Measuring Mobile Data Service Usage and Traffic in Mobile Multi-Access Networks

ANTERO KIVI
Helsinki University of Technology, Networking Laboratory
P.O. Box 3000 FI-02015, FINLAND
antero.kivi(a)tkk.fi

Abstract

New wireless access networks, managed by multiple operators, are emerging as alternatives to mobile cellular networks. Thus, having a holistic view on usage and traffic patterns will be increasingly difficult. A range of methods exist for collecting data on mobile data usage and traffic. Surveys and terminal measurements provide very detailed sample based data, whereas mobile network data is less granular but is based on an operator's entire subscriber base. Server measurements give detailed data on usage patterns of a fairly focused user population. In the future of heterogeneous multi-access networks, multi-radio terminals seem the most promising place to measure usage and traffic, while mobile cellular networks are likely to lose most as traffic partly leaks to other networks.

1 Introduction

Accessing the Internet using advanced mobile handsets and mobile usage of laptop computers is increasing rapidly. Reliable and transparent information on true mobile service usage and traffic patterns is of value to many stakeholders, including business development and marketing, product development, network planning and management, as well as academics studying topics such as consumer behavior and usability.

New wireless access technologies, such as WiMAX and Flash-OFDM, are emerging as alternatives to existing mobile cellular network technologies, while the coverage of WLAN hotspots of various providers is ever increasing. The emergence of multi-radio terminals featuring radio interfaces in addition to basic mobile cellular radio is supporting this trend. Meanwhile, the performance of mobile cellular networks is increasing with the HSPA technologies, keeping mobile cellular as a viable alternative for wireless Internet access. In the future, good interoperability between different access technologies will be essential, and terminals should be able to select the best available access network automatically without user input.

In this heterogeneous environment where both licensed and unlicensed radio spectrum is used and the network has no centralized point, understanding the system as a whole becomes hard. Moreover, as traffic will diverge to multiple networks managed by multiple access providers, having a holistic view on usage and traffic patterns will be increasingly difficult.

In this paper, the alternative sources for collecting information on mobile data usage and traffic are presented and compared, after which the effects of the trend towards the use of multiple wireless access network technologies on these methods are considered.

2 Sources of Data on Mobile Data Service Usage

Usage data can be obtained from many alternative sources (see Figure 1). The most straightforward way available to all researchers is to do a survey or a panel study on a sample of real end-users, while actual usage measurements can be conducted directly at (a sample of) terminals, at the mobile network, and at various servers. Moreover, the charging and billing functions in networks and servers can also provide data on service usage.

Some secondary sources can also be used to acquire data on usage. Expert interviews typically provide information either derived from other data sources or just based on the expert's educated guesses, although such interview studies do enable a large scope as specialists and managers from different organizations and functions can be interviewed. Mobile operators and handset manufacturers publish data on usage and sales in their quarterly or annual reports. Background data, albeit with somewhat narrow scope, is published by retail and wholesale organization such as the Finnish Association of Electronics Wholesalers [5] and Kotek [12], and organizations formed due to regulatory requirements, as is the case with number portability and the Finnish NUMPAC organization [16].

2.1 Surveys and Panels

Surveys are the most widely used method of data collection for studying mobile data service usage. Surveys can be implemented using e.g. telephone, postal mail, email, web, and various face-to-face questioning methods, depending on the available resources and the objectives of the research. Thus, the used method also defines the scalability and accuracy of the survey. Time series data can be produced by repeating a certain set of questions. Surveys are an efficient way of collecting

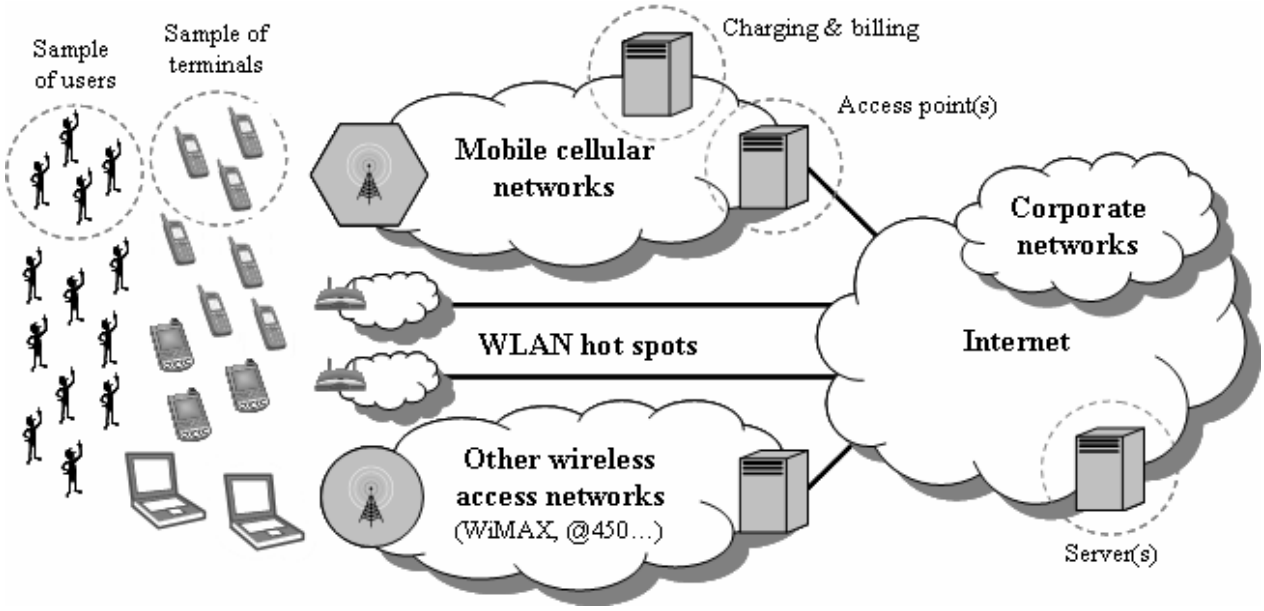

Figure 1 Sources of data on mobile data service usage

information from a large number of respondents, with statistical techniques to determine issues such as statistical significance of the results. Surveys are also flexible as a wide range of information can be collected to study e.g. attitudes, values, beliefs, and past behavior. However, survey responses always depend on the respondents' motivation, honesty, memory, and ability to respond. While a random sample of subjects is often selected for the survey, the actual respondents are usually self-selected, meaning that the characteristics of the whole population cannot be obtained from the sample. Finally, survey question answer-choices could lead to error, as they are often relative only to a personal abstract notion, for both the respondent and the interpreter of the results.

A continuous panel study is a series of measurements on the same sample of test units over an extended period of time. Panel research is used extensively both in Europe and the U.S. to study consumer purchase patterns [1], and it has been used to study mobile service usage as well [10]. Participating panelists use the diary method and register usage events manually to an online or paper "diary", which results in data of high accuracy and granularity. The basic limitations of the panel method are similar to those of surveys, although the continuous nature of panels significantly limits sample sizes.

2.2 Terminal Measurements

Panel studies can also be conducted at the terminal level, where the manual registering of usage events is replaced with the logging functionality of monitoring software or hardware installed in the terminal. Recruiting a representative panel of people is one of the main challenges also in terminal level measurements. Explanatory background variables on the terminal user(s) are also often collected at beginning of the panel, with the reliability issues typical to all survey data. In case of terminals with multiple users (e.g. PC, TV), measures need to be taken to distinguish the usage of each individual.

Terminal level measurements are quite common in the PC world. Commercial analysts, such as Nielsen NetRatings, have panels with hundreds of thousands of monitored participants. The software watches what a user does with the computer and sends that information further. While being otherwise quite similar to benign PC monitoring software, malicious spyware monitors the PC without the consent of the user. Spyware can collect information ranging from tracking the types of visited websites to recording the user's keystrokes to intercept passwords or credit card numbers.

Television viewing can also be measured at the terminal level. In Finland television viewing research is conducted with a continuous measurement of in-home viewing of a panel of 1000 households. The TV sets in panel households are equipped with metering devices to monitor changes in set status and viewer appearance. [6]

The monitoring of advanced mobile handsets has become possible due to recent developments in handset operating systems and processing capability. A handset monitoring software can measure the usage frequencies, durations and volumes of all terminal features and applications. Communication-specific data (e.g. voice calls, SMS, MMS) can be broken down between different callers/senders and recipients, while visited browsing destinations can also be logged. Usage of the handset's different radio interfaces (e.g. GSM/WCDMA, WLAN) for packet data transmission by different applications can be differentiated, and the possible *ad hoc* connections (e.g. Bluetooth, WLAN) with other terminals can also be measured. Usage of the handset's offline features (e.g. camera, multimedia player, games) and broadcast multimedia (e.g. FM radio, DVB-H based TV) is also captured. Moreover, location information (cell identity code) connecting user location and mobility to usage might also be collected, though the actual

geographic location of the cell is not known by the terminal. The monitoring software can be further augmented with triggered "real time" pop up questions sent to the panelist after certain pre-specified event (e.g. at 12 o'clock, browsing session ended). While handset monitoring is not yet very common, at least two such software for the Symbian S60 platform with slightly differing functionality have previously been used by the academic community ([11] and [20], [14] and [3]). A major drawback of these studies is that their scope is limited to the users of a certain, albeit widely used, handset operating system (OS) and software platform. As the OS and software platform might also have a large effect on usage behavior, and the results are thus not generalizable to users of significantly differing mobile handsets. Another limitation of such monitoring software is that all non-standardized applications are seen as "black boxes" with no information on the usage "inside" the application. Commercial spyware specifically made for mobile handsets with some of the above functionalities has also emerged recently [7].

Mobile handset and PC monitoring methods have been combined in some rare cases to study the general behavioral patterns related to computer and handset usage of certain focused user populations [8].

2.3 Mobile Cellular Network Measurements

Mobile cellular network is a logical measurement point, as it is a point of convergence of mobile data traffic and covers all subscribers of the mobile operator.

Mobile operator's charging and billing systems provide a great source of service usage data, as information on all chargeable events generated by the subscriber is registered in them. In principle, the time-stamped GSM/UMTS charging data records (CDRs) identify the mobile subscriber (by IMSI code), the used mobile terminal (by IMEI code), and the volume of packet data traffic to/from different external packet networks (by the used GGSN). The billing system uses aggregated CDR data, and also contains other data on subscribers from the customer register. Subscriptions vary in type (e.g. postpaid / prepaid, consumer / business, fixed-term / continuing) and in tariffs of different services (e.g. voice calls, SMSs, packet data transfer), for instance. Consumer subscriptions are registered to an identity number or social security number and business subscriptions are identified by a business ID, both of which only refer to the bill payer of the subscription and not to the actual end-user. Whether or not the CDR and subscriber data can be combined in detail depends very much on the implementation of the reporting functionalities of the operator's information systems. Unrestricted access to the CDR databases or customer registers could also enable the use of sophisticated data mining techniques to uncover usage patterns (e.g. [21]).

Mobile network packet data traffic measurements are most easily conducted at the access points to external packet data networks (GGSN Gi interface), such as the Internet, capturing the traffic of all subscribers going via the access point. Capturing TCP/IP packet headers only and thus avoiding sensitive application level user data keeps the amount of measurement data manageable, and enables the analysis of generic traffic patterns. The volume of usage (bytes, flows) can be broken down by application protocol, destination host IP addresses (e.g. for web servers), and by day and time of usage. While individual usage sessions can be separated, they cannot be accounted to individuals as subscriber terminal IP addresses are typically allocated dynamically. Usage of mobile handsets can be distinguished from laptop usage. While browser based identification does necessitate application level header data, the distinct TCP fingerprints of different operating systems enable identification with just the IP and TCP headers. Similar measurements are often used in more technologically-oriented research while studying e.g. network or protocol performance in mobile or fixed networks.

In GSM/UMTS networks, packet data traffic header data could in fact be linked to the charging and billing data, at least in theory. This would necessitate keeping a continuous log of the IP addresses allocated to each subscriber terminal at the SGSN. Thus, a terminal identifier (TAC code) and/or subscriber identifier (IMSI) could be linked to traffic by the used IP address and time of usage. The subscriber identifier in turn could be linked to the data in the customer register. In principle, the network is also aware of the cell identity the mobile is using at all times. Combining location information at to traffic data at this level, however, would be a yet more complicated process. These rather laborious efforts have been done to some extent in [9].

The presented methods apply to GSM/UMTS mobile cellular networks. The extent to which similar measurements can be repeated in other mobile cellular networks or e.g. in WiMAX networks is not clear.

2.4 Server Measurements

Service usage and traffic patterns by mobile terminals can be studied at the server level at various points, including portals and individual web/wap sites/servers, search engines, proxy serves, as well as by with a service provider's billing data. The scope of server level measurements is naturally limited to the users of the service in question. Background data on the registered users of a service might also be available.

Web portals are a place where usage and traffic converge. Usage of web portals and individual sites can be monitored in a similar manner. A typical method includes placing small pieces of code on all pages of a web site. Each time a page is loaded by a user, the code executes and sends data to another server. New and repeated visitors can be distinguished using browser

cookies. At best, the method enables the identification of individual users and their detailed usage patterns. By identifying the used browser version, the use of a mobile device can be identified as well. Another way of obtaining web site usage data is to use server software (e.g. Apache web server) log files. Log file data is typically less detailed, but can cover several sites, as web hotel and hosting services often locate multiple web sites on the same server machine. [18], [2]

Analyst companies (e.g. TNS, Nielsen/NetRatings) providing web site usage analysis services can also combine the data measured on multiple individual web sites. Such data is typically published, if published at all, covering the service providers in one country regardless of the origin of the actual users (see [17]). No mobile-specific data has yet been published, though such data should already be available. The representativeness of these studies is somewhat questionable, as they only cover the clients of the analyst company who also allow the publication of the data of their own site. This means that all web sites are (by far) not included in the data, as analysis services are typically purchased only by well-established service providers. Thus, providers of services such as advertising and adult content, for instance, are typically not covered by the method. Other analytics software (e.g. Google Analytics) could possibly be used similarly to aggregate browsing data.

Search engine companies have another source of usage data at their disposal at the server level, as analysis on the most used search words provides information on service popularity. Mobile usage can be identified from accesses to mobile-adapted search sites as well as from the use of services specifically made for searching the mobile web (e.g. Google Mobile). Moreover, a mobile device accessing standard PC search site can also be identified, again by its browser type, while the searches of individual users can be distinguished with the help of browser cookies. As search companies also offer a range of other Internet services they are able to relate the background data on registered users to search behavior, regardless of which terminal is used as long as browser cookies identify the individual in question. The potential of such data was demonstrated in autumn 2006, when AOL released about 2.2 Gb of search logs to the general public with the seemingly good intention of providing the research community with hard search engine usage data. While the identities of individual people were not revealed by AOL, many of them could be easily deduced from the used search words. [19] In another less controversial case, search engine data has been specifically used to analyze mobile search patterns [3].

Traffic of multiple users converges also at proxy servers. Caching web proxies, for instance, can be used to measure web site popularity. The Opera Mini java browser for mobile devices serves as a specifically mobile related example of proxy based usage measurement. As Opera Mini fetches all requested content through an Opera proxy, detailed statistics on the browsing behavior of all Opera Mini users are available to Opera Software (see e.g. [13]).

Methods comparable to those used in mobile network specific measurements can also be used at server level. The service provider's billing data gives indications on service usage related to monetary transactions, while traffic measurements with similar methods to identify mobile device usage can also be applied at servers.

2.5 Comparison of Alternative Data Sources

The alternative data collection methods have fundamental differences regarding the typical researcher, research scope, and the nature and granularity of collected data. A summary of the characteristics of each method is presented in table 1.

The selection of a usage data collection method depends very much on its availability, i.e. on the access of the researcher to the data. Surveys are far the most used method as they are available to anybody. Implementing a terminal monitoring panel is harder, as a functioning monitoring client for mobile devices is needed in addition to the capability of recruiting a representative panel. Only mobile operators are able to measure at the mobile network, as direct access to operator reporting systems and the network itself are both required. Server measurements can only be conducted by the service provider in question, or someone on their behalf.

The selected research method dictates the scope and size of the studied sample. While survey sample sizes vary a lot depending on the used survey method (doorstep vs. web surveys), panel studies typically have smaller sample sizes. Mobile network measurements cover a large number of people, as the studied sample can be up to the entire subscriber base of the operator. Server level measurements can have varying sample sizes, at maximum encompassing all users of the measured service.

The main difference between surveys and different measurements is that surveys provide subjective data on usage as perceived by the respondents, whereas measurements generally provide objective data. The accuracy and granularity of the data depend on the method. Survey respondents are likely able to provide information on aggregate usage of different services, whereas diary based panel and actual measurements are able to register even individual usage events and transactions.

The ability to obtain variables potentially explaining usage varies a lot by method. Survey studies can again include any background variable (e.g. gender, handset type, pricing scheme) the respondent is capable of providing. Generic data on the perceived time (e.g. evening) and context (e.g. at home) of usage can also be obtained, while diary based methods might achieve

Table 1 Summary of alternative mobile data usage data collection methods

Methods / Attributes	Surveys and panels	Terminal measurements	Mobile cellular network measurements		Servers measurements
			Charging & billing	Packet data traffic	
Researchers with access to data	Anybody (analysts, consulting firms. investment banks, academics…)	Those with a monitoring client and ability to recruit the sample	Mobile operators	Mobile operators	Service providers (mobile operators, 3rd party providers, search engines…)
Research scope and sample size	Survey respondents: $10^1 – 10^5$ Panel participants: $10^1 – 10^3$	Sample of panelists using certain handset OS and software platform: $10^1 – 10^3$	Operator's entire subscriber / terminal base: $10^5 – 10^7$	Operator's entire subscriber / terminal base: $10^5 – 10^7$	All users of the measured service(s): $10^2 – 10^7$
Nature and description of data	Subjective data on perceived aggregate service usage, diary method could sort out individual events	Accurate and objective data on handset application and feature usage	Accurate and objective data on mobile network transactions	Quite accurate and objective profile of mobile uplink / downlink packet data traffic	Objective data on service-specific usage, accuracy depending on the method
Explanatory / independent variables • User • Terminal device • Time • Location	Any background variables on the respondents Perceived time and context of usage Panels using a diary method enable more accuracy in time and location of usage	Many background variables on the panelists Handset model and access network used Time of usage Location of usage (cell ID)	Some background variables on subscription (type, tariff), no data on real end-users Terminal model Time of usage No location data	No data on individual subscribers or terminals, different terminal operating systems identifiable Time of usage No location data	All depends on the used method (identification of individual users, background data on registered users, separation of mobile usage, and time of usage possible) No location data
Usage / dependent variables • Usage volume • Usage frequency • Data granularity	Perceived amount and frequency of service usage	Volume, frequency, and duration of usage per panelist	Volume and frequency of chargeable service usage per terminal or subscriber type	Volume of usage (bytes, flows) per application protocol and traffic destination	Depends on the used method
			Additional measurements could link traffic data to charging and billing data		

higher level of accuracy in this respect. Terminal measurements typically begin with a survey study collecting relevant background data on the panelists. The monitoring software itself is aware of the handset model and features, the used access network, as well as time of each usage event. The location of usage might also be collected (cell ID).

The mobile network charging and billing based data can associate subscription information (type of subscription, tariff scheme) and terminal type to usage, though no data on the actual end-users is typically available. A timestamp of each transaction is also registered, as service pricing might depend on it. Sophisticated data mining techniques might include other variables. In mobile network packet data traffic measurement individual subscribers or terminals cannot be separated from each other. However, additional analysis methods enable the separation of different terminal operating systems, i.e. the separation of mobile handset usage from laptop originated traffic. The exact time of usage is also registered in packet data traffic measurements. While the mobile network is aware of the cell covering terminals with active packet data connections and the operator could actually relate the cell identities to geographical locations, actually combining this data to usage would be very laborious. For server level measurements, it again depends on the type of method used. Individual users might be separated and identified, and previous registration could also provide explanatory background variables for each user. Mobile handset originated usage can also be distinguished from PC usage in some cases, as should if the method is used for measuring mobile data usage. The time of service usage is also often obtainable. The user location cannot be known at the server, though the operator and country of the user could be derived from the user terminal IP address.

The usage volume and frequency of different services can be obtained at different levels of granularity. Surveys and panels provide data on the amount and frequency of service usage, as perceived by the

respondent. Terminal measurements register the volume, frequency and duration of usage per application for each panelist. Charging and billing data can produce summary data on usage of different chargeable services, typically aggregated to subscriber or terminal type level. Packet data traffic measurements provide byte and flow counts on the volume of usage, accountable to used application protocols and external network traffic destinations (Internet hosts/servers). As stated before, by conducting additional measurements in the mobile network, traffic data and charging and billing data could be linked. The type of data obtainable on server level usage depends completely on the used measurement method.

The major limitation of any type of usage measurement is the fact that one can only measure what a user has done, not why the user has done something or what he/she actually intended to do. Thus, determining the motivation of usage or the real end-user need is hard based on measurement data only. Surveys are more flexible in this, as data on both perceived usage and motivation of usage can be collected, provided that the survey questions are formulated accordingly to obtain valid data. Terminal measurements facilitating the pop up question mechanism provide a promising new method to uncover the true reasons of usage, as the panelist can be asked triggered questions immediately following a certain type of usage. However, use of the pop up method must be very focused, as only a limited number of questions and only related to unsensitive type of usage can be asked to avoid irritating the panel participants. While mobile network measurements can't basically obtain any data on the motivation of usage, server level measurements have some means to understand it for instance by analyzing the user's site-specific browsing or search patterns.

As each method has its advantages and disadvantages, multiple methods are often used to gain further knowledge on usage. For instance, mobile operators can complement charging data based subscriber segmentation by conducting a survey study to better understand the motivation and usage behavior of each segment. This is also recommendable from the theoretical point of view, as method triangulation can improve both internal and external validity of the data [15].

3 Effect of Mobile Multi-Access to Usage and Traffic Measurements

The use of multiple alternative wireless access methods will diverge traffic to multiple networks. While almost all mobile handset users and many laptop users are currently (2006) still using mobile cellular networks for accessing the Internet, traffic will increasingly leak to alternative access networks in the near future. This will have a profound effect on the methods to measure mobile data usage and traffic.

Survey and panel studies should work quite similarly in collecting data on usage behavior, as perceived by the users. However, if terminals are capable of changing access networks automatically, the users might not be aware of such changes and, thus, of the network they are using at any given time. Relating usage behavior to the performance or other characteristics of a particular network will therefore be hard.

As the same terminal is still used irrespective of the access network used, the mobile terminal is still the point where all usage of an individual user converges. Moreover, if the terminal monitoring software is aware of the changes in network usage, behavior can also be associated with the used network, while the motivation of usage and user experience can also be measured by triggering pop up questions based on the network in use. Thus, the terminal is still a promising place for conducting usage and traffic measurements.

As mobile cellular networks will only represent one of the many alternative wireless access networks, usage and traffic will no longer be holistically measurable in them. While the spreading WLAN hotspots might rapidly take a large share of the total mobile data traffic volume, new HSPA techniques will still be in the scope of the presented mobile cellular network measurements.

Most server level measurements should function in a similar manner regardless of the used access technology. As mobile handsets are converging with PCs in functionality and will be capable of using all the services previously used exclusively by PCs instead of the mobile-specific/adapted versions, distinguishing the two might become more difficult. This also applies to some handset applications (e.g. Opera Mini browser) whose architecture has enabled server level measurements. Some of the methods in identifying mobile clients are still applicable, but the used access method might also be of less interest to service providers in the future when adapting the service to the lesser capabilities of the mobile terminals is no longer necessary.

4 Conclusions

A range of alternative methods for collecting data on mobile data usage and traffic have been presented in this paper. Each of the methods has its advantages and disadvantages, and the applicability of a particular method depends on the research objectives as none of the presented methods is suitable for all purposes. Thus, multiple methods are often used. In general, surveys and terminal measurements provide very detailed sample based data, whereas mobile network data is less granular but is based on the entire subscriber population of an operator. Server level measurements are a compromise between the above, as detailed usage patterns can be uncovered from a fairly focused user population.

In the future of heterogeneous multi-access networks, terminals seems to be the most promising place to

measure usage and traffic, while mobile cellular networks are likely to lose most as traffic will partly diverge to other wireless access networks. In addition, the mechanism of pop up questioning in terminal measurements presents a potential research approach.

The usage of money might also provide a centralized view for understanding mobile service usage. While accesses to bank, auction and e-commerce web sites, for instance, already give some indications, authentication in the electronic and mobile environment might be another convergence point of money usage. Although universal authentication is still a thing of the future, authentication is already performed by various actors, including banks, credit card companies, mobile operators, as well as purely Internet based actors such as PayPal. Complementing the time stamp and "location stamp" with a "price stamp" presents an interesting future research prospect for understanding mobile user behavior.

References

[1] Aaker David A., Kumar V., Day George S. (2004). Marketing Research, 8th ed., John Wiley & Sons Inc.

[2] Bluemetrix Inc web site [online]. Bluemetrix – Business intelligence at your fingertips, updated 2006 [accessed September 14, 2006]. Available in www-format: <URL: http://www.bluemetrix.com/engsite/webanalysis/methodology.htm>

[3] Kamvar M., Baluja S. (2006). A Large Scale Study of Wireless Search Behavior: Google Mobile Search, presented at CHI 2006, April 22-27, 2006, Montréal, Québec, Canada.

[4] Eagle N. (2005). Machine Perception and Learning of Complex Social Systems. Doctoral dissertation, Program in Media Arts and Sciences, Massachusetts Institute of Technology, USA.

[5] Elektroniikan Tukkukauppiaat ry web site [online]. Elektroniikan Tukkukauppiaat RY, updated 2006 [accessed September 8, 2006]. Available in www-format: <URL: http://www.etkry.com/>

[6] Finnpanel Oy web site [online]. Finnpanel Oy, updated 2006 [accessed September 14, 2006] Available in www-format: <URL: http://www.finnpanel.fi/>

[7] Vervata Co Ltd web site [online]. FlexiSpy – Spy Software for mobile / cell phones, updated 2006 [accessed September 15, 2006]. Available in www-format: <URL: http://www.flexispy.com/>

[8] Heimo Consulting web site [online]. Heimo Tribe, updated January 2005 [accessed October 5, 2005]. Available in www-format: <URL: http://www.heimotribe.com>

[9] Kalden R. (2004). Mobile internet traffic measurement and modeling based on data from commercial GPRS networks. Doctoral dissertation, University of Twente, The Netherlands.

[10] Kim, J. (2001). Analyzing mobile Internet users: Results from a monitoring study, an experiment, and a survey study. In CHI 2001 Workshop, Mobile Applications: Understanding Users, Adoption & Design.

[11] Kivi A. (2006). Mobile Internet Usage Measurements - Case Finland. Master's Thesis Series. Networking Laboratory. Department of Electrical and Telecommunications Engineering. Helsinki University of Technology, Espoo, Finland, April 2006.

[12] Kotek web site [online]. KOTEK – Kodintekniikka-alan tiedotusfoorumi, updated 2006 [accessed September 8, 2006]. Available in www-format: <URL: http://www.kotek.fi>

[13] Opera Software ASA press release August 8, 2006. Mobile Browsing Bonanza: One billion Web pages viewed with Opera Mini.

[14] Raento, M., Oulasvirta, A., Petit, R., Toivonen, H. ContextPhone – A Prototyping Platform for Context-Aware Mobile Applications. IEEE Pervasive Computing, Vol. 4, No. 2 (April-June 2005), pp. 51-59.

[15] Scandura T., Willians E. (2000). Research Methodology in Management: Current Practices, Trends, and Implications for Future Research. Academy of Management Journal, Vol. 43, No. 6 (Dec. 2000), pp. 1248-1264.

[16] Suomen numerot NUMPAC Oy web site [online]. Suomen numerot NUMPAC Oy, updated 2006 [accessed September 10, 2006]. Available in www-format: <URL: http://www.numpac.fi/>

[17] TNS Gallup web site [online]. TNS Metrix, updated 2006 [accessed April 4, 2006]. Available in www-format: <URL: http://www.gallupweb.com/tnsmetrix.htm>

[18] Tähtinen M. (2006). Tunne web-käyttäjäsi, Tietokone, No. 9, 2006, pp. 55-57.

[19] USATODAY.com [online]. AOL search data release reveals a great deal, updated August 6, 2006 [accessed September 10, 2006]. Available in www-format: <URL: http://usatoday.com/tech/columnist/andrewkantor/2006-08-17-aol-data_x.htm>

[20] Verkasalo, H & Hämmäinen, H. (2006). Handset-Based Monitoring of Mobile Subscribers. Presented at Mobility Roundtable, 1-2 June, 2006, Helsinki, Finland.

[21] Wei, C-P & Chiu I-T (2002). Turning telecommunications call details to churn prediction: a data mining approach, Expert Systems with Applications, Vol. 23 (2002), pp. 103-112.