

MSCE: A Mobile Statistical Computing Environment

YUE-SHAN CHANG, YI-SHIN KUO

Department of Computer Science and Information Engineering

National Taipei University

151, University Road, Sanhsia, Taipei County,

237 TAIWAN, R.O.C.

<http://web.ntpu.edu.tw/~yssc/>

Abstract: - In the paper, we design and implement a simple architecture to supporting statistical computing for novice or professional statisticians executing computing and analyzing using wired/wireless device, such as Notebook, PDA, PDA phone, or Java enabled cell-phone, that with embedded browser or e-mail and without any statistical software or tools. Users can submit their statistical program (SAS in this prototype) via various familiarized user interfaces to backend server and get the response immediately. It is believable that the environment is useful for users who need statistical tools. The environment can tie other tools for supporting various specialties E-/M- science.

Keywords: Statistical Computing, Mobile computing, mobile statistical, SAS, J2ME

1 Introduction

Statistical tools are important tools for social science to analyze and examine collected data to mine the information behind the data. The information can immediately affect the decision-making lately. There are some famous and expensive statistical tools, such as SAS [1], SPSS [3], and MINITAB [2], on the world. The expensive and bulky software are not suitable to be installed on the mobile device that have limited storage and computing power.

It has the trend of E-/M-science for all specialties to cope with their domain issues on the wired/wireless Internet anytime and anywhere [9-12]. The M- science is necessary based on the emerging technology-mobile computing. Mobile computing [13] systems are constrained in important ways relative to static systems. These constraints are intrinsic to mobility as follows:

1. Mobile devices are resource-poor relative to static devices
2. Mobile devices are more prone to loss, destruction, and theft than static devices.
3. Mobile devices must operate under a much broader range of networking conditions.

Imaginatively, biologist makes study in the field. He/She is inconveniently taking heavy equipments. Statistical instructor needs to compute and analyze data that is collected from class where without any statistical tools. Or market investigator investigates in the street. These users maybe not have any desktop computer and statistical tool in hand. There are some network-based statistical environments for supporting statisticians to execute statistical computing on the Internet [4-7]. There are surveyed in next section. Their do not support computing using mobile devices.

Obviously, supporting a mobile statistical computing environment is necessary for various specialties to cope with their domain problems. In the paper, we design and implement a novel application for E-/M-science, that is a simple architecture and environment, named MSCE (Mobile Statistical Computing Environment) to support mobile statistical computing for novice or professional statisticians executing computing and analyzing using wired/wireless device, such as Notebook, PDA, PDA phone, or Java enabled cell-phone, that are with embedded browser or e-mail and without any statistical software or tools. Users can submit their statistical program (SAS in this prototype) via various familiarized user interfaces, such as web browser, e-mail client, or designated java based agent, to backend server and get the response immediately no matter what kind of computing devices used. The technology used in the system is Java related technologies such as Servlet, JSP, J2ME, and JavaMail, so that the system can be installed and executed anywhere.

In addition, we evaluate the effectiveness and efficiency of the system by various samples. It is believable that the environment is useful for users who need statistical tools. The environment can tie other tools for supporting various specialties E-/M- science.

This paper is organized as follows. Section 2 we survey some works related to the networked based statistical systems. Section 3 examines some design and implementation issues. Section 4 demonstrates the operation using the system to submit SAS procedure and get the response in three ways. Section 5 gives a conclusion.

2 Related Works

There are some networked statistical environments designed for statisticians as follows:

Junji Nakano [7] implemented a JASP (Java based Statistical Processor) - a Statistical package. JASP was based on the client-server model, and used Java RMI (Remote Method Invocation) and distributed computing technology to implement the statistical package. Its client UI (User Interface) applet can use only the server on the Web server computer. The server to which the client connects firstly is called as a main server, and let a main server communicate with other remote servers.

StatCrunch [4] is statistical software for data analysis on the Web developed by Dr. West. The main aim of the menu driven package was to offer a freely accessible alternative to commercial statistical software via the WWW with the only minimal requirement on the part of the user being a Java capable Web browser. The current edition is StatCrunch 4.0.

MMM (Method Management system) MMM [5] is proposed by German National Research Center. The Center promotes interdisciplinary projects among its members, who include mathematicians, statisticians, econometricians, economists, and computer scientists. The key features of the MMM system for sharing statistical computing modules are 1) its implementation of stateful services to support the exploratory nature of statistical data analysis, 2) its use of middleware services to enable interoperability between proprietary statistical computing packages, and 3) its provision of publishing support to help authors with their interface definition.

Chooichiro Asano et. al. develop a web-based statistical system [6]. The basic idea is to separate the interface from statistical engine and to apply the web browser as the interface and Dynamic Link Library (DLL) as statistical library. The application of a web browser has great advantages as an interface, e.g. ability of addition of documents, graphical objects, hyperlinks to other relative sites on the web, and so on. These capabilities make it easy to develop an interface to fit users' wide needs by using the system. Combining a web-based interface page with an appropriate site by hyperlinks, the users can access the site quickly and analyze data while reading related documents.

We also proposed a flexible and scalable statistical environment called MASS (stand for Multi/web-Agent Statistical Service) [8] based on the multi/web-agent and offered a unified XML-based programming interface for statistical computing. We separated various statistical functionalities into groups of services and design its interface and architecture. All of services were defined as an agent and

implemented using web service technology. And all the statistical operations, data, and results were formulated as a XML document.

All the systems mentioned above only supported web-based statistical service. It did not yet support other user interface, such as E-mail, and mobile devices that with small screen. The system we support not only the web-based environment but also mobile statistical service via e-mail or java based user agent. User can submit movably their SAS program.

3 System Design and Implementation

3.1 System Architecture

In general, the transmission on the Internet can be classify into two modes; one is synchronous and the other is asynchronous. The applications, such as Web, MSN, and general client/server based applications, are synchronous mode applications because two sides of application should be on-line simultaneously. On the contrary, the asynchronous mode, the receiver of message/request can be off-line temporarily. The e-mail is a well-known asynchronous mode application. The synchronous mode application can immediately get the response while the asynchronous mode needs to wait for response in a time period. As we known, statistical computing sometimes is a time-consuming for large data set or complex procedure. Such computing we prefer to initial asynchronous mode computing so that we can save time for getting the response. Another advantage of asynchronous mode is location-independent. We can submit request on one location and get response on another location.

As mentioned above, the environment can support users submitting their SAS procedure using versatile devices. We consider the possible way user interface to submit the statistical computing for mobile users. The system will offer three types user interface to connect to, as follows:

- (1) Web browser: Notebook or embedded on the mobile device.
- (2) The J2ME cell-phone with PDA function: connect to the backend server utilizing wireless communication, such as WiMax, WiFi, or GPRS (3G).
- (3) E-mail: any devices (involving handheld device) that have e-mail application can send the e-mail to backend server.

Fig. 1 shows the system architecture of MSCE. It is a three-tier architecture. First tier is a user interface involving browser, e-mail client, or designated java based agent that can submit the SAS file to and get response from backend server. Second tier is the service tier to accept and forward request from client to server, then to get and deliver response to client. In this tier, we adopt Java related technologies such as

Servlet, JSP, J2ME, and JavaMail, so that the system can be installed and executed anywhere. Third tier is backend server serving the request.

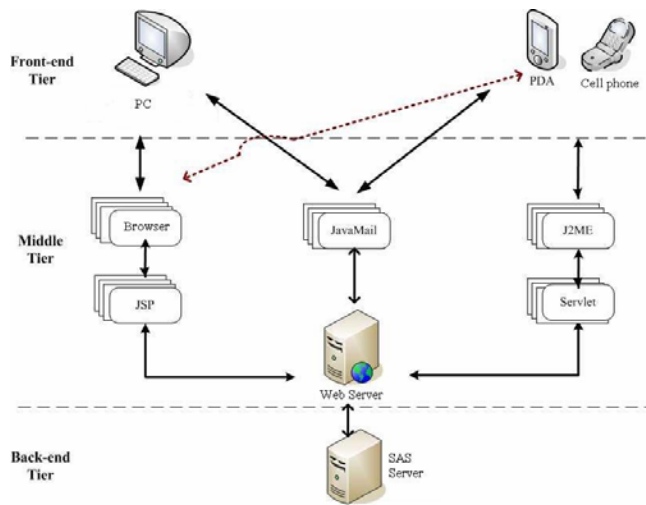


Fig. 1 System Architecture

3.2 Implementation issues

As mentioned above, the implementation of MSCE adopts the java related technologies in order to cross platform easily. The web page is constructed using JSP. The user agent of cell-phone or PDA is built by the J2ME and MIDP that are with *FileConnection* API to access data file stored on these devices. The motivation is that almost all cell phones are J2ME enabled. In the subsection we depict the implementation of the system separated three manner: web page, J2ME enabled mobile device and E-Mail respectively.

3.2.1 Web page

Fig. 2 shows the operation flows based on JSP. User using browser to submit SAS program can be uploading either a SAS file or a text that edited on line. After uploading, server will save the request as a SAS file, and then invoke SAS program to execute the SAS file. Finally SAS will save the result into a file that user can browse it.

The file naming scheme applied in MSCE is well known time stamp approach. That is the file name is combined from the received time, that is “dd:hh:mm:ss:ms.SAS” scheme. Here hh is hour, mm is minute, ss is second, and ms is mini second, For example, a file is named as 08203025956.SAS that represents the file is received at 20:30:25:956 on date is 8. User can then view the result file or log file immediately.

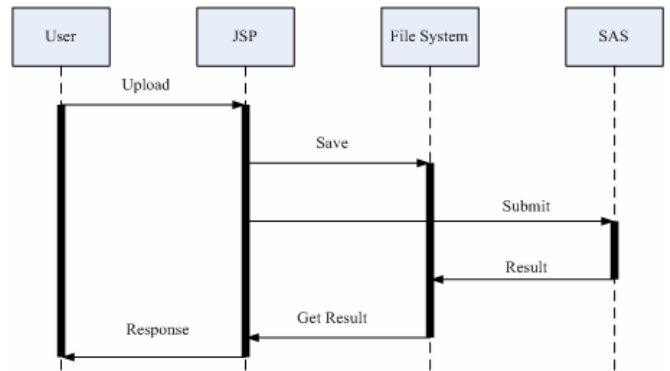


Fig. 2 JSP webpage flow chart

3.2.2 J2ME cell phone and PDA

The subsection will depict the implementation of user agent for Java enabled cell-phone or PDA. The user agent includes three functions as follows: 1) open and view a file, 2) submit a SAS file to remote server, and 3) edit a SAS program and then submit it to remote server. In the J2ME, open a file can use *FileConnection* API to visit directory and browse file on the cell-phone or PDA. User can view SAS file first and then submit it or edit SAS program and submit.

The server side is implemented using Java Servlet. So that submitting a file is using HTTP protocol to convey the file to remote server. First, user agent use *Connector.open()* to open an URL connection to remote server, and get a *HttpConnection* connection entity. Then it will call *setRequestMethod()* to set the URL Request as POST and call *openOutputStream()* to create an Output Stream to remote server. And finally, the agent will call *write()* to output SAS file to stream buffer. On the other side, Servlet create a *HttpConnection* entity and then call *openInputStream()* to create a Input Stream to receive request from user agent. After Servlet receiving a request and saving it as SAS file, it then invoke SAS to execute the SAS file. And finally reply the result to the user agent.

3.2.3 E-Mail

Use of web browser and J2ME user agent to submit SAS file are all synchronous modes. Here we will present the asynchronous mode- E-mail. Without doubt, we need an e-mail account that support POP3 and SMTP protocol to submit the SAS file. So that, we can submit a request using various devices that offer e-mail application.

In the design, MSCE adopts general mail server instead of implementing a new mail server. In addition, we implement a background process named SASmailDeamon that is responsible for get and send mail in the server. The daemon will periodically scan the mail box in a constant time period and check the mail that is SAS request or not. In the prototype we set

the time period is one minute. When a mail in the mail server has a leading word “SAS” in the subject field, the daemon will get the mail and retrieve the content in mail body, then save the content as SAS file. The daemon is a Java application and call JavaMail API to get and reply mail from mail server. After SAS execution, the mail daemon will pack the result of SAS into a mail and use *Transport.send()* to reply result to sender mail account. If there are many mails in the mail box, the daemon will process sequentially all mail.

4 Example

Here we demonstrate the operation using MSCE to submit SAS procedure and get the response in three ways mentioned above. The example we used will analyze the relation between study time (X) and achievement (Y) under $\alpha = 0.05$ of six students. The data set are as Table 1.

Table 1: Data set for operations

Y	60	72	74	78	80	82
X	2	2	3	3	4	4

The model is

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

$$H_0 : \beta_1 = 0 \quad \text{vs.} \quad H_1 : \beta_1 \neq 0$$

$$t^* = 3.16 > t(1 - \alpha / 2, v - 2) = t(0.975, 4)$$

From SAS OUTPUT, we can obtain
 P-Value = $P(t(v-2) > |t^*|) = 0.0343 < \alpha$
 is equal to 0.05

4.1 JSP dynamic web page

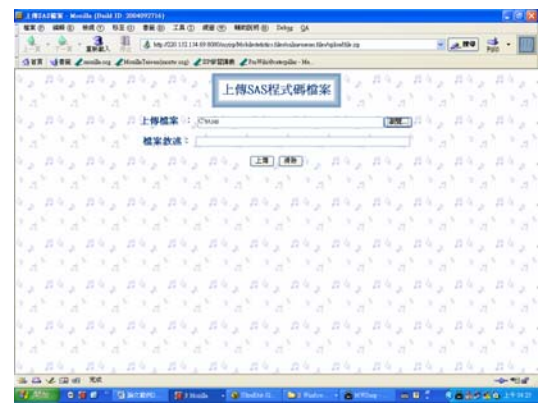
In this subsection we show briefly the operations using browser on the notebook or embedded browser on the mobile device. Fig. 3(a) shows the operation uploading SAS file. Fig. 3(b) shows the response after upload. User can option to view the result file or log file. Fig. 3(d) shows the result file. If user does not have presently SAS file. He can also edit SAS file for submission, as shown in Fig. 3(c).

4.2 J2ME cell phone submission

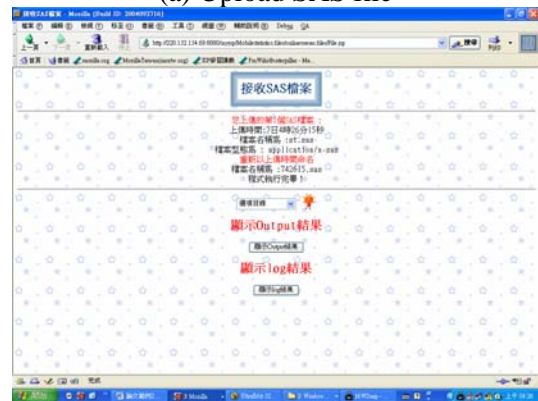
In the emulation of cell phone and PDA, we are using Qwerty Device simulator of J2ME Wireless Toolkit 2.2, which have wide-screen for showing the result completely.

Similarly, Fig. 4 shows the operations of uploading SAS file using the Qwerty Device. First, user select a SAS file and the press the “exeSAS” button to submit the request, as shown in Fig. 4(a). And then get the response from MSCE as shown in

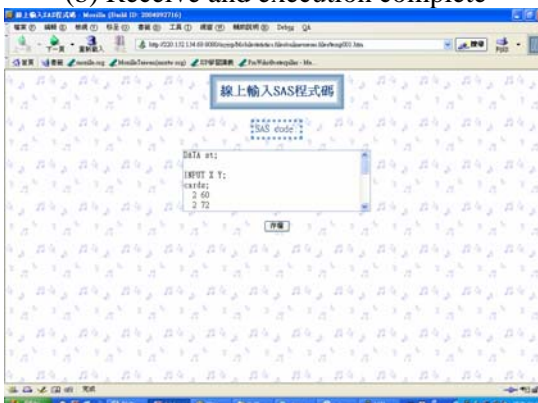
Fig. 4(b).



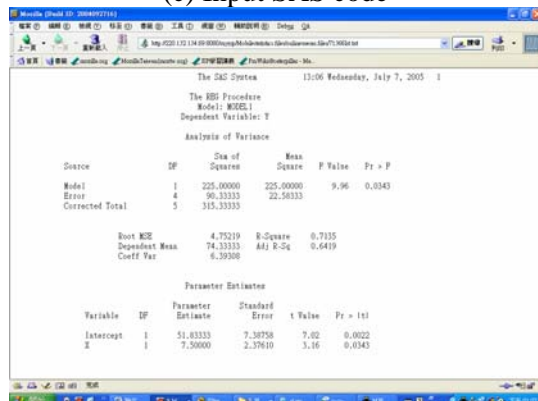
(a) Upload SAS file



(b) Receive and execution complete



(c) Input SAS code

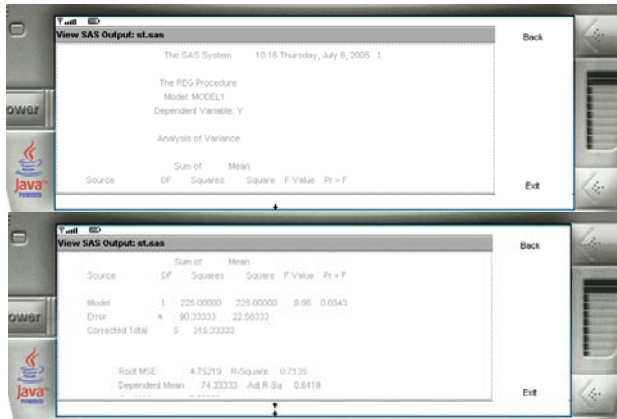


(d) SAS OUTPUT

Fig. 3 browse upload SAS code



(a) File transmission



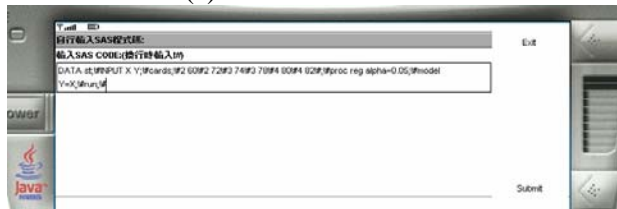
(b) Display result

Fig. 4 J2ME uploading SAS code

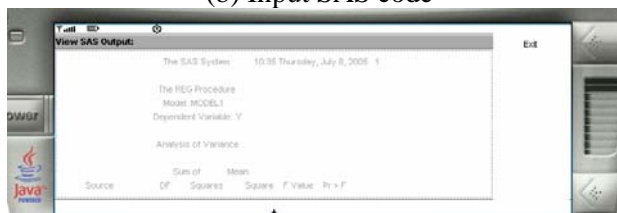
If user does not have ready a SAS file. He can also edit SAS procedure immediately in order to submit. First, as shown in Fig. 5(a), user choice “textFiledCapture” function, and then input SAS code, as shown in Fig. 5(b). Then press “Submit” button to submit the request. Finally, get the response from the backend server.



(a) Function selection



(b) Input SAS code



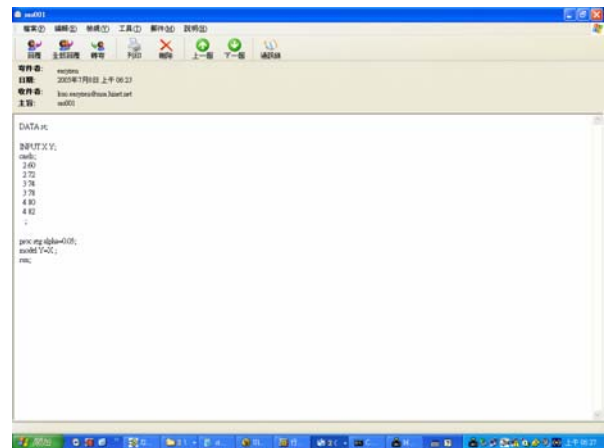
(c) Display result

Fig. 5 J2ME phone input and upload SAS code

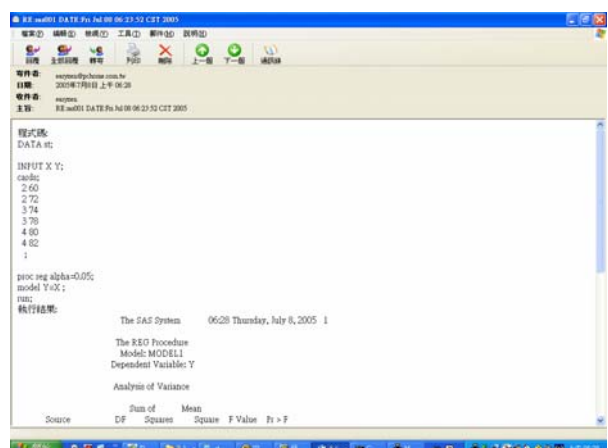
4.3 E-MAIL submission

As we know, e-mail is one of most popular application on the Internet, which can be found in most of Internet device. The application belongs to asynchronous mode application. User can submit their statistical computing at one location and get the response at another location. The submitted SAS code in this manner can be placed at the “Body” field of the e-mail and the “Subject” must have a leading term “SAS”. Fig. 6 shows the operations using OUTLOOK to submit the request. Fig. 6(b) shows the result from the backend server.

Fig. 7 and Fig. 8 shows the operations of using e-mail to send SAS code and receive its result by Motorola E680 cell-phone that with PDA function. That can send e-mail to mail server via GPRS. Due to the limited of screen, we use three pictures to show the content of SAS code and four pictures to show the content of result.



(a) Sending SAS code using OUTLOOK



(b) Receiving response using OUTLOOK

Fig. 6 Sending request & receiving response using OUTLOOK



Fig. 7 Sending SAS code using Motorola E680's e-mail

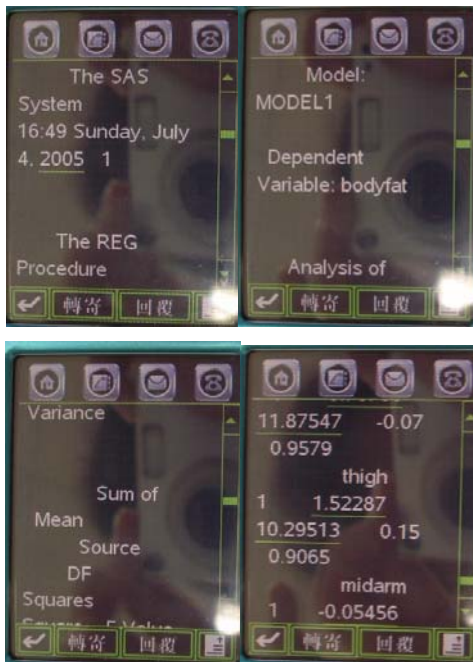


Fig. 8 Receiving SAS OUTPUT using Motorola E680's e-mail

5 Conclusion

In the paper, we have designed and implemented a novel application for E-/M-science, that is a simple architecture and environment, named MSCE (Mobile Statistical Computing Environment) to support mobile statistical computing for novice or professional statisticians executing computing and analyzing using wired/wireless device. Users can submit their statistical computing using various devices, such as Notebook, PDA, PDA phone, or Java enabled cell-phone, that with embedded browser or e-mail and without any statistical software or tools. Users can submit their SAS program via various familiarized user interfaces, such as web browser, e-mail client, or designated java based agent, to backend server and get the response immediately no matter what kind of computing devices used any time and any where.

Acknowledgements:

This work was supported by the National Science

Council of the Republic of China under Grant No. NSC 93-2118-M-305-003 and No. NSC 94-2752-E-009-006-PAE.

References:

- [1] SAS: <http://www.sas.com/offices/asiapacific/taiwan/index.htm>
- [2] MINITAB: <http://www.minitab.com>
- [3] SPSS: <http://www.spss.com>
- [4] StatCrunch: <http://www.statcrunch.com>
- [5] Oliver Gonther, Rudolf Moller, Peter Schmidt, Hemant K. Bhargava, and Ramayya Krishnan, "MMM: A Web-Based System for Sharing Statistical Computing Modules," *IEEE Internet Computer*, May-June 1997, pp.59-68.
- [6] Chooichiro Asano and Akinobu Takeuchi, "Web-based statistical system by using the DLL," *Computational Statistics & Data Analysis*, 44(2003), 409-418.
- [7] Jasp: <http://jasp.ism.ac.jp>
- [8] Yue-Shan Chang, Tong-Ying Juang, "MASS- A Multi/web-Agent based Statistical Service," *WSEAS Transaction on Information Science & Applications*, Vol. 3, No. 6, June 2006, pp. 1122-1129.
- [9] Hai Zhuge, "Exploring an epidemic in an e-science environment," *Communications of the ACM*, Vol. 48, No. 9, Sept. 2005, pp. 109-114.
- [10] Lars Erik Holmquist, "Welcome to the mobile life!" *Interactions*, Vol. 13, No. 3, May 2006, page. 57 and 71.
- [11] Andreas Holzinger, Alexander Nischelwitzer, Matthias Meisenberger, "Lifelong-learning support by m-learning: example scenarios," *eLearn*, Vol. 2005, No. 11, November 2005, pp. 2.
- [12] Fiona Fui-Hoon Nah, Keng Siau, Hong Sheng, "The value of mobile applications: a utility company study," *Communications of the ACM*, Vol. 48, No. 2, Feb. 2005, pp. 85-90.
- [13] Satyanarayanan, M, "Mobile computing." *Computer*, Vol. 26, No. 9, Sept. 1993 pp. 81 - 82