韓國IT서비스學會誌
第13卷 第4號
2014年 12月, pp.189-203

Journal of Information Technology Services
http://dx.doi.org/10.9716/KITS.2014.13.4.189

광고를 동반한 소셜 네트워크 이름-디렉터리 서비스의 실험적 데이터 분석

김 영 복*

Empirical Data Analysis of a Social Network Name-Directory Service with Advertisements

Yung Bok Kim*

■ Abstract ■

With the evolution of Internet technologies and the increasing variety of Internet devices, advertisements in various web services have also expanded. Interactive web services often go hand in hand with effective advertisements for a business model. We estimated statistical parameters of the interactive web server for service monitoring and advertisement-effect. In the web pages, we integrated the plugins of social networking services (SNSs) (e.g. Facebook, Twitter) and an advertisement scheme (e.g. Google AdSense) that regards social name-directory contents.

Empirical data analysis and statistical results are presented with the implementation of estimations of parameters (e.g. utilization-level and serviceability) and advertisements in a social networking name-directory service (http://ktrip.net or http://한국.net). We found that estimated parameters were applicable to service monitoring of web-server as well as to synthesis of advertisement-effect in our social-web name-directory service.

Keyword : Analysis, Parameter Estimation, Serviceability, Name-directory, Advertisement

Submitted : July 22, 2014 1st Revision : November 28, 2014 Accepted : December 1, 2014

* 세종대학교 컴퓨터공학과 교수

1. Introduction

A variety of social networking services (e.g. Facebook, Twitter, Google+) that use interactive mobile devices such as iPhones/iPads and Android phones have been proliferating rapidly around the world. Social networking involves human interactions in a virtual environment, where one can do things such as chat with friends, show photo albums, watch movies, and listen to music (Ferreira, 2010). Researchers can better understand how SNSs work and assess their influence on social interactions as well as the associated benefits and risks (Lampe and Ellison, 2012).

Corporate companies actively use Facebook as a tool for corporate marketing; an implication of value where corporate company use it as a marketing tool was presented, through factors from continued usage of corporate company's Facebook Fan Page users (Shim et al., 2013). Social network service become one of the most successful web-based business, and recommendation in social network sites that assist people to choose various products and services was also widely adopted; Park (2014) reviewed and compared research works about recommendation using social network analysis and collaborative filtering in social network sites.

In this study about SNS, we have researched an interactive social business-card service with a unified name-directory http://ktrip.net using HTML5 (Taivalsaari and Systa, 2012) web-based implementation, in regard to efficient interaction as well as for the integrity of consistent name-based contents that integrate with social plugins and advertisements. Social-web business-card contents integrated with Facebook's 'Like' and Twitter's 'Tweet' application program interface (API) for mobile applications (for smart phone/TV/PC) as well as with Google advertisements (i.e. AdSense) were studied with parameter estimation. For service monitoring, detection of server overload, and managing advertisements, real-time estimation and application of measurable parameters related to utilization status and the service capability of a specific web server were fundamental motivations of this research, especially for what should be estimated, how to estimate, and how to apply.

Usability as a rather static (i.e. time-invariant) characteristic has been investigated by many researchers (Garcia et al., 2011). Our previous research (Kim, 2010) focused on how the abstraction of accessibility and usability evolved. We expanded into real implementation, including the concept of *serviceability* and advertisements in interactive web services. We defined *utilization-level* as the level of current utilization of an interactive web server and tried to differentiate the time-variant *utilization-level* from the conventional time-invariant *usability* (Kim, 2012). A time-variant parameter representing the current utilization of a web server for management in interactive services is necessary, as is a more appropriate parameter representing the current level of utilization of interactive web information services for service monitoring. Estimation of quantitative parameters of the *utilization-level* and *serviceability* in the case of server overload was researched in terms of real-time estimation and analysis of interactive name-directory content delivery with advertisements.

The measurability and manageability of interactive systems is essential, the quality of which depends largely on the resources that the system uses (Tai et al., 2012). The *utilization-level* was defined as the level of utilization of a web server with measurable parameters such as session characteristics of the web server (e.g. number of sessions, session arrival rate and session duration). We defined the *serviceability* as the service-capability of a web server with measurable parameters such as inter-arrival time/rate of sessions that can be serviced by the web server. A common *time* metric was used to compare the time-variants *utilization-level* and *serviceability* between different and heterogeneous web services including expanding interactive mobile-web services. Presenting *utilization-level* and *serviceability*, this research implemented an interactive social name-directory portal site http://ktrip.net (or http://한국.net) to find frequently used social contents, such as social business-cards showing SNS URLs (e.g. Facebook IDs, Twitter IDs), lecture bulletin-boards, and blog URLs, as application examples of interactive social name-directory.

Online social networks such as Facebook and Twitter can be built around the concept of user activity (Park et al., 2011) related to sessions. In our research, the session information is stored on the social name-directory 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 and analysis based on a session (i.e. HTTP session) in a social name-directory server was researched. The important parameters of *utilization-level* and *serviceability* with social name-directory interactions as they relate to sessions from the service perspective are presented.

The standards for locating web services provide interoperability at the syntactic level search, i.e. keyword-based search (Modica et al., 2011). We considered searching schemes appropriate to mobile devices using many single character alphabet domains (Kim, 2006) with a personalized social name-based contents directory (Kim, 2010). We refined the scheme presented in this research more specifically for *utilization-level* and expanded to *serviceability* with simple implementation for social networking contents delivery service. We researched the *utilization-level* and *serviceability* of a social-web interaction server with smart-phone/tablet-PC/PC/smart-TV, especially using a social name-directory accessible with many simple (e.g. single-character) multilingual domain names (i.e. top level domains : http://ㄱ.com http://ㅏ.com http://김.net http://이.net http://박.net etc.). We applied the statistical parameters estimated from sessions to the analysis of server overload and the effect of advertisements with social plugins. We present real-time estimation schemes for *utilization-level* and *serviceability* based on social interaction sessions for social name-directory contents delivery with social plugins (e.g. Facebook, Twitter) and advertisements (e.g. Google AdSense).

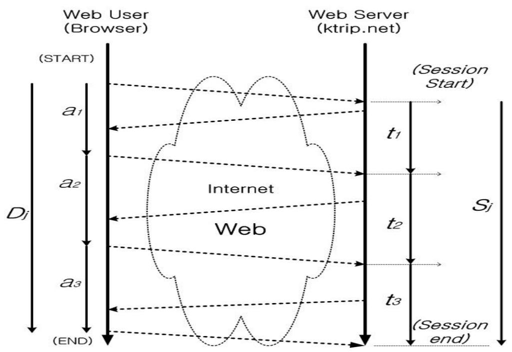
In the following sections, the *utilization-level* and *serviceability* of an interactive web server are defined and presented. The real-time estimation schemes of *utilization-level* and *serviceability* with interactions for name-directory contents with advertisements are presented. The empirical results (http://ktrip.net/display.asp) and data analysis based on real-time estimation schemes with the implementation of an interactive social name-directory service http://ktrip.

net (or http://한국.net) are presented, then the synthesis regarding advertisements (i.e. Google AdSense) is discussed along with other issues of the research. Finally, this study concludes with results of implementation and considerations of future work.

2. Real-time Estimation Schemes for Analysis

2.1 Utilization-level and Serviceability

For service monitoring, detection of server overload, and managing advertisements that generate revenue, we need measurable parameters related to utilization status and the service capability of a specific web server. We define *utilization-level* and *serviceability* of an interactive web server, and we present real-time estimation schemes for *utilization-level* and *serviceability* based on interaction sessions.



〈Figure 1〉 Web Interactions in a Session in Web Activity

In a web activity, there are several interactions in a session as shown in <Figure 1>, which shows the web interaction sequence including accessing an web server (a_1) with an Internet browser, reading a business-card list, reading the contents (a_2) after clicking a listed name in an web server, then finishing the session (a_3) after closing the browser. For real-time estimation, the session duration time can be estimated consistently in a server program running on the social-web server, instead of estimation on the user's side. Instead of estimation of each time interval (t_1, t_2 ⋯) in <Figure 1>, whole session duration time is easy to estimate within a web server.

Session Duration Time =

$$D_j \approx \sum_{i=1}^{n} t_i \approx S_j \ [sec] \qquad (1)$$

A web server program is the appropriate place for the real-time estimation/analysis of session time and usage frequency (i.e. number of sessions within a time period) for simple implementation. The *utilization-level** (we used the notation* to differentiate from the actually estimated *utilization-level* based on sampling discussed in next section) 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].

$$\text{Utilization-level}^* = \sum_{j=1}^{f} D_j \approx \overline{D} \times \overline{f} \qquad (2)$$

Moving average model is suitable for real-time estimation of the statistics (i.e. mean or average) and efficient real-time calculation. We used an exponentially weighted moving average model with the smoothing parameter α (in Eq. (3)). In general, a mean (i.e. average) value of a random variable A with k*th* sample can be estimated in real-time as follows :

$$\overline{A_k} = \alpha A_k + (1-\alpha)\overline{A_{k-1}} \text{ where } 0 < \alpha < 1 \quad (3)$$

We used 0.1 as the smoothing parameter α in our implementation of real-time estimation (i.e. giving 10% weight to the most recent sample and 90% weight to the recent average as an example in equation (4) and following equations) to estimate the mean value of the random variable (i.e. session duration time, utilization-level*).

The mean (i.e. average) value of the session duration time with k*th* sample : $\overline{D_k}$, required for the real-time estimation, can be estimated with a smoothing scheme as follows :

$$\overline{D_k} = 0.1 \times D_k + 0.9 \times \overline{D_{k-1}} \quad (4)$$

The utilization-level* with kth sample can be estimated approximately as follows :

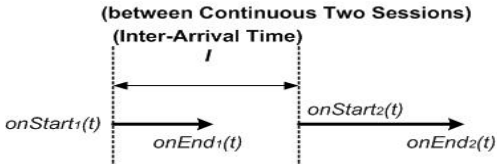
$$\text{Utilization-level}^* = U_k \approx \overline{D_k} \times \overline{f_k} \quad (5)$$

In another way, the mean value of utilization-level* : $\overline{U_k}$, for the real-time estimation can be estimated as follows :

$$\overline{U_k} = 0.1 \times U_k + 0.9 \times \overline{U_{k-1}} \quad (6)$$

$$\overline{U_k} = 0.1 \times (\overline{D_k} \times \overline{f_k}) + 0.9 \times \overline{U_{k-1}} \quad (7)$$

We defined the serviceability (i.e. service ability at the point of performance) of a web-based name-directory server in the interactive service as the maximum number of experienced (serviced) sessions in one second (i.e. [sec]), which can be estimated with the reciprocal of the minimum mean inter-arrival time between continuous sessions, as shown in <Figure 2>.



〈Figure 2〉 Inter-arrival Time between Continuous Sessions in a Web Server

In <Figure 2>, the stochastic inter-arrival time can be estimated as follows.

(Inter-arrival time)

$$I = onStart_2(t) - onStart_1(t)$$

The mean (average) value of the stochastic inter-arrival time I can be estimated. Serviceability [1/sec] of a web server can be estimated on the basis of the reciprocal of minimum mean inter-arrival time I between continuous sessions. The mean value of inter-arrival (i.e. interval) time : I, for the real-time estimation can be estimated with a same smoothing scheme as follows.

$$\overline{I_k} = 0.1 \times I_k + 0.9 \times \overline{I_{k-1}} \quad (8)$$

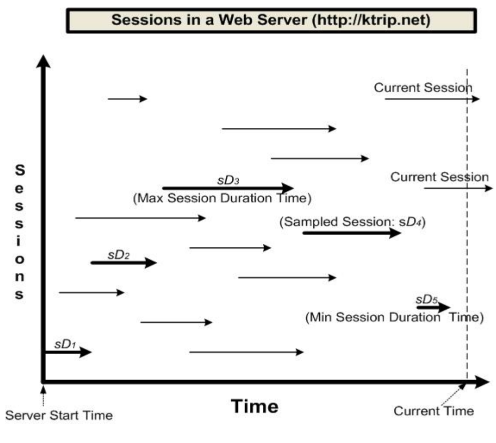
The real-time estimation and analysis of the session arrival rate at the interactive web server for the interactive contents delivery is estimated from the reciprocal of the mean of inter-arrival time I. The serviceability [1/sec] is estimated from the minimum mean of I [sec] to find the maximum ability of web-server (i.e. name-directory server) performance, as follows :

$$Serviceability = \frac{1}{\min(\overline{I_k})} \quad [1/sec] \quad (9)$$

2.2 Implementation for Real-time Estimation Schemes

The right place for implementation of the real-time estimation/analysis was studied considering the requirements and implementation. For real-time estimation in the interactive web services, the session duration time as well as the inter-arrival time between sessions can be estimated consistently in the programs running on the web server, instead of estimation on the user's side. Also, the server program was implemented to estimate the serviceability in real-time.

<Figure 3> shows sessions on a time axis, estimating many aspects of sessions in an interactive name-directory server. In a program in the web server, max/min session duration time, starting/ending time of sessions, inter-arrival time between adjacent sessions could be estimated in real-time for the estimation of the utilization-level and serviceability of an interactive name-directory server.



〈Figure 3〉 Arrival Stream of Sessions and sample Sessions in Server

Estimation with whole session durations instead of samples is difficult and inefficient in real implementation because of many concurrently active sessions (need many timers to calculate all durations). With sampling of session duration time, we used only a single timer efficiently to calculate the sampled duration time in our real implementation. The utilization-level*, which should be estimated in a real-time way as a quantitative parameter, was implemented with following Eq. (10) for the prior Eq. (2). Sequentially sampled sessions for the real-time estimation of the utilization-level* are shown in <Figure 3>. Utilization-level as a stochastic random variable is asymptotically approximated with the multiplication of the mean (i.e. average) of sequentially sampled session duration time sD [sec] and the mean of usage (session) frequency f [1/sec], as following Eq. (12). As an efficient implementation scheme for various interactive web servers, estimations were in real-time for the mean of sampled session duration time sD with only sampled sessions as shown in <Figure 3>, rather than estimation of session duration time D with all sessions :

<Figure 3> shows sessions on a time axis, estimating many aspects of sessions in an interactive name-directory server. In a program in the web server, max/min session duration time, starting/ending time of sessions, inter-arrival time between adjacent sessions could be estimated in real-time for the estimation of the utilization-level and serviceability of an interactive name-directory server.

$$\text{Utilization-level}^{*} = \sum_{j=1}^{f} D_j \approx \overline{D} \times \overline{f} \approx \overline{sD} \times \overline{f} \quad (10)$$

The mean (i.e. average) value of the sampled session duration time with (sequentially sampled) mth sample (instead of whole kth sample) $\overline{sD_m}$ required for the real-time estimation can be estimated with a smoothing scheme in an interactive name-directory server as follows :

$$\overline{sD_m} = 0.1 \times sD_m + 0.9 \times \overline{sD_{m-1}} \qquad (11)$$

The utilization-level based on sampling is defined and can be approximately estimated as follows :

$$\text{Utilization-level} = U_m = \overline{sD_m} \times \overline{f_k} \qquad (12)$$

With the real-time estimation scheme, the mean (i.e. average) value of the utilization-level with (sampled as shown in <Figure 3>) mth sample : $\overline{U_m}$ is estimated in real-time as follows :

$$\overline{U_m} = 0.1 \times U_m + 0.9 \times \overline{U_{m-1}} \qquad (13)$$

$$\overline{U_m} = 0.1 \times \overline{sD_m} \times \overline{f_k} + 0.9 \times \overline{U_{m-1}} \qquad (14)$$

The estimation of frequency with whole session events (with kth sequential sample) is easy in implementation with single timer. The mean (i.e. average) value of the usage (session) frequency with (whole) kth sample $\overline{f_k}$ for the real-time estimation can be estimated in an interactive web server as follows :

$$\overline{f_k} = \frac{1}{mean\,(\text{inter} - \text{arrival time})} = \frac{1}{\overline{I_k}} \ [1/\text{sec}] \qquad (15)$$

Eq. (14) with Eq. (15) can be completed as follows.

$$\overline{U_m} = 0.1 \times \overline{sD_m} \times \frac{1}{\overline{I_k}} + 0.9 \times \overline{U_{m-1}} \qquad (14')$$

The statistics of the inter-arrival time (to get the usage frequency) are estimated with all sessions on a web-based name-directory server as shown in previous <Figure 2>, and the statistics (i.e. mean, max/min of session duration time) are estimated in real-time with sampled sessions as shown in <Figure 3>. A server program in the http://ktrip.net server was implemented for the real-time estimation of stochastic random variables (i.e. session duration time, utilization-level, usage frequency, serviceability). Statistics for the mean/max/min value of each stochastic parameter are estimated in real-time.

3. Empirical Results and Data Analysis

Within a server program for a social-web name-directory server : http://ktrip.net, many interesting parameters can be estimated in real-time. Then, with the program, http://ktrip.net/display.asp, the estimated statistics of the interesting parameters can be displayed on the screen. In <Table 1>, the following results (No. 1~24) for a social-web site (http://ktrip.net or http://한국. net), which is accessed frequently (mean inter-arrival time : 0.0545~60 [sec] as shown at No. 15/16 in <Table 1>), are displayed in real-time at the URL http://ktrip.net/display.asp. These empirical results were gathered 7 times over two years (from Oct. 31, 2010 to Nov. 12, 2012) because the server program restarted over 7 times after the rebooting of the server (http://ktrip.net or http://한국.net). We differentiated the 7 sample groups with Group (A), Group (B), Group (C), Group (D), Group (E), Group (F) and Group (G) in <Table 1>.

Each parameter for real-time estimation of the empirical data is explained as follows : A1 (i.e. No. 1 in estimated value in sample group A) is for the cumulative total number of sessions (in the same way, B1 is from sample group B; C1, D1, E1, F1 and G1 are from sample group C, D, E, F and G, respectively); A6 is for the current mean of usage frequency of sessions based on the reciprocal of mean inter-arrival time (A8) between continuous sessions (with kth sample in Eq. (15)); A10 is for the current number of actually sampled (for real-time estimation of session duration time; related to the sequentially sampled mth sample) sessions; A18 is for the (cumulative) maximum value of mean session duration time (related to the sequentially sampled mth sample in Eq. (11)); A20 is for the (cumulative) maximum value of utilization-level (related to Eq. (12)); A21 is for the (cumulative) maximum value of mean utilization-level (related to equation (14 or 14')); and finally, A24 is for the estimated (cumulative maximum) serviceability. The utilization-level and serviceability are estimated, analyzed and discussed in terms of the overload of name-directory server and the revenue from advertisements with Google AdSense.

The utilization-level from Eq. (12) correlates to the number of active sessions in the interactive web server (i.e. http://ktrip.net or http://한국.net).

(Utilization-level)

$$U_m = \overline{sD_m} \times \overline{f_i} \approx Number-of-Sessions \quad (16)$$

From Eq. (16) showing our estimation scheme 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.

The mean value of the utilization-level can be compared to the mean number of active sessions. <Figure 4> shows the relationship between the mean utilization-level and the mean number of sessions in the web server. The following relationship (Eq. (17)) (between the mean utilization-level and the mean number of active sessions) can be derived from Eq. (16) without statistical outliers.

(Mean Utilization-level)

$$\overline{U_m} \approx \overline{\overline{D_m} \times \overline{f_k}} \approx \overline{Number-of-Sessions} \quad (17)$$

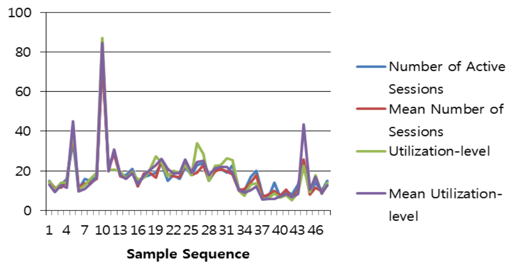
As examples, we can compare the mean utilization-level (14.43 in G7; <Table 1>) to the mean number of current active sessions (14.05 in G2) in the name-directory server http://ktrip.net. Similarly, we can compare the mean utilization-level (14.15 in F7) to the mean number of current active sessions (14.85 in F2; <Table 1>).

In <Figure 4> (sampled every 30 minutes on Nov. 9, 2012) with the maximum value of the utilization-level between samples no. 9 and 10 among recent 48 samples (among sample group (G)), there is a rapid increase in mean number of sessions, 75.59, compared to the prior mean number of sessions, 16.08. The rapid increase in the utilization-level is caused by the rapid increase in the active number of sessions.

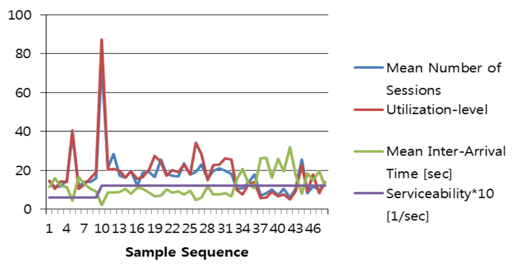
In <Figure 5>, the minimum mean inter-arrival time approached 2.06 [sec] from 9.33 [sec] between samples No. 9 and 10; therefore the increased number of sessions affected the utilization-level and serviceability (from 0.61 to 1.2) in the interactive name-directory server.

〈Table 1〉 Estimated Parameters from Real-time Estimation Schemes

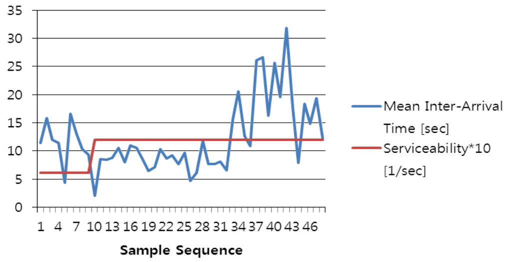
No	Estimation Parameter	Group (A)	Group (B)	Group (C)	Group (D)	Group (E)	Group (F)	Group (G)
1	Cumulative Number of Sessions	26,978	261,588	174,868	572,916	723,059	1,068,752	65,635
2	Mean Number of Active Sessions	20.2	9	17.8	12.72	14.6	14.85	14.05
3	Start Date of Server	Oct 31, 2010	Nov 27, 2010	Jan 19, 2011	Mar 9, 2011	Jul 10, 2011	May 24, 2012	Nov 1, 2012
4	End Date of Server	Nov 5, 2010	Jan 17, 2011	Feb 19, 2011	May 16, 2011	Oct 9, 2011	Oct 28, 2012	Nov 12, 2012
5	Mean Session Duration Time [sec]	147.3	155.5	172.3	158	177	181.86	175.56
6	Mean Usage Frequency [1/sec]	0.275	0.056	0.086	0.062	0.083	0.076	0.077
7	Mean Utilization-level	24.07	7.8	14.49	8.49	14.74	14.15	14.43
8	Mean Inter-arrival Time (between Sessions) [sec]	3.63	17.8	11.57	16	12.0	13.1	13.0
9	Number of Completed Sessions	26,950	261,577	174,847	572,901	723,044	1,068,736	65,620
10	Number of Sampled Sessions	2,739	24,237	15,109	32,648	43,342	74,616	5,300
11	Session Duration Time [sec]	147	163	175	177	179	176	176
12	Max Number of Active Sessions	81	188	109	224	188	181	132
13	Min Number of Active Sessions	1	1	1	1	1	1	1
14	Max Mean Active Sessions	77.5	181.5	102.7	217.4	50.3	174.4	125.6
15	Max Mean Inter-arrival Time [sec]	55.82	58.99	54.22	45.99	50.6	60.0	50.92
16	Min Mean Inter-arrival Time [sec]	1.12	0.235	0.356	0.201	0.4595	0.0545	0.83
17	Max Inter-arrival Time (between Sessions) [sec]	237	529	256	183	194	234	222
18	Max Mean Session Duration Time [sec]	579.3	1057.7	975.4	646.6	788.5	1,959	629
19	Min Mean Session Duration Time [sec]	73	64.6	44	69.5	71	76.7	89.1
20	Max Utilization-level	137.3	724.7	481.7	845.7	370.8	2,827	215
21	Max Mean Utilization-level	91.18	470.45	362.33	542.69	263.2	1,656.3	175.3
22	Max Session Duration Time [sec]	950	1233	1076	660	831	2,237	635
23	Min Session Duration Time [sec]	64	10	44	14	64	14	70
24	Serviceability [1/sec]	0.893	4.256	2.806	4.977	2.176	18.355	1.2



〈Figure 4〉 Real-time Estimation of Active Sessions and Utilization-level



〈Figure 5〉 Real-Time Estimation of Mean Inter-Arrival Time and Serviceability



〈Figure 6〉 Serviceability and Mean Inter-Arrival Time

From <Figure 5>, <Figure 6> and Eq. (18), the serviceability is the reciprocal of minimum of mean Inter-Arrival Time.

(Serviceability)

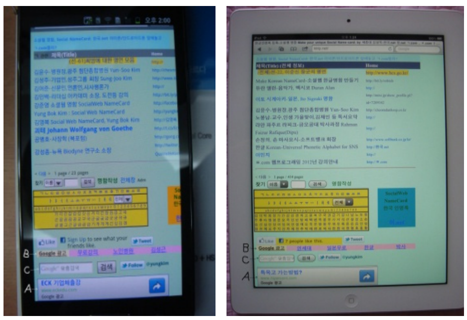
$$S_k = \frac{1}{\min(\overline{I_k})} = \frac{1}{\min(meanInterArrivalTime)} [1/sec] \quad (18)$$

The serviceability of the http://ktrip.net server was 1.2 in the sample group (G); it had been 4.977 [1/sec] until Oct. 9th, 2011 as shown in sample groups (A) through (E), and later became 18.355 [1/sec] (in the sample group (F)) with 1,068,736 (as shown in F9; <Table 1>) experienced sessions within a 5 month period in 2012. The maximum value of serviceability experienced should be stored as a statistical value of an estimated parameter during server's operation.

We used the measurable parameters about utilization status and the service capability of the web server (http://ktrip.net or http://한국.net) for service monitoring, detection of server overload and managing advertisements. As a specific example for practical application, many organizations outsourcing their servers with cloud computing can monitor data on utilization-level and serviceability of the servers for their quantitative service-level contracts with outsourcing companies or upgrading quantitatively the performance of own web-server with proper investment.

4. Advertisements and Discussion

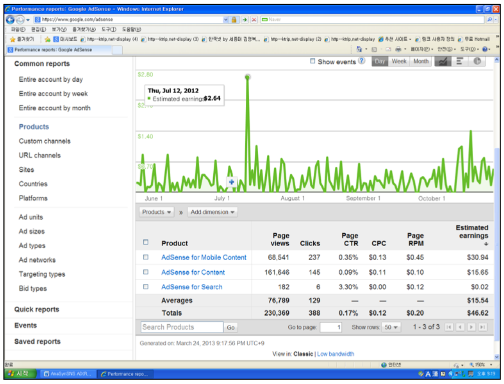
Social name-based contents in the interactive name-directory server http://ktrip.net have social plugins with Facebook/Twitter as well as Google AdSense. Our web site used Google AdSense advertisements because of their convenient plugins as well as customer support web-service. Moreover, Google provided the revenue related statistics well. In <Figure 7>, we placed three kinds (left bottom notation A, B, C) of Google AdSense advertisements in our name-directory server http://ktrip.net (or http://한국.net). A was for the mobile-content AdSense; B was for the content AdSense; and, C was for the search Ad Sense. The position of A, B and C could be changed for better exposure of advertisements. This data for the specific account http://ktrip.net was shown in the performance reports of Google AdSense http://google.com/adsense. Among the three kinds of advertisements, mobile advertisements (i.e. mobile-content AdSense) were the major revenue source.



〈Figure 7〉 Advertisement in Web Service (ktrip.net) with Smart Phone/iPad

<Figure 8> shows specifically a daily revenue of three kinds of advertisements (data from May 24, 2012 until Oct. 28, 2012) related to sample Group (F) in <Table 1>. The impact of social plugin SNS messages to the interactive name-directory server http://ktrip.net (or http://한국.net) was related to the advertisement, and eventually relates to revenue.

Estimated earnings (i.e. revenue) was the sum of Clicks×CPC as shown in <Figure 8>. If we want to estimate the upper bound of expected revenue of an arbitrary web site with Google AdSense, then we may forecast the upper bound revenue synthetically on the basis of the max mean utilization-level, max serviceability, and the planning period of an advertisement.



〈Figure 8〉 Daily Revenue of different AD Kinds (Sample Group (F))

This is an application to advertisement (AD) and revenue on the basis of the estimated parameters of a targeted web server. The advertiser pays the advertisement fee to the AD service provider (e.g. Google), and Google pays the AD site depending upon the ADrevenue : R, which becomes the revenue of the AD site such as http://ktrip.net (or http://한국.net). We studied the RevenueFactor with a refined scheme. To increase the ADrevenue : R, the service provider (web site) should try to increase the mean utilization-level and serviceability of the name-directory server within a same condition of the ADperiod and RevenueFactor.

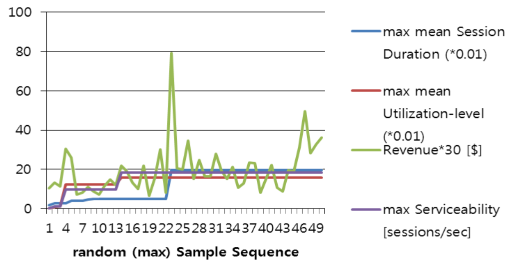
We tried to estimate synthetically the upper bounds of revenue. If we can find the revenue factor from Google AdSense statistics for a certain period, then we can forecast the upper bound of advertisement revenues of a web site considering AdSense advertisements with social plugins. To obtain the upper bound of estimated revenue ADrevenue, we can consider the following equation.

(ADrevenue [$])

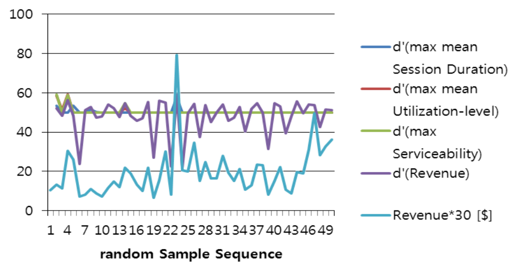
$$R \leq Max(\overline{Utilization-level}) \times Serviceability \times ADperiod \times RevenueFactor \quad (19)$$

In <Figure 8>, we found that the largest daily revenue (2.64 USD) occurred on July 12, 2012 with the largest number of clicks (19) and page CTR (1.52%) in sample Group (F) (in <Table 1>). An SNS message (related to social-plugin business-cards of leading candidates in web server http://ktrip.net in a Facebook Timeline (1 : 30 PM on July 12, 2012) caused the web-based name-directory server to overload because of instantly increased traffic. The total number of our Facebook friends (5,000 people) and Facebook subscribers (around 2,400 following people) amounted to 8,400 including 1,000 Twitter followers, and a portion of them simultaneously clicked the linked contents on the interactive web server http://ktrip.net. The contents of the message may have aroused interest to a number of Facebook/Twitter friends, because the message contained social business-cards for leading candidates of the Korean presidency.

With sample data among sample Group (F) in <Table 1>, we compared the estimated parameters (such as max mean utilization-level and max serviceability) to revenue around the sample with the largest value (occurred on July 12, 2012, sample sequence no. 23 in <Figure 9>) as shown in <Figure 8>. <Figure 9> shows the revenue increase with (randomly) sampled (cumulatively) max mean session duration, (cumulatively) max mean utilization-level and (cumulatively) max serviceability (among sample Group (F)) around sample sequence no. 23 of largest revenue and overload. In <Figure 10>, for easy comparison around samples of revenue increase with (cumulatively) maximum parameters, we compared the differential ratio of revenue to the differential ratio of estimated parameters (such as max mean session duration time, max mean utilization-level and max serviceability).



〈Figure 9〉 Revenue Increase vs. Maximum Parameters (among Sample Group (F))



〈Figure 10〉 Revenue vs. Differential Ratio Parameters (among Sample Group (F))

<Figure 10> shows modified differential ratio parameters (max mean session duration time, max mean utilization-level, serviceability, revenue) from samples among sample group (F) (in <Table 1>) derived from Eq. (20) as an example with the max mean utilization-level : $Max(\overline{utilization-level})$. For easy comparison, the Eq. (20) is modified to Eq. (21) called a modified differential ratio parameter.

$$d(Max(\overline{utilization-level}))_m = \frac{(Max(\overline{utilization-level}))_m - (Max(\overline{utilization-level}))_{m-1}}{(Max(\overline{utilization-level}))_m} \quad (20)$$

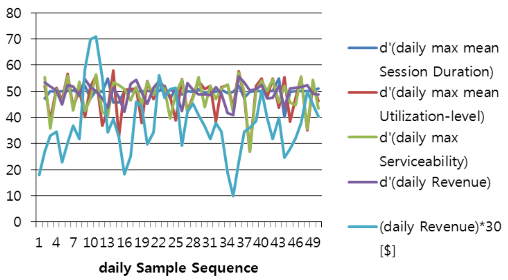
$$d'(Max(\overline{utilization-level}))_m = d(Max(\overline{utilization-level}))_m \times 10 + 50 \quad (21)$$

In <Figure 9> (and <Figure 10>), the large increase of revenue seems to be correlated with the (modified differential ratio of; in <Figure 10>) max mean session duration time as shown in the sample sequence No.23. Overload in web server occurred in sample sequence No.23, and the session duration time became abnormally long. In the sample sequence No.4 and No.13, the increase of revenue seems to be correlated with the (modified differential ratio of; in <Figure 10>) max mean utilization-level as well as to the differential ratio of max serviceability.

Daily estimated statistics of parameters are needed for better synchronized comparison with daily revenues reported in Google AdSense.

<Figure 11> shows modified differential ratio parameters (daily max mean session duration time, daily max mean utilization-level, daily max serviceability, daily revenue) from daily estimated samples of parameters from Jan. 2nd to Feb.

21st in 2013 for synchronized comparison to daily estimated revenue in Google AdSense reports.



〈Figure 11〉 Revenue vs. Differential Ratio Parameters (Samples Jan.~Feb., 2013)

With our social name-directory service, we will gather more sufficient data with bigger advertisement revenue as a future research work for further justification. Other issues for improvement and refinement will be also considered, for example, daily statistics of parameters for daily revenue analysis, detailed correlation analysis (between revenue and estimated parameters), statistical outlier analysis, cloud computing service (to prevent server overload), semantic search (Yndurain et al., 2012) and geo-location application (Goth, 2013). As the influence of SNS has grown extensively, potential threats to privacy have also become pervasive (Kim et al., 2014); we will consider the privacy and security as further research issues.

5. Conclusion

For service monitoring, detection of server overload, and examining advertisements affecting revenue, we defined and estimated the utilization-level and serviceability of an interactive name-directory server. Real-time estimation schemes for utilization-level and serviceability based on interaction sessions were presented for social name-directory contents delivery with social plugins (e.g. Facebook, Twitter) and advertisements (e.g. Google AdSense). The empirical results were presented on the basis of the implementation in a web-based name-directory server, http://ktrip.net (or http://한국.net) which has social plugins with Google AdSense advertisements. A practical application based on the presented utilization-level and serviceability in social name-directory services was researched with the Google AdSense advertisements affecting revenue. The normal increase of utilization level increased AdSense advertisements and affected revenue. The maximum serviceability of the name-directory server limited the increase of advertisements affecting revenue. For upgrading own server after monitoring of the utilization-level and serviceability (as well as for outsourcing server with cloud computing), we could plan to improve the performance of the name-directory server http://ktrip.net quantitatively. A semantic search with advertisements and privacy/security issues will be also considered in our future work.

References

Ferreira, N., "Social Networks and Young People : A Case Study", *International Journal of Human Capital and Information Technology Professionals*, Vol.1, No.4, 2010, 31-54.

Garcia, J., F.J. Garcia-Penalvo, R. Theron, and P. O. Pablos, "Usability Evaluation of a Visual Modelling Tool for OWL Ontologies", *J. UCS*, Vol.17, No.9, 2011, 1299-1313.

Goth, G., "Where It's At Mapping Battle Highlights New Era of Revenue and Development Models", *IEEE Internet Computing*, Vol.17, No.1, 2013, 7-9.

Kim, H., K.K. Kim, and H. Lee, "Factors Affecting Intention to Use Security Functions in SNS", *Journal of Information Technology Services*, Vol.13, No.2, 2014, 1-17.

Kim, Y.B., "A Ubiquitous Social Community Portal Service for Social Networking with Convenient Accessibility", *LNCS*, Vol.4061, 2006, 263-270.

Kim, Y.B., "Real-time Analysis of Time-based Usability and Accessibility for Human Mobile-Web Interactions", *J. UCS*, Vol.16, No. 15, 2010, 1953-1972.

Kim, Y.B., "Accessibility and Usability of User-centric Web Interaction with a Unified-Ubiquitous Name-based Directory Service", *World Wide Web*, Vol.13, No.1/2, 2010, 105-120.

Kim, Y.B., "Utilization-level and Serviceability of a Social Name-card Portal for QoS in a Cloud Social Networking Service", *J. UCS*, Vol.18, No.11, 2012, 1523-1543.

Lampe, C. and N. Ellison, "Understanding Facebook : Social Computing Isn't 'Just' Social", *IEEE Computer*, Vol.45, No.9, 2012, 98-100.

Modica, G.D., O. Tomarchio, and L. Vita, "A P2P Based Architecture for Semantic Web Service Discovery", *International Journal of Software Engineering and Knowledge Engineering*, Vol.21, No.7, 2011, 1013-1035.

Park, J., R. Sandhu, and Y. Cheng, "A User-Activity-Centric Framework for Access Control in Online Social Networks", *IEEE Internet Computing*, Vo.15, No.5, 2011, 62-65.

Park, S., "Comparison of Recommendation Using Social Network Analysis with Collaborative Filtering in Social Network Sites", *Journal of Information Technology Services*, Vol.13, No.2, 2014, 173-184.

Shim, E.S., J.H. Lee, and B.G. Jeong, "A Study on influence factors of User's Continuous Use Intension in Firm Promotion Social Network Service : by Mediating the Moderator of the Operation Objective Types in Facebook Fan Page", *Journal of Information Technology Services*, Vol.12, No.4, 2013, 41-57.

Tai, S., P. Leitner, and S. Dustdar, "Design by Units : Abstractions for Human and Compute Resources for Elastic Systems", *IEEE Internet Computing*, Vol.16, No.4, 2012, 84-88.

Taivalsaari, A. and K. Systa, "Cloudberry : An HTML5 Cloud Phone Platform for Mobile Devices", *IEEE Software*, Vol.29, No.4, 2012, 40-45.

Yndurain, E., D. Bernhardt and C. Campo, "Augmenting Mobile Search Engines to Leverage Context Awareness", *IEEE Internet Computing*, Vol.16, No.2, 2012, 17-25.

◈ About the Authors ◈

Yung Bok Kim (yungbkim@sejong.ac.kr)

Professor Yung Bok Kim is in Dept. of Computer Engineering at Sejong University in Seoul, Korea. He received a BS degree in Electrical Engineering from Seoul National University, an MS degree in Electrical Engineering from the Korea Advanced Inst. of Science and Technology (KAIST) and a PhD in Electrical and Computer Engineering from Illinois Institute of Technology (IIT), Chicago, USA. Dr. Kim's areas of interest are computer networks, mobile Internet, unified-ubiquitous web-based information service, and real-time name-based directory service for SNS (http://한국.net http://ㄱ.com http://ㅏ.com http://김.net etc.).