A Case Study on Mobile Internet Use Analysis Implications from Long-term Mobile Internet Observation

Toshihiko Yamakami[†]

The mobile Internet is rapidly growing. The methodologies to capture the mobile Internet use are still in the early stage. The constraints in the mobile handsets lead to the requirements to capture the contexts surrounding the small display. Based on the proposed context cue exploration approach, a case study is performed to capture the time context factor over a long span of time. Cross-carrier comparisons show the differences of markup languages and network characteristics do not affect much over a long span. The real world issues to impede 2-year observation are discussed. The implications of the mobile Internet usage analysis are presented.

1. Introduction

The mobile Internet covers a wide variety of Internet access from wireless LAN to mobile handsets. The author focuses the mobile handset based Internet because it occupies the significant part of the today's Internet in Japan. The active mobile handset based Internet users exceed 50 million, easily outnumbering other Internet access methods.

Rapid penetration of internet-enabled mobile handsets is witnessed in worldwide[3], especially in Asia. However, the systematic observation on the user behavior is still to be explored. It is highly likely that headset-based Internet users show different user behaviors. However, the methodologies to capture the significant characteristics are still immature. The reason why the mobile Internet diffusion differs from country to country is still an interesting research topic. The context cue exploration approach and the paradigm behind it are proposed. As an example of the compensating context, the lifetime session is analyzed for a long span of time using the identifier tracking analysis method[8]. The method is simple and effective, however, the follow-up analysis over a long span of time encounters several challenges. The long-term analysis needs a careful and consistent logging design, however, the real world issues in the mobile Internet with dynamism and unpredictable

† ACCESS, 2-8-16 Sarugaku-cho, Chiyoda-ku, Tokyo, JAPAN, e-mail:yam@access.co.jp

factors make it hard. This paper presents the follow-up analysis and implications of the long-term transitions of the mobile Internet in Japan.

2. Challenges and Issues

2.1 Constraints and Markup Languages

The micro-browser in the mobile handsets presents the significant constraints in the mobile Internet services.

- Small display space
- Small memory capacity
- Low-power CPU
- Narrow bandwidth and low-speed network

The constraint directly leads to the new skills of the content authoring and content marketing in the mobile-handset Internet services. Japan is the first country to reach 50 million active users with its handset-enabled Internet users. The subscribers for handset-enabled Internet are depicted in **Fig. 1**. It is important to use this opportunity to understand the user behaviors to learn improvement requirements from this rapidly changing mobile Internet application domain.

In the past, the content language for a small device was parted into two categories, based on HTML or Wireless Markup Language(WML[6]), a language specified in WAP Forum. In Japan, to make the things worse, there have been three languages, NTTDoCoMo's i-mode HTML(derived from Compact HTML[5]), au's Handheld Device Markup Language(HDML)
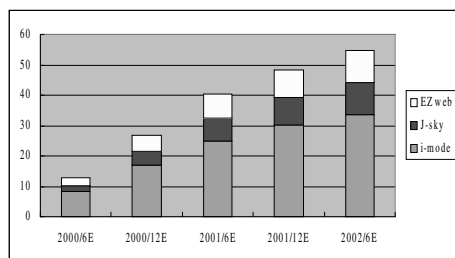
Fig. 1 Subscriber Growth during 2000-2002 (million users)

for EZweb, and J-PHONE's Mobile Markup Language(MML) for J-sky. HDML is an origin language for WML. Recently, XHTML Basic[1] has been established for the next generation small device markup language. WAP Forum uses XHTML Mobile Profile[7], a slightly enhanced version of XHTML Basic. The mobile Internet content language will be based on XHTML Basic or XHTML Mobile Profile. The comparison between different markup languages in the mobile Internet was rarely available in the past literature. Cleary the compact HTML had the content authoring advantage for minimized time to market. However, Fig. 1 shows each carrier continuously increases its mobile Internet subscribers over a long span of time despite of the markup language.

3. Purpose of the Study and Related Works

3.1 Purposes

This study aims at understanding the mobile Internet form the context cue exploration approach. The context cue exploration approach is based on the assumption that the surrounding contexts define the fundamental characteristics on the mobile Internet use. Especially, this paper aims at exploring the time factor in the mobile Internet using the long-term log observation including cross-carrier observations.

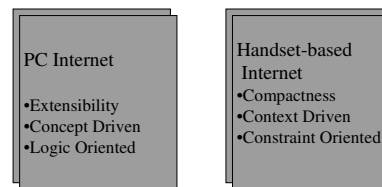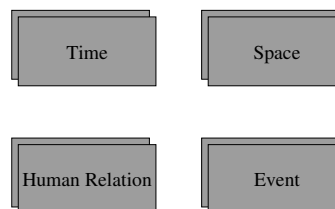**Fig. 2** Paradigm Comparison between PC Internet and Handset-based Internet

Fig. 3 Context Factors to Compensate the Handset-based Constraints

3.2 A Context Cue Exploration Approach

It is important to note that the Japanese culture had the significant effect on the mobile Internet usage from news/transaction to socialization/entertainment. It is partly because that the mobile handset's limitation of display just needs the additional context cues in the mobile Internet. The difference between PC Internet approaches and mobile-handset based Internet approaches are depicted in **Fig. 2**. The role of context cue should be investigated for further examination of the mobile Internet usage. The context cues may include various aspects of social life. **Fig. 3** illustrates some of the major context cues for the mobile Internet.

Fig. 2 shows the comparison between the PC Internet and the handset-based Internet.

Fig. 3 shows factors to compensate the constraints in the handset-based Internet.

3.3 Related Works

The literature about the success of the mobile Internet is rarely available at the moment.

The user interface study such as[2] illustrates the user interface constraints on the mobile handsets, however, it did not provide any cue for the rapid mobile Internet growth mainly in Asia. The social approach like[9] is still in the early stage to identify the culture factor in the mobile Internet. The author is inspired by Grudin's idea[4] that the digital world is too large and that we need to pursue how to split the digital world in an effective way. He performed the multiple display user interface research, however, the concept to pursue how to effectively split the digital world is the most important issue in the mobile Internet. The author performs the initial study on the mobile Internet usage patterns[8] with the approach unique identifier tracking analysis to make use of the unique mobile Internet characteristics. However, this research was done on relatively short-term log (6 months). In addition, the cross-carrier comparison was not performed.

4. A Case Study of Cross-Carrier Unique Identifier Tracking Analysis over a Long Span of Time

4.1 Analysis System Configuration

One of the key factors in the mobile Internet is the user track ability using the unique user identifier from the mobile carriers. It is distinguished that the mobile Internet is aimed at the paid services. It is unique in the mobile Internet service domain. The unique user identifier is an important tool to track the user behavior. **Fig. 4** shows the outline of the analysis system configuration. The system is implemented by PHP3. The analysis system currently runs on PHP 4.1.2. The server logs occupies the significant amount of server storage space. The log is compressed in a per-month manner.

- tracking each user behavior individually
- tracking each user even when the device is upgraded
- tracking each user over different application services

However, during the unique user identifier tracking analysis, the author encountered the following challenges:

- user identifier can be affected by the service upgrade by the carrier
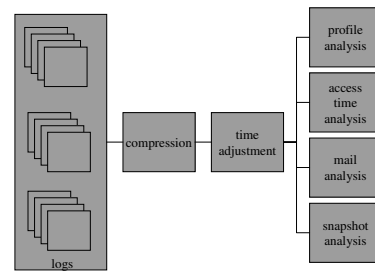- each carrier provides a different access mechanism to the user identifier

Fig. 4 Analysis System Configuration

- due to the privacy policy, multiple different user identifiers may be provided by the carrier
- due to the different user identifier policy, user identifier tracking may be affected by the service menu structure

It is important to make the consistent observation design for the accuracy of the long-term observation. However, the real world issues in the mobile Internet makes it hard to be consistent for long years. First, the burst traffic characteristics make the mobile Internet server system complicated with backup servers. The increasing number of the servers makes the time synchronization difficult. In addition, the mobile Internet runs very long time without halting because the always-on characteristics of the mobile Internet demand 24-hour operation. Also, the carrier may update the user identifier policy when the major system upgrade is planned. Such a change is hardly predictable, and it is frequent that such a change will not make public in a timely manner. Details of the carrier operation may require of the change of the operations, leading to the changes of the observation plan. The dynamism exists in the content service itself. The menu structure and charging structure may affect the observation plan. In addition, in the long-term observation, it is uncommon to obtain the perfect operation records. The details of the server system operations like shutdowns and restarting are difficult to trace after a long span of time. At the time of the launch, it was hardly predictable to witness such a rapid growth in Japan. The server configuration is frequently changed to cope with the rapid growth.

4.2 How to Deal with the Time Synchronization Problem

During the long-term observation, it is common to capture the time discrepancy between servers. There are several methods to identify such discrepancy.

- using the database information
- using the order of the frequent used commands
- using the information on the symphonized commands

The database keeps the single track of the user subscription records. The frequent used commands have natural orders of the menu structure of each service. Some management commands have the synchronized nature because the statistics commands needs queries to the other servers. In this research, the author analyzes one of the management commands that have the synchronized nature. The result is shown in **Table 1**. It shows an example of how a distributed server environment has time synchronization problem over a long span of time. Over more than 2 years, the server time synchronization problem may affect the long-term observation especially on the time context.

There are two factors: an incremental factor and an operational factor. Many operational details are difficult to identify after a long span of time. In this example, the major discrepancy seems to occur in July 2001. The details are for further studies.

Table 1 An Example of Distributed Server Environment Time Synchronization Problem

YYMM	server-A and server-B (sec)	server-A and server-C (sec)
0105	6	2
0106	15	5
0107	24	10
0108	32	975
0109	41	954
0110	50	970
0111	59	968
0112	67	964
0201	76	963
0202	85	958
0203	93	955
0204	102	942
0205	111	949
0206	121	946
0207	131	940

Note: YYMM denotes the least significant 2 digits
of year and 2 digits of month. "0105" means May 2001.

4.3 A Case Study

Using the time synchronization information obtained in the long-term operation, the i-mode, EZweb and j-sky service usage is analyzed to identify the time context transitions in the mobile Internet. This study aims at analyzing the easy-come and easy-go characteristics of the mobile Internet using the same metrics on the cross-carrier analysis. The case study is performed on the medium size news site officially listed on i-mode, EZweb, and j-sky. The large size system with more than a million users needs a further analysis tool performance study. Usually, the million user level server logs are too large to analyze. The content of these services are slightly different among carriers partly due to the carriers' intentions to make differentiation. **Table 2** shows the session duration in i-mode without time synchronization technique from September 2000 to May 2001. The heavy users heavily affect the average session length within 256 sec interval. Therefore, the medium value is also shown. When the command interval is less than 256 seconds, the next command is considered to belong to the same session. The interval between the first command and the last command in the session is observed. When the command has more than 256 seconds from/to the adjacent commands, it is considered to be isolated commands. The relative ratio of the isolated commands is also shown.

Table 3 illustrates the same result from May 2001 to June 2002 with time synchronization technique. The May 2001 result is consistent with the previous result. However, the consistent increasing trend shown in Table 3 is not found in Table 2. The contributing factor of this change is still for further studies.

Table 2 An Example of i-mode session length at 256 sec session interval

YYMM	average length	medium length	medium session duration (sec)	ratio of isolated sessions (%)
0009	2.81	1	92	100
0010	4.01	1	100	42
0011	4.30	1	101	34
0012	4.77	1	101	29
0101	4.41	1	95	31
0102	5.58	2	99	21
0103	5.64	2	96	22
0104	5.64	2	99	19
0105	5.84	2	98	18

Table 4 shows the EZweb result from February 2001 to June 2002. The increasing session length corresponds to the increase of the number of the commands in a session. There is a trade-off between session duration and the ratio of isolated sessions. From this result, the author assumes that the increasing session duration, time context, comes from the departures of low-use users. This effect can be commonly observed in the paid mobile Internet services.

Table 5 depicts the session length analysis result on a service on j-sky. The trend is also similar in j-sky partly because that the content is similar. The difference between markup languages does not play a significant role in the usage in this use case.

The results consistently show the longer session. However, even the trend is same among carriers, the duration is only 300 seconds, relatively shorter than that of PC Internet. The mobile Internet is characterized as 24-hour frequently visited short sessions. The further studies will cover the followings:

- effect of content length constraints
- users' learning effect
- effect of the core user effect (the infrequent users leave, then core users remain)

5. Implications from the long-term dynamism in the mobile Internet

The mobile Internet grows in a surprisingly rapid manner. However, from the long-term analysis, we can witness some of the key issues of the saturation of the mobile Internet services. From 2001, it is consistently observed that the longer session duration in a cross-carrier manner. This is partly due to the loss of the infrequent users. It is remarkable to see the walled garden business model in the mobile Internet. The carriers collect the charges for the paid services for the content providers. Even with the high growth of the mobile Internet in Japan, the ratio of the content increase is higher than the increase of the subscribers. This shows the consistent decrease of the low usage paid users

Table 3 An Example of i-mode session length at 256 sec session interval

YYMM	average length	medium length	medium session duration (sec)	ratio of isolated sessions (%)
0105	6.01	2	98	19
0106	7.45	2	106	22
0107	12.12	4	127	23
0108	6.68	2	135	11
0109	7.21	2	155	10
0110	7.78	2	165	10
0111	7.59	2	178	9
0112	7.24	2	195	8
0201	6.86	2	208	8
0202	7.22	2	223	8
0203	7.76	2	242	9
0204	7.50	2	260	8
0205	7.79	2	268	9
0206	7.13	2	276	11

Table 4 An Example of EZweb session length at 256 sec session interval

YYMM	average length	medium length	medium session duration (sec)	ratio of isolated sessions (%)
0102	2.48	1	70	100
0103	3.04	1	72	81
0104	3.03	1	69	80
0105	3.53	1	68	80
0106	5.07	1	80	88
0107	10.80	2	103	73
0108	5.37	1	107	34
0109	4.85	1	131	49
0110	5.89	1	141	40
0111	5.99	1	158	35
0112	5.21	1	176	30
0201	4.72	1	189	34
0202	5.26	1	212	29
0203	5.57	1	226	31
0204	6.02	1	245	26
0205	5.87	1	257	26
0206	6.66	1	269	24

Table 5 An Example of j-sky session length at 256 sec session interval

| YYMM	average length	medium length	medium session duration (sec)	ratio of isolated sessions (%)
0012	4.51	2	165	100
0101	4.67	2	165	68
0102	5.12	2	176	50
0103	4.98	2	168	48
0104	5.45	2	168	48
0105	5.49	2	160	52
0106	7.20	3	174	49
0107	10.83	6	203	51
0108	5.24	2	192	41
0109	5.77	2	197	59
0110	7.33	3	202	40
0111	6.60	2	208	37
0112	6.59	3	237	28
0201	6.09	2	244	30
0202	7.26	4	272	27
0203	7.62	4	284	25
0204	7.13	3	285	25
0205	8.13	3	307	23
0206	8.26	4	309	15
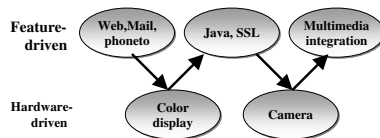
Fig. 5 Stage Transitions in the mobile Internet

in each service. **Fig. 5** shows the stage transitions in the mobile Internet.

This trends show the step-by-step diffusion of the mobile Internet in features and hardware improvement in turn. Using these step-by-step driving forces for the mobile Internet, the services shift from news/transaction services in the early stage, to socialization, then entertainment in the current stage.

- news/transaction services
- socialization services (bulletin boards, chat room)
- entertainment services (games, movie)

The time analysis over a span of time does not reflect these stage transitions. The methodology to capture these cross-service analysis is for further studies. The in-depth analysis in this field may contribute to the worldwide expansion of the mobile Internet by providing the rigid cues for the service transition identification.

6. Conclusions

The methodologies to identify the mobile Internet characteristics are still present research challenges. The uniqueness of the mobile Internet is that the constraints of the handset environment inevitably invite the heavy effect of the context to compensate the constraints. The author presents the context cue exploration approach to understand the situation. In addition, the rapid growth presents an interesting research topic to analyze the long-term user behavior transitions for the first time in the world. A case study of a long-term observation is presented in a cross-carrier manner to make use of the identifier tracking method unique to the mobile Internet. The 2-year observa-

tion about the session lifetime analysis is performed. The real world problems like a time discrepancy problem are presented. The future methodologies should be well prepared against these issues. The results are consistent among carriers. This indicates that the various differences among mobile Internet services are not crucial. The increasing session duration is observed, however, it is about 300 second per session, typical to the mobile Internet. In addition, the stage changes of the mobile Internet are outlined. The long-term mobile Internet logs are only recently available in the world. The challenges include the methodologies to identify the long-term usage transitions.

References

1) Baker, M., Ishikawa, M., Matsui, S., Stark, P., Wugofsky, and T., Yamakami, T.: XHTML™ Basic, W3C Recommendation 19 December 2000, (Dec. 2000). (available at http://www.w3.org/TR/xhtml-basic)
2) Buchanan, G., Farrant, S., Jones, M., Thimbleby, H., Marsden, G., and Pazzani, M.: Improving mobile internet usability, The tenth international World Wide Web conference on World Wide Web ACM Press, (2001).
3) Cerf, V.: Beyond the Post-PC Internet CACM, Vol. 44, No. 9, pp. 34–37 (Sep. 2001)
4) Grudin, J.: Participating digital worlds: focal and peripheral awareness in multiple monitor use ACM CHI ' 01, ACM Press, pp. 458–465 (2001)
5) Kamada, T.: Compact HTML for Small Information Appliances W3C Submission, W3C Note, (Feb. 1998) (available at http://www.w3.org/TR/1998 /NOTE-compactHTML-19980209)
6) WAP Forum™: Wireless Markup Language Version 1.3, WAP-191-WML-20000219-a (available at http://www.wapforum.org/), (Feb. 2000)
7) WAP Forum™: XHTML™ Mobile Profile, WAP-277-XHTMLMP-20011029-a.pdf, (available at http://www.wapforum.org/), (Nov. 2001).
8) Yamakami, T.: Unique Identifier Tracking Analysis: A Methodology To Capture Wireless Internet User Behaviors IEEE ICOIN-15, Beppu, Japan, (Feb 2001)
9) Yamakami, T.: An AIMS model-based approach: Toward Understanding Wireless Minds IPSJ DiCoMo2002, Toi, Japan, (July 2002)