{Utilization-level}} = \sum_{j=1}^{f} D_j \approx \overline{D} * \overline{f} \tag{2}$$

For the sake of simplicity, the *utilization-level* as a stochastic random variable could be asymptotically approximated with the multiplication of the *mean* (i.e. average) of *session duration time* ***D*** *[sec]* and the *mean* of usage frequency ***f*** [1/sec]. In equation (2) the magnitude without dimension is dependent on the dimension of frequency ***f***, which we used with 'per second' (i.e. [1/sec]) in our implementation, as discussed in the following section. The greater the value of *utilization-level* is, the more utilized the social-web service for a user/group in a cloud SNS is.

The utilization-level (Equation (2)) could be used for both offline and real-time analysis. We present the real-time estimation of statistics (e.g. utilization-level, usage frequency). For the real-time estimation, we used an exponentially weighted moving average model with the smoothing parameter (we used 0.1 in our simplified implementation of real-time estimation, i.e. giving 10% weight to the most recent sample and 90% weight to the recent average with our social-web experience) to get the mean value of the random variable (i.e. utilization-level). The *mean (i.e. average)* value of *Session Duration Time*: $\overline{D_k}$, required for the real-time estimation can be estimated with a smoothing model as follows:

$$\overline{D_k} = 0.1 * D_k + 0.9 * \overline{D_{k-1}} \tag{3}$$

The utilization-level can be estimated approximately as follows:

$$Utilization\text{-}level = U_k \approx \overline{D_k} * \overline{f_k} \tag{4}$$

In another way, the *mean (i.e. average)* value of *Utilization-level* $\overline{U_k}$, for the real-time estimation can be estimated or forecasted (if needed for resource management in the cloud computing service) as follows:

$$\overline{U_k} = 0.1 * U_k + 0.9 * \overline{U_{k-1}} \tag{5}$$

$$\overline{U_k} = 0.1 * (\overline{D_k} * \overline{f_k}) + 0.9 * \overline{U_{k-1}} \tag{6}$$

The right place for the real-time estimation/analysis is proposed considering the requirements and implementation based on the server program. For example, the *global.asa* ASP web server program is the right place for the simple implementation of the real-time estimation/analysis of session duration time and usage (session) frequency (i.e. number of arrival sessions within a time period), especially for a cloud computing service.

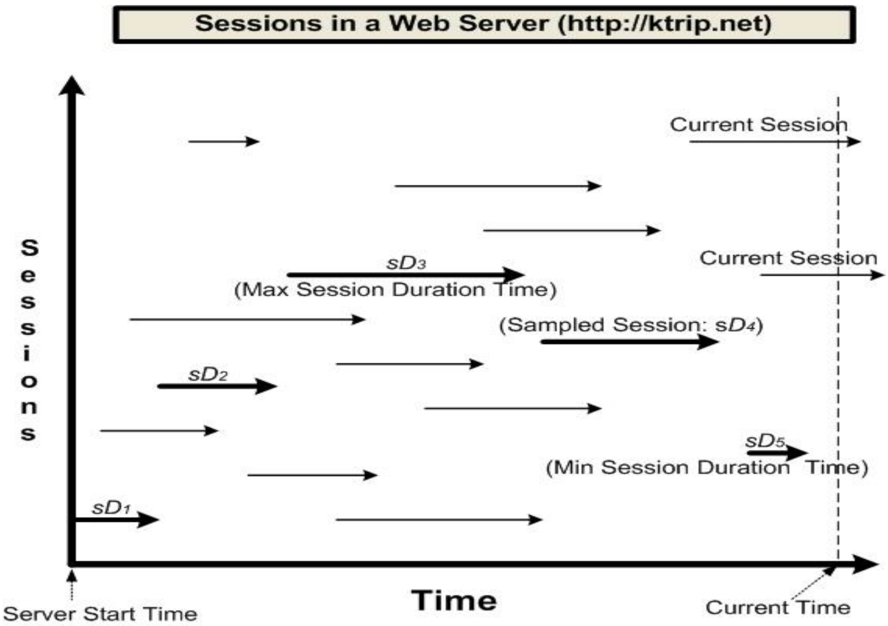

Figure 2: Arrival Stream of Sessions in a Web Server (ktrip.net)

The *utilization-level*, which should be estimated in a real-time way as a quantitative stochastic parameter, was defined and approximated with the equation (7). The *utilization-level* as a stochastic random variable is asymptotically approximated with the multiplication of the *mean* (i.e. average) of *sampled session duration time **sD** [Sec]* and the *mean* of usage (session) frequency ***f*** *[1/Sec]*. As an efficient implementation scheme for various (in terms of load) web servers in the cloud computing service, estimations were in real-time of the *mean* of *sampled session duration time **sD*** with only sampled sessions as shown in *Figure 2,* rather than estimation of session duration time ***D*** with all sessions

$$\textit{Utilization-level} = \sum_{j=1}^{f} D_j \approx \overline{sD} * \overline{f} \tag{7}$$

The greater the value the *utilization-level* is, the more utilized the social-web service (activity) for a user or group in a cloud SNS on the basis of time metric. With this concept, comparison is possible of the *utilization-level* between different types of service. For example, the utilization-level of *Facebook* for a specific user as well as worldwide users can be compared to the utilization-level of *Twitter* (or *Google+*) for a specific user and for worldwide users, on the basis of the real-time/offline analysis depending upon the implementation.

The *mean (i.e. average)* value of *sampled session duration time:* $\overline{sD_k}$ required for the real-time estimation can be estimated with a smoothing model in a web server for a cloud SNS as follows:

$$\overline{sD_k} = \delta * sD_k + (1-\delta) * \overline{sD_{k-1}} \text{ where } 0 < \delta < 1 \tag{8}$$

$$\overline{sD_k} = 0.1 * sD_k + 0.9 * \overline{sD_{k-1}} \text{ when } \delta = 0.1 \tag{9}$$

The utilization-level can be approximately estimated as follows:

$$\textit{Utilization-level } U_k = \overline{sD_k} * \overline{f_k} \tag{10}$$

The *mean (i.e. average)* value of usage (session) frequency: $\overline{f_k}$, for the real-time estimation can be estimated in a web server as follows:

$$\text{Slow tracking model: } \overline{f_k} = \text{(total number of sessions/server life time)} \quad [1/\text{Sec}] \tag{11}$$

$$\text{Fast tracking model: } \overline{f_k} = \frac{1}{mean(\text{inter-arrival time})} = \frac{1}{\overline{I_k}} \quad [1/\text{Sec}] \tag{12}$$

With the real-time estimation model, the *mean (i.e. average)* value of the *Utilization-level*: $\overline{U_k}$ is estimated in real-time in a web server as follows:

$$\overline{U_k} = \delta U_k + (1-\delta)\overline{U_{k-1}} \text{ where } 0 < \delta < 1 \tag{13}$$

$$\overline{U_k} = 0.1 * U_k + 0.9 * \overline{U_{k-1}} \text{ when } \delta = 0.1 \tag{14}$$

$$\overline{U_k} = 0.1 * \overline{sD_k} * \overline{f_k} + 0.9 * \overline{U_{k-1}} \text{ when } \delta = 0.1 \tag{15}$$

Equation (15) can be completed with equation (11) or (12), depending upon the slow/fast tracking estimation model of the usage (session) frequency $\overline{f_k}$, here both were implemented.

3.3 Real-time Estimation Model of Serviceability

The *serviceability* (as a QoS parameter) of a web server in the cloud computing service is defined as the experienced maximum serviceability in one [sec], which can be estimated with the reciprocal of the minimum *mean* inter-arrival time between continuous sessions as shown in *Figure 3*. For real-time estimation, the inter-arrival time between sessions could be estimated consistently on the basis of easy implementation in the programs (e.g. ASP) running on the web server in the cloud computing service. For the implementation, the inter-arrival time is easy to estimate with the global.asa program. Also, the global.asa program (e.g. with the ASP program in MS IIS server) is proposed here as the right place to estimate the serviceability in real-time with a simple implementation for service monitoring.

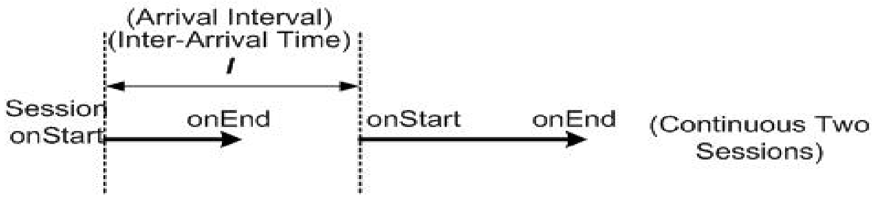

Figure 3: Inter-arrival Time between Continuous Sessions in a Web Server

In *Figure 2* and *3*, the stochastic inter-arrival time can be estimated as follows:

$$\textit{Inter-arrival time } I = onStart_2(t) - onStart_1(t)$$

The *mean* (average) value of the stochastic inter-arrival time ***I*** can be estimated, as presented in the following section. The right place for the real-time estimation/analysis is proposed considering the requirements and implementation based on the server program. The *global.asa* ASP web server program is the right place for a simple/easy implementation of the real-time estimation/analysis of the *serviceability*, and *serviceability* [1/Sec] based on the reciprocal of minimum *mean inter-arrival time* ***I*** between continuous sessions is proposed here. The *maximum Serviceability*: $\boldsymbol{S}_{\max}$ of a web server in the cloud computing service can be estimated with smoothing model: The *Mean* (i.e. Average) value of *inter-arrival* (i.e. *interval*) *time*: ***I***, for the real-time estimation can be estimated as follows:

$$\overline{I_k} = \varphi I_k + (1-\varphi)\overline{I_{k-1}} \text{ where } 0 < \varphi < 1 \tag{16}$$

$$\overline{I_k} = 0.1 * I_k + 0.9 * \overline{I_{k-1}} \text{ when } \varphi = 0.1 \tag{17}$$

The real-time estimation and analysis of the Markovian rate of arrival sessions at a web server for the service-based contents delivery is estimated from the reciprocal (i.e. rate with the dimension [1/Sec]) of the *mean* of inter-arrival time ***I***. The *serviceability* $\boldsymbol{S}_{\max}$ [1/Sec] is estimated from the minimum *mean* of ***I*** [Sec], as follows:

$$S_{\max} = \frac{1}{\min(\overline{I}_k)} \text{ [1/Sec]} \tag{18}$$

4 Implementation and Empirical Results

4.1 Implemented Source Code for Real-time Estimation Models

Figures 2 and 3 show sessions on a time axis, estimating many aspects of sessions in a web server in the cloud computing service. In the global.asa program in on the http://ktrip.net server, *max/min* sesssion duration time, starting/ending time of sessions, *inter-arrival* time between adjacent sessions, etc. could be estimated in real-time for the real-time estimation of *utilization-level* and *serviceability* of a specific web server in cloud computing environments. A greater number of sessions implies more test preparation and execution time, techniques for filtering sessions such as the reduction and clustering techniques [Elbaum 05] were suggested; that filtering concept with sampled sessions was considered in our implemention for estimation.

Continuous sample sessions (**S**) for the real-time estimation of utilization-level are shown in *Figure 2*; this sampling scheme is proposed as a reduction technique of samples, and the empirical results are presented in the following section. The statistics (i.e. *mean, max/min of session duration time*) are estimated in real-time with sampled sessions, and the statistics of the inter-arrival time (to get the usage frequency based on fast tracking) are estimated with all sessions on http://ktrip.net, as shown in *Figure 2* and *3*. Source codes of global.asa in the http://ktrip.net server for a cloud SNS were implemented for the real-time estimation of stochastic random variables, i.e. *session duration time*, *utilization-level*, usage frequency, serviceability, etc. Statistics for *mean*, *max/min* value of each stochastic random variable are estimated in real-time. Equations (9), (10), (12), (15), (17) and (18) in the previous section have been implemented for estimation of parameters. Another estimation parameter ("*mean*InterArrivalTime") is for a usage frequency based on inter-arrival time between continuous sessions in a web server for a cloud SNS.

4.2 Empirical Results from Real-time Estimation Models

Within the program, *global.asa*, for the social-web name-card portal site, http://ktrip.net, numerous interested parameters can be estimated in real-time. Then, with the following program, http://ktrip.net/display.asp, the estimated statistics of interested parameters can be displayed on the screen at anytime and anywhere.

Source Code Listing for Displaying Estimated Parameters

```
<body>
…….. (omit)…………….
Start time of Server <%=Application("serverdate")%>   <br>
Mean Current Number of Sessions=<%=Application("MeanCurrentSes")%> <br>
Mean Session Duration Time=<%=Application("meanD")%> [sec] <br>
Utilization-level(fast tracking)=<%=Application("Ufast")%>  <br>
```

```
Mean Usage Frequency(fast tracking) =
<%=1/Application("meanInterArrivalTime")%> [per sec]  <br>
Mean Usage Frequency(slow tracking)=<%=Application("F")%> [per sec]  <br>
Mean Inter-arrival Time between Sessions=<%=
Application("meanInterArrivalTime")%> [sec] <br>
…….. (omit)…………….
Current Session ID=<%=Application("sid")%> <br>
Maximum Session Duration Time= <%=Application("Dmax")%>  [sec]<br>
Minimum Session Duration Time = <%=Application("Dmin")%>  [sec]<br>
Maximum Utilization-level (fast)= <%=Application("Ufastmax")%> <br>
Minimum Utilization-level (fast)= <%=Application("Ufastmin")%> <br>
Serviceability=<%=Application("Smax")%> [per Second]<br>

…….. (omit)…………….
</body>
```

In *Table 1*, the following results (no. 1~36) for the social-web site (http://ktrip.net), which is accessed frequently (*mean* inter-arrival time: 0.4595~52.77 [sec] as shown at no. 24/25 in *Table 1*), are displayed in real-time in the URL: http://ktrip.net/display.asp. The empirical results of Estimated Value (A) in *Table 1* are displayed around 2010-10-20 PM 2:24, and the global.asa program (same initialization as server start) started around 2010-10-18 PM 2:26:38. Similarly, the empirical results of Estimated Value (B) in *Table 1* are displayed around 2011-10-9 AM 8:24, and the global.asa program (same initialization as server start) started around 2011-07-10 AM 11:27:10.

Some parameters for the real-time estimation of empirical data are explained as follows: A2 is for currently active (i.e. in use) number of sessions, A3 is for the *mean* number of recently active sessions, A6 is for the current *mean* of usage frequency of sessions based on the reciprocal of *mean* inter-arrival time (A11) between continuous sessions (for fast tracking *Utilization-level*), A7 is for the current ***Utilization-level*** (*fast tracking*), A8 is for the current *mean* ***Utilization-level*** (*fast tracking*), A11 is for the current *mean* of inter-arrival time between continuous sessions, A15 is for the current number of actually sampled (for real-time estimation of session duration time) sessions, A16 is for the current session duration time, A17 is for the starting time of actually sampled (recent) session (to estimate *sD* with A18 as in the *Figure 2*), A18 is for the ending time of actually sampled (recent) session (to estimate *sD* with A17 as in the *Figure 2*), A25 is for the minimum *mean* of inter-arrival time between continuous sessions (to estimate *serviceability*), and A36 is for the estimated ***Serviceability*** of the social-web name-card portal server.

Below, Estimated Value (A) in *Table 1* shows the statistical results (88 random samples) with 8,899 sessions (from sequential session number: 1 to 8,899) of the ktrip.net server at 48 hours after the starting time of the server (i.e. ktrip.net). From Estimated Value (A) in *Table 1*, 12 active (not completed yet) sessions (as shown in A2) and 928 sampled sessions (as in A15) among 8,887 completed sessions (as in A14) from 8,899 total sessions (as in A1) can be observed in Estimated Value (A). The real-time estimation/analysis was the prime focus. The maximum inter-arrival time, A26 (244[sec]) can be observed. The maximum session duration time of A34

(655[sec]) can also be observed. Estimated Value (A) (acquired October, 2010) can be compared to Estimated Value (B) (acquired October, 2011). Estimated Value (B) in *Table 1* shows the statistical results with 723,059 sessions of ktrip.net server during three months. The characteristics and stability of the estimation models were observed in the really implemented trial service for a cloud SNS; the models worked well. *Utilization-level* can be estimated with the *number of current active sessions*; similarly, *mean Utilization-level* with the *mean number of active sessions* (refer to A2, A3, A7, A8, B2, B3, B7 and B8 in *Table 1*) after smoothing of data fluctuation.

4.3 Empirical Data Analysis and Discussion

In *Figure 4* with the maximum value of session duration time, between sample no. 32 (sequential session number: 489 in the 8,899 sessions) and 33 (sequential session number: 1278) among 88 samples, there is a rapid increase in session duration time: 626 [sec] compared to the prior maximum session duration time: 180[sec] in ktrip.net server/site. The rapid increase in the session duration time is caused by the rapid increase of the active number of sessions (24/19.8 as max/mean number of sessions in the sample no. 32, 49/42 as max/mean number of sessions in the sample no. 33) between the two sequential session numbers (489 and 1278 among 8,899 sessions). In *Figure 5*, the minimum mean inter-arrival time approached 3.02 [sec] from 5.42 [sec] between the two sample no. 32 and 33; therefore the increased number of sessions affected the session duration time in the web server.

From the previous equation (10), estimations were made for the *utilization-level*, which is a dimensionless and stochastic random variable. The implications of the utilization-level were sought. From equation (15), estimation of the *mean utilization-level* in real-time was possible. The implications of the utilization-level was also calculated. The utilization-level is strongly related to the number of active sessions in the web server/site (e.g. http://ktrip.net in our case) for the cloud computing service. *Figure 6* shows the relationship between the utilization-level and the active number of sessions in the web server/site at an arbitrary moment. The following relationship between the utilization-level for fast tracking (13.55 in A7, *Table 1*) and the number of current active sessions (12 in A2, *Table 1*) in the mobile-web server/site http://ktrip.net at a moment of time. The relationship is similar for fast tracking (15.6 in B7, *Table 1*) and the number of current active sessions (15 in B2, *Table 1*) in the mobile-web server/site in a moment of time.

Utilization-level: $$U_k = \overline{D_k} * \overline{f_k} \approx Number - of - Sessions \qquad (10a)$$

From equation (10a) as our estimation model for the *Utilization-level*, with the *mean session-duration time* and *mean inter-arrival-time*, the current *Utilization-level* can be estimated. More conveniently, the *Utilization-level* can be estimated approximately with the current *number of active sessions* as shown in *Figure 6*. The accuracy of proposed estimation models may be interrogated with the deviation of *utilization-level* from the actual *number of sessions*. The strict analysis of estimation error was left as further research.

no	Estimation Parameter	Estimated Value (A)	Estimated Value (B)	Unit
1	Cumulative Number of Sessions	8,899	723,059	
2	Number of Current Active Sessions	12	15	
3	Mean Number of Active Sessions	11.3	14.6	
4	Start Time of Server	2010-10-18PM2:26:38	2011-07-10AM11:27:10	
5	Mean Session Duration Time	159.45	176.98	sec
6	Mean Usage Frequency (fast tracking)	0.085	0.083	1/sec
7	**Utilization-level** (fast tracking)	13.55	15.6	
8	Mean **Utilization-level** (fast tracking)	11.5	14.74	
9	Mean Usage Frequency (slow tracking)	0.052	0.092	1/sec
10	Mean Utilization-level (slow tracking)	7.94	16.0	
11	Mean Inter-arrival Time (between Sessions)	11.8	12.0	sec
12	Front Session Starting Time	2010-10-20PM2:23:38	2011-10-9AM8:23:27	
13	Back Session Starting Time	2010-10-20PM2:23:42	2011-10-9AM8:23:32	
14	Number of Completed Sessions	8,887	723,044	
15	Number of Sampled Sessions	928	43,342	
16	Session Duration Time	179	179	sec
17	Session (sampled) Starting Time	2010-10-20PM2:18:39	2011-10-9AM8:18:10	
18	Session (sampled) Ending Time	2010-10-20PM2:21:38	2011-10-9AM8:21:09	
19	A Current Session ID	720483364	142274258	
20	Max Number of Active Sessions	56	188	
21	Min Number of Active Sessions	1	1	
22	Max Mean Active Sessions	50.3	50.3	
23	Min Mean Active Sessions	1	1	
24	Max Mean Inter-arrival Time	52.77	50.6	sec
25	Min Mean Inter-arrival Time	1.6	0.4595	sec
26	Max Inter-arrival Time (between Sessions)	244	194	sec
27	Min Inter-arrival Time (between Sessions)	0	0	sec
28	Max Mean Session Duration Time	216.28	788.49	sec
29	Min Mean Session Duration Time	120	70.99	sec
30	Max Utilization-level (fast tracking)	95.3	370.8	
31	Min Utilization-level (fast tracking)	0	0	
32	Max Mean Utilization-level (fast tracking)	8.86	263.2	
33	Min Mean Utilization-level (fast tracking)	6.14	0	
34	Max Session Duration Time	655	831	sec
35	Min Session Duration Time	64	64	sec
36	**Serviceability** (fast tracking QoS)	0.625	2.176	1/sec

Table 1: Parameters from Real-time Estimation Models

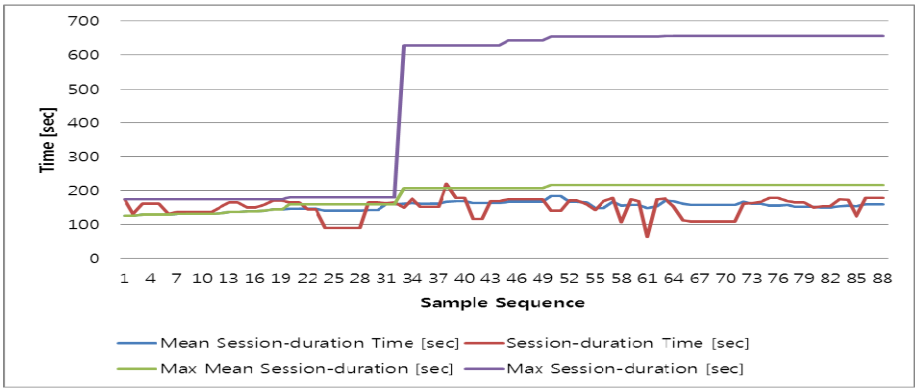

Figure 4: Real-time Estimation of Mean Session-duration Time (in ktrip.net)

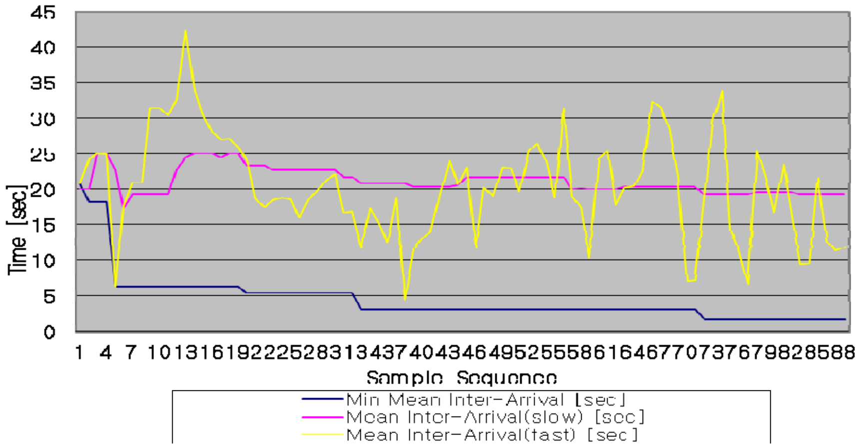

Figure 5: Mean Inter-arrival Time between Sessions (in ktrip.net)

In *Figure 6* with the utilization-level and number of active (in use) sessions, between the sample no. 37 (sequential session number 1299 in the 8,899 sessions) and 38 (sequential session number: 1324) among 88 samples, there is rapid increase of utilization-level (from 8.58 to 38.56). In *Figure 5*, we found that the rapid increase is caused by *mean* inter-arrival time (18.79 [sec] in the session number 1299, 4.35 [sec] in the session number 1324) between the two sequential session numbers (1299 and 1324; i.e. the sample no. 37 and 38). If the arrival rate (i.e. the reciprocal of *mean* inter-arrival time) approached the increased rate, then the session duration time (from 153[sec] to 220[sec]) and the utilization-level increased (from 8.58 to 38.56) with the increased number of sessions (from 14 to 24) with the proposed estimation model in http://ktrip.net server/site.

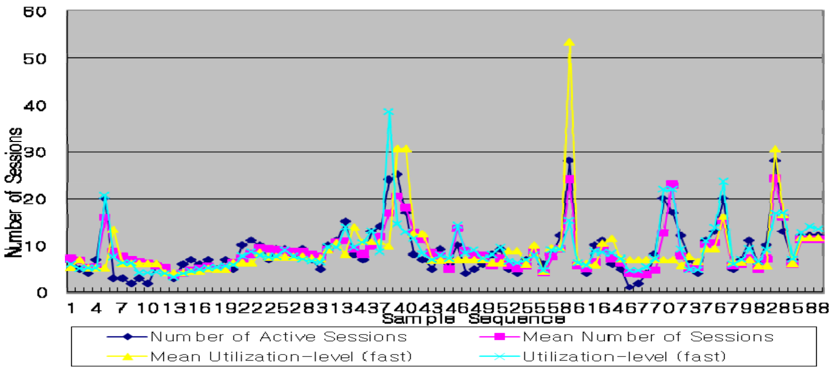

Figure 6: Utilization-level and Number of Active Sessions (in ktrip.net)

In *Figure 6* with *mean* utilization-level and *mean* number of active (in use) sessions, between sample no. 5 (sequential session number: 48 in the 201 sessions) and no. 6 (sequential session number: 58) among 88 samples, there is rapid increase in the utilization-level (from 17.92 to 37) and maximum utilization-level (from 32.3 to 38.05). We found that the rapid increase is caused by the *mean* inter-arrival time (7.43[sec] to 4.7[sec]) between the two sequential session numbers (48 and 58; i.e. the sample no.: 5 and 6). *Figure 6* shows the relationship between the utilization-level and the active number of sessions in a web server for a cloud SNS. The *mean* value of the utilization-level can be compared to the *mean* number of active sessions. The following relationship (*equation (19)*) was found between the *mean* utilization-level and the *mean* number of active sessions in the web server/site http://ktrip.net at a moment of time.

Mean Utilization-level: $\overline{U_k} \approx \overline{\overline{D_k} * \overline{f_k}} \approx \overline{Number - of - Sessions}$ (19)

The formula (i.e. equations (1) through (20)) put the same weight for sessions generated by the same user and sessions generated by different users. From the comparison between A3 and A8 as well as B3 and B8 in *Table 1*, *Figure 6*, and equation (19), the estimation model for *mean Utilization-level* seems to work well. The accuracy of proposed estimation models may be also interrogated with the deviation of *mean utilization-level* from the actual *mean number of sessions*. Further research is required for refinement of proposed estimation model with the trade-off analysis between the accuracy of real-time estimation and the complexity of implementation in a web server for the cloud computing service.

Serviceability: $S_{\max} = \dfrac{1}{\min(\overline{I_k})} = \dfrac{1}{\min(meanInterArrivalTime)}$ [1/Sec] (20)

From *Figure 6* and *equation (18) and (20)*, the *serviceability* is the reciprocal of *minimum of MeanInterArrivalTime*. The *serviceability* of the http://ktrip.net server for a cloud SNS was 0.625 [1/Sec] until it reached 8,899 experienced sessions during 48 hours in Oct. 2010, and later became 2.165 [1/Sec] with 723,059 experienced sessions during 3 months in 2011. As a QoE parameter, the record of estimated serviceability is essential, because the maximum value of serviceability is experienced value during server's operation. The experienced serviceability, as shown in *Figure 7*, had been 4.977 [1/sec] in the year of 2011. The memory size for storing serviceability is negligible, but the most important work is to store and transfer to the global.asa program with non-volatile memory.

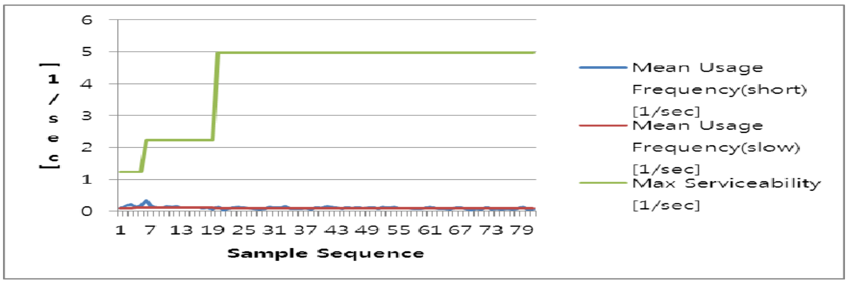

Figure 7: Max Serviceability and Mean Usage Frequency (in 2011)

If the experienced maximum arrival rate (as a QoE parameter), which a social-web server can operate with a number of active sessions and acceptable session duration time for QoS, is known, then we can detect or forecast the situation of the beginning time of over capacity of a social-web server in the cloud computing service. The control scheme will be researched to protect the web server from over capacity as a further research issue, because such over-capacity problems have happened so frequently in rapidly growing social-web services such as Twitter or Facebook. If the two values (i.e. the session duration time and inter-arrival rate) are known, then the number of current active sessions in a social-web server can be estimated. For a cloud computing service, the mean value of the utilization-level can be used for resource-allocation/cost-analysis for billing of each web service/site by a cloud service (e.g. SaaS: Software as a Service) provider. The billing scheme of advertisements, which is related to the utilization-level and serviceability, in a specific social-web server/site for the cloud computing service will be researched further.

As a summary of implementation, the social-web name-card portal server (ktrip.net) for a cloud SNS was developed for ubiquitous social-web name-card services in cloud computing environments. The implementation with server script (e.g. the ASP) was the focus in the real-time estimation and analysis of *utilization-level* and *serviceability* in our social-web name-card directory for a cloud SNS. As an example of ubiquitous social-web information services with various smart phones serviced by many mobile service operators, the ASP server program based on Microsoft IIS web server and DBMS was used.

With this application program and the global.asa for real-time estimation/analysis, the http://ktrip.net server/site for a cloud SNS can be accessed in a unified way by different smart phones for social-web interaction with a quantitative service level (e.g. service level agreement in SaaS) of the *utilization-level* and *serviceability*. The statistical significance of retrieved results is that the *utilization-level* may be estimated from the *number of active sessions*, and the experienced *serviceability* should be estimated through whole service period and should be kept permanently and used for resource management and billing in the cloud computing environment.

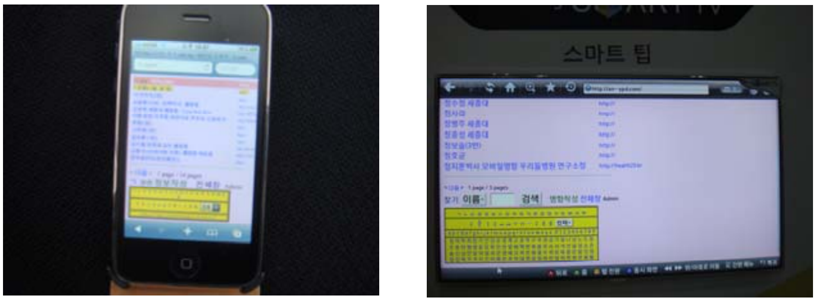

Figure 8: Social-web Name-card Service of ktrip.net with Smart Phone/TV

For Internet domain names for the unified social-web name-card portal service in cloud computing environments, over 300 single-character multilingual domain names (i.e. TLDs) were used, including tens of multilingual (Korean) alphabet domain names as simple domain names to find information as well as to notify information in a ubiquitous and unified way. With the social-web name-card directory site http://ktrip.net and many other single-character TLDs, it was possible to register and search social-web information in real-time with any Internet device such as a smart phone/TV (as shown in *Figure 8*) as well as with PCs using recent versions of browsers (e.g. MS Explorer, Firefox, Safari, Opera, Chrome, etc.). With the proposed scheme for real-time estimation models, it was possible here to monitor/understand in real-time the utilization-level and serviceability of ktrip.net social-web name-card portal server for a cloud SNS.

5 Conclusion

We defined the *utilization-level* and *serviceability* of a web service/server for service monitoring of cloud computing applications. Real-time estimation models for *utilization-level and serviceability* based on social-web interaction sessions were proposed for service monitoring as well as for service-based name-card contents delivery with a social-web name-card portal. *Utilization-level* and *serviceability* are defined based on *mean session duration time* and *mean inter-arrival rate*, which can be easily estimated in real-time. The empirical results, at http://ktrip.net/display.asp, are presented on the basis of the implementation within the global.asa program running in a social-web name-card portal server http://ktrip.net for a cloud SNS, and the estimation models seems to work well. For future work, the practical applications based on the proposed *utilization-level* and *serviceability* in social-web services will be researched for real-time social-web advertisements for various business models based on ubiquitous social-web applications in cloud computing environments (e.g. SaaS service). In the cloud computing environment, the application of the *utilization-level* and *serviceability* to service monitoring/control as well as to the resource-allocation/billing scheme will be researched further for cloud computing services (e.g.

SaaS/PaaS/IaaS and contents as a service). Google voice search and Apple iOS5 SiRi Voice assistant application will be considered in our social-web name-card portal service, and the *utilization-level* and *serviceability* in the cloud computing environment is intended to be improved.

References

[Brooks 10] Brooks, P., Hestnes, B., Telenor: "User Measures of Quality of Experience: Why Being Objective and Quantitative Is Important"; IEEE Network, March/April (2010), 8-13.

[Castro 11] Castro, V., Genero, M., Marcos, E., Piattni, M.: "Empirical study to assess whether the use of routes facilitates the navigability of web information systems"; IET Softw., 5, 6 (2011), 528–542.

[Elbaum 05] Elbaum, A., Rothermel, G., Karre, S., Fisher II, M.: "Leveraging user-session data to support web application testing"; *IEEE Transactions on Software Engineering*, 31, 3 (2005), 187-202.

[Ferreira 10] Ferreira, N.: "Social Networks and Young People: A Case Study"; International Journal of Human Capital and Information Technology Professionals, 1, 4 (2010), 31-54.

[Garcia 11] Garcia, J., Garcia-Penalvo, F.J., Theron, R., Pablos, P.O.: "Usability Evaluation of a Visual Modelling Tool for OWL Ontologies"; J.UCS (Journal of Universal Computer Science), 17, 9 (2011), 1299-1313.

[Ibarrola 11] Ibarrola, E., Xiao, J., Liberal, F., Ferro, A.: "Internet QoS Regulation in Future Networks: A User-Centric Approach"; IEEE Communication Magazine, October (2011), 148-155.

[Kelly 09] Kelly, D., Dumais, S., Pedersen, J.O.: "Evaluation challenges and directions for information-seeking support systems"; *IEEE Computer*, 42, 3 (2009), 60-66.

[Kim 10a] Kim, Y.B.: "Accessibility and Usability of User-centric Web Interaction with a Unified-Ubiquitous Name-based Directory Service"; World Wide Web, 13 (2010), 105–120.

[Kim 10b] Kim, Y.B.: "Real-time Analysis of Time-based Usability and Accessibility for Human Mobile-Web Interactions"; J.UCS (Journal of Universal Computer Science), 16, 15 (2010), 1953-1972.

[Lane 11] Lane, N.D., Xu, Y., Lu, H., Campbell, A.T., Choudhury, T., Eisenman, S.B.: "Exploiting Social Networks for Large-Scale Human Behavior Modeling"; IEEE Pervasive Computing, October-December (2011), 45-53.

[Lytras 07] Lytras, M.D., Rafaeli, S., Downes, S., Naeve, A., Pablos, P.O.: "Editorial"; *Int. J. Knowledge and Learning*, 3, 4/5 (2007), 367-370.

[Lytras 11] Lytras, M.D., Pablos, P.O.: "Software Technologies in Knowledge Society J.UCS Special Issue"; J.UCS (Journal of Universal Computer Science), 17, 9 (2011), 1219-1221.

[Park 09] Park, K.L., Yoon, U.H., Kim, S.D.: "Personalized service discovery in ubiquitous computing environments"; *IEEE Pervasive Computing*, 8, 1 (2009), 58-65.

[Park 11] Park, J., Sandhu, R., Cheng, Y.: "A User-Activity-Centric Framework for Access Control in Online Social Networks"; IEEE Internet Computing, September/October (2011), 62-65.

[Paulson 09] Paulson, L.D.: “W3C adopts web-accessibility specifications”, *IEEE Computer*, 42, 2 (2009), 23-26.

[Seffah 06] Seffah, A., Donyaee, M, Kline, R.B., Padda, H.K.: “Usability measurement and metrics: A consolidated model”; *Software Qual J*, 14 (2006), 159-178.

[Skehin 11] Skehin, N., Chung, J.: “Mobile Computing Systems and Applications”, IEEE Pervasive Computing, July-September (2011), 80-83.

[Stefani 11] Stefani, A., Xenos, M.: “Weight-modeling of B2C system quality”, Computer Standards & Interfaces, 33 (2011), 411–421.

[Sutcliffe 05] Sutcliffe A., Angeli, A.D.: “Assessing interaction styles in web user interface”; in *Proc. of INTERACT 2005*, LNCS 3585 (2005), 405-417.

[Trier 09] Trier, M., Bobrik, A.: “Social search: Exploring and Searching Social Architectures in Digital Networks”, *IEEE Internet Computing*, 13, 2 (2009), 51-59.

[Yang 09] Yang, X., Xiang, P., Shi, Y.: “Finding user’s interest using significant implicit evidence for web browsing on small screen devices”, *Springer, World Wide Web: Internet and Web Information Systems*, 12, 2 (2009), 213-234.

[Yau 11] Yau, S.S., An, H.G.: “Software Engineering Meets Services and Cloud Computing”; IEEE Computer, October (2011), 47-53.