GeoWeb in a Role of Routine Information System

Jitka Komarkova (University of Pardubice, Czech Republic) -
jitka.komarkova@upce.cz
Martin Novak (University of Pardubice, Czech Republic) - martin.novak@upce.cz

An easy and remote access to geographic information belongs to nowadays challenges in the field of geographic information systems because geographic information and their analyses play a significant role in decision-making process in public administration, private sector, and peoples' lives. Thus, technologies geo-enabling the Web have become important and today they are used as a part of regular information systems. It means significantly higher demands are made on them – they must be of the same quality as all other information systems. Security and usability of the system belong to the quality metrics. An influence of selected security measurements on server response time is studied. Next, a possible way of usability testing of GeoWeb is proposed in the paper.

Keywords: GeoWeb, Internet GIS, response time, usability testing

1 Introduction

Utilization of geoinformation (or spatial information) by everyone who needs it, belongs to the contemporary needs – almost everything happens somewhere. So, geographic information is more and more demanded not only by the business sphere but by the public as well. For many people it has become normal to use their services like searching the best route, realty, and many others. At the same time, today it is impossible to manage a territory, property, utilities, etc., without using geographic information. But because of the special nature of geographic information a special software tools are required to its treatment and analyses – geographic information systems (GIS) [13] - [21].

Traditional (professional) GIS are primarily from the users point of view quite complicated and costly software packages. It is estimated that about 90% of traditional GIS users use less than 10% of functions provided by their application [21]. This was the reason why cheaper and more user-friendly solutions were looked for to allow end users to use geographic information. The new solutions are based on Internet and mobile technologies and they are spreading rapidly over the world and Internet [16], [21].

Nowadays, many various technologies which allow access to geographic information in user-friendly environment exist. They allow remote and easy access of end-users to geographic information without any special education in the field of geoinformation science. Access is possible by means of variety devices, e.g. computers, notebooks, pocket PCs, mobile phones, etc. At the same time, one can find a lot of various terms which are used while talking about this kind of technology. Some examples are: GIS applications on Internet/intranet, GIS on-line, distributed geographic information [22], [23], Web-based GIS [20], Internet GIS, mobile GIS [16], [20], interactive mapping [5], [7], distributed GIservice [20], geo-enabled Web [19], GeoWeb, Internet map servers, and many others. These terms are sometimes understood as synonyms but it is not the best way of their understanding. For example, Internet does not provide only WWW service so Internet GIS has a different meaning from Web-based GIS. Mobile GIS is not the same as Internet GIS because mobile devices like PDA, and mobile phones use different protocols and technologies, and so on [20]. In the framework of this paper the term GeoWeb and the term Internet GIS will be used equally for Web-based solutions.

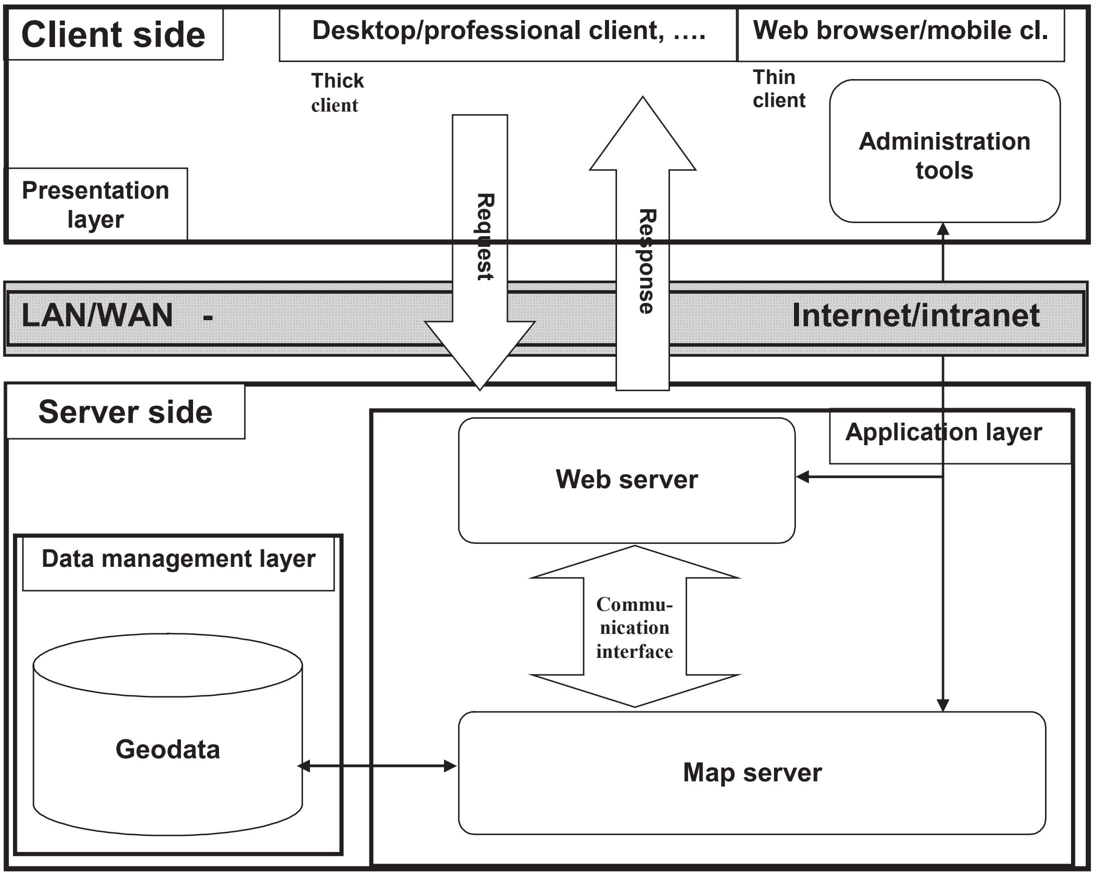
Historical development of methods used for interactive publishing of geographic information and an overview of contemporary methods can be found in [13], [16], [20], [22], and many others.

Architecture of Internet GIS follows possibilities of Internet technologies and architecture of information systems in general. So, they are usually based on n-tier client/server architecture and HTTP protocol is used as the most common communication protocol. Web services are used very often as well. It means the whole solution is divided at least into the next three main parts (see the Fig. 1) [1]:

- *Presentation layer* – simple software tools for users' work (client). They should be user-friendly but provide all needed functions. Usually, they are for free or for the lowest costs per user. Common Web browser or simple viewers (e.g. ArcExplorer) can be given as an example.

- *Application layer* – application logic; functions which obtain user questions, do all the work and return answers to the users.
- *Data management layer* – an important part which deals with data storing and data management. Both spatial and non-spatial data are treated.

Fig. 1 - Typical n-tier client/server architecture of Internet GIS (source: authors, based on [1], [20]).



Internet GIS are supposed to quickly provide a large amount of data from various sources including terabytes of data from remote sensing. Thanks to an existence of interoperability standards which are accepted by wide community, Internet GIS were found as a suitable domain for application of ideas of parallel and distributed computing [9], [20].

The complexity of spatially oriented problems, like land management, urban planning, and public administration, leads even to integrated systems approach, such as a spatial decision support system [26].

2 Related works

Various geo-enabling solutions have become a widely used and inherent part of information systems both in private and public sector. The aim of GIS including Internet GIS is the same as the aim of each information system - to collect, verify, treat, store, analyse and publish information. Just in the case of GIS a special type of information can be treated. It means life cycle of all kinds of GIS solutions should follow the same rules as life cycles of 'normal' information systems and GIS must meet the same quality requirements as the other information systems.

There has been done plenty of research work in the field of the software quality, quality of software development, and quality of information systems, e.g.

[2], [3], [4], [15], [17], [27], [28]. But no attention has been paid particularly to Internet GIS solutions although many facts should be taken into account during the process of choice of software package too [11].

Today, many software quality models exist. Each specialist can propose his/her own quality model which will meet the needs of the given situation [2]. Anyway, there is one quality model standardized by ISO/IEC: ISO/IEC 9126 - Information technology - Software Product Quality. This quality model is widely used for both measuring architectures and intranet applications [15], [17]. Its extended version was proposed for measuring quality of intranet/Internet applications so it can be suitable at least for the first studies devoted to the quality of Internet GIS too. This model uses functionality, reliability, efficiency, usability, maintainability, and portability as main software quality characteristics [15].

Usability is today a very important metric for measurement of software quality. The reason for its fast wide spreading is that there are many sophisticated information and communication technologies available but it is often hard to use them. So, today it is preferred to “make the design to fit the users” to the previous attempt: “make the user to fit the design” [24]. Contemporary users do not want to think. Applications must be intuitive. Designers of WWW pages should even remember that users do not read precisely the entire page and they are very often in a hurry [14]. Previously, attention was mostly paid to the system (e.g. software package) not to the end users and their needs during the system design. During the time users have changed but developers are sometimes slow to react to this evolution of the users and their needs. It is difficult and demanding to develop a usable system because it should be useful, effective, learnable, and likeable [24]. The aim of usability testing is to collect empirical data about the tested application. It means representative end users are observed when they use the application to perform selected typical tasks [24].

Concerning performance management and planning, a common approach for client/server systems, e.g. [6], can be used. As far as Internet GIS use Web server for communication, queuing model theory and other server location techniques should be taken into account while evaluating performance of the solution. Web server is modeled as an open queuing network. An upper limit of a performance capability of a server can be clearly found. This limit depends on a size of a published file [8], [18]. Agent technologies have been more and more used for speeding and improving information retrieval. But there are many languages which use specific features. Czech language can be given as an example – it uses many specific features and signs which must be considered [10].

Remote administration can be useful in some situations too but it is not so easy to implement it for all available software solution in the environment of the contemporary computer networks [12].

While talking about an information system, target groups of the users have to be considered. The classifications of the target groups of the GeoWeb users vary from author to author and they are dependent on the purpose of classification too. Anyway, following basic types of users are usually distinguished [16], [20], [25]:

- *High-end users*, e.g. data treatment specialists (GIS specialists), who can run spatial analyses as well and provide the results of their work to the other users.
- *Regular users*, e.g. civil servants, managers, controllers, regular customers, cooperating partners, etc. Regular, everyday use of Internet GIS is typical for this group of users. They usually need only several functions too. All

needed functions are known in advance and are used repeatedly. It can be supposed that they access Internet GIS by means of appointed Web browser or other defined client. It means their working environment is known in advance as well and can be influenced in the case of necessity.

- *Casual users*, e.g. tourists, residents, businessmen, etc. They use Internet GIS solution irregularly and casually. They are usually not educated in computer science and their skills how to use computer, may be very low. On the other side, only a few functions are interesting for these users. They usually need to select region of interest, select appropriate data layers, view geographic information, change scale, run very simple queries, and print outputs or save result maps. It is supposed that they can use various Web browsers, they may not be able to install any software, and their Internet connection can be slow.
- *Mobile users*, i.e. people who usually use wireless technologies and mobile devices like PDA or mobile phones to connect to a server and access geographic information and/or GIS functionality provided by the server. The users can vary from casual low-end users to high-end users. Utilities management (e.g. water, electrical utilities) can be given as an example of a branch where mobile GIS solutions are very often used as a regular tool by non GIS specialists.

Correct identification of target groups of the users and their needs is very important for the proper design of Internet GIS application and its subsequent utilizability.

3 Methodology

The ISO/IEC 9126 quality model was selected as a framework for studying possibilities of deployment of GeoWeb solutions in the role of regular information system (see Fig. 2). In according to the study [15] security (a sub-characteristics of functionality), availability (a sub-characteristics of reliability), and user-friendliness (a sub-characteristics of usability) belong to the very important software quality characteristics. Experimental measurements were done to study an influence of selected security measurement on response time of selected Internet GIS solution. Then, a set of typical tasks was proposed to perform a usability testing by means of method proposed by Rubin [24].

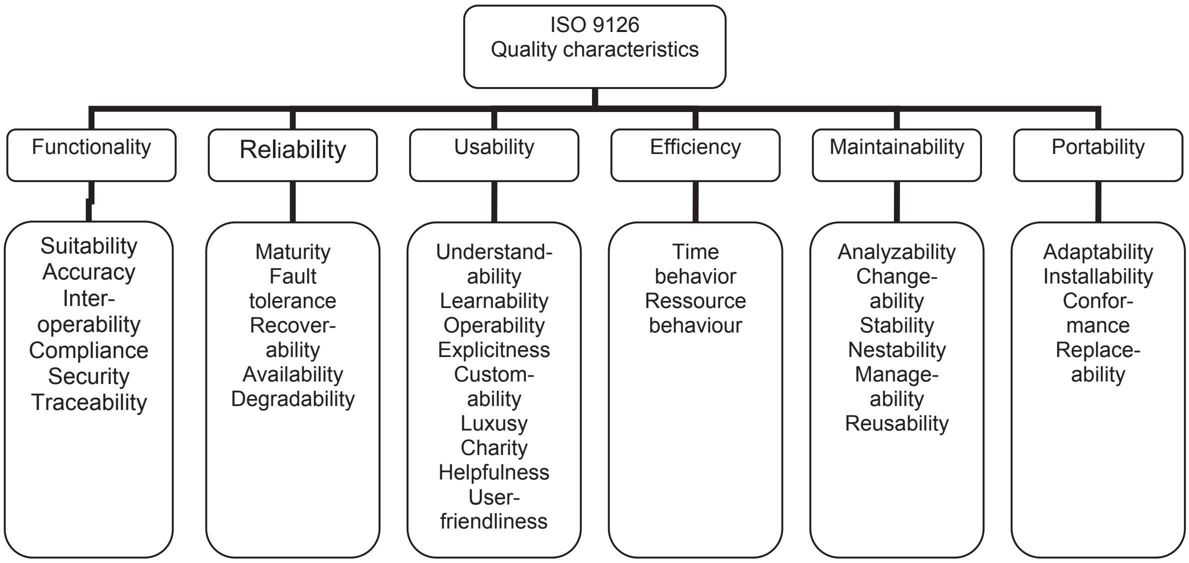


Fig. 2 – Extended ISO/IEC 9126 quality model [15]

4 Experimental Work

Security is a representative of functionality in extended ISO/IEC 9126 model. It is the first used sub-characteristic. A presence or an absence of a security measurement is used as a metric. The first set of measurement was done with an absence of a security measurement. Then, following security measurements were tested:

- Packet filtration – **http-filtr**
- Internet Protocol Security (IPSec), Authentication Header (AH), check sum (HMAC) calculated by means of MD5 algorithm (128 bits), transport regime – **http-ipsec-ah (transport-MD5)**
- IPSec, AH, SHA2 (168 bits), transport regime – **http-ipsec-ah (transport-SHA2)**
- IPSec, AH, MD5 algorithm (128 b), tunnel regime – **http-ipsec-ah (tunel-MD5)**
- IPSec, AH, SHA2 (168 b), tunnel regime - **http-ipsec-ah (tunel-SHA2)**
- Internet protocol security (IPSec), Encapsulating Security Payload (ESP), data encryption by means of AES256 algorithm (256 b), check sum calculated by means of MD5 algorithm (128 b), transport regime – **http-ipsec-esp (transport-aes256-md5)**
- IPSec, ESP, Blowfish128 (128b), MD5 (128b), transport regime – **http-ipsec-esp (transport-blowfish128-md5)**
- IPSec, ESP, AES256, MD5, tunnel regime - **http-ipsec-esp (tunel-aes256-md5)**
- IPSec, ESP, Blowfish128, MD5, tunnel regime - **http-ipsec-esp (tunel-blowfish128-md5)**
- Secure Sockets Layer (SSL), Diffie-Hellman algorithm (DHE) for key exchange, authentication protocol: RSA (length of private key: 1024 bits),

data encryption by means of AES256 algorithm (256b), check sum (HMAC) calculated by means of SHA1 (160 b) – **https (DHE-RSA-AES256-SHA)**

- SSL, RSA (1024b), RSA (1024b), AES256, SHA1 – **https (RSA-RSA-AES256-SHA)**
- SSL, DHE, RSA, AES256, SHA1 + IPSec, ESP, Blowfish128, MD5, tunnel regime – **https (DHE-RSA-AES256-SHA + ipsec-esp-tunnel-blowfish128-md5)**.

Time behavior is the second used sub-characteristic. It represents another quality characteristic in the quality model – efficiency. A real end-user response time is used as a metric in this case. It is measured in seconds on the client side as a real time between sending the user's command to obtaining the result from the server. Service time of the server and transmission time are used as supplemental characteristics. They were measured in milliseconds. An existence of the queuing model theory must be mentioned here. It takes into account that Web servers must process concurrent jobs. Each arriving job joins the queue and waits there until it is done. Response time of server is then calculated as a sum of a queuing time and a service time [18].

4.1 Experimental Conditions

UMN MapServer (see http://mapserver.gis.umn.edu/) was chosen as an Internet GIS solution for measurement purpose. Measurements were run on the following computers: server configuration: Intel Celeron (Coppermine) 600MHz, MB MSI 6309 (VIA 694x), 384MB SDRAM 100MHz, HDD Seagate 60GB 5400 rpm, Edimax 9130TXA PCI with Realtek 8139d chip (100Mbps, full duplex), operating system Debian GNU/Linux 3.0r1 with kernel 2.4.19, Apache HTTP Server 2.0.47, MapServer 4.0.1. Client configuration: AMD Athlon XP-1700+, MB ECS K7VTA3B (VIA KT333), 256MB DDR 266 MHz, HDD Western Digital 80GB 7200 rpm, Edimax 9130TXA PCI with Realtek 8139d chip (100Mbps, full duplex), operating system Debian GNU/Linux 3.0r1 with kernel 2.4.19.

For the purpose of this study data were stored directly on the application server to exclude a possible influence of network communication between application server and data storing server. On both computers only necessary processes were running during the measurements (e.g. httpd, syslogd, kernel processes).

Measurements were done in an environment of a real computer network: Ethernet based LAN with the full support of the speed 100 Mbps. Both server and client were dedicated to the experimental measurements only and requests were not generated simultaneously so the queuing time was not taken into account in this phase of research.

Requests were generated on the client side and send by means of 'wget' program which is a common part of Linux distributions. On the server side, Netfilter (a part of kernel for packet filtering) together with syslogd were used for logging the time. Both end-user response time (TC) and server response time (TS) were measured (see Fig. 3).

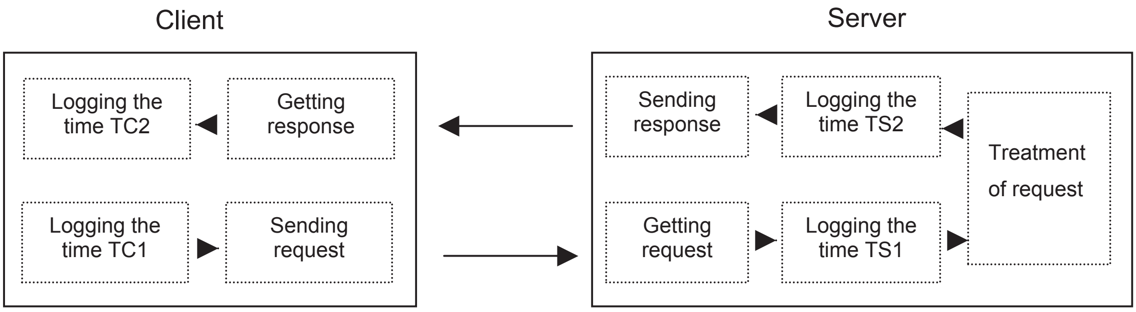


Fig. 3 – Measured times: end-user response time TC and server response time TS

4.2 Experimental Results

For each security measurement 20 measurements of response time were done. Obtained experimental results were statistically treated by means of exploratory analysis. The results are shown on the following figures Fig. 4 (average end-user response times) and Fig. 5 (relative end-user response time – related to http response).

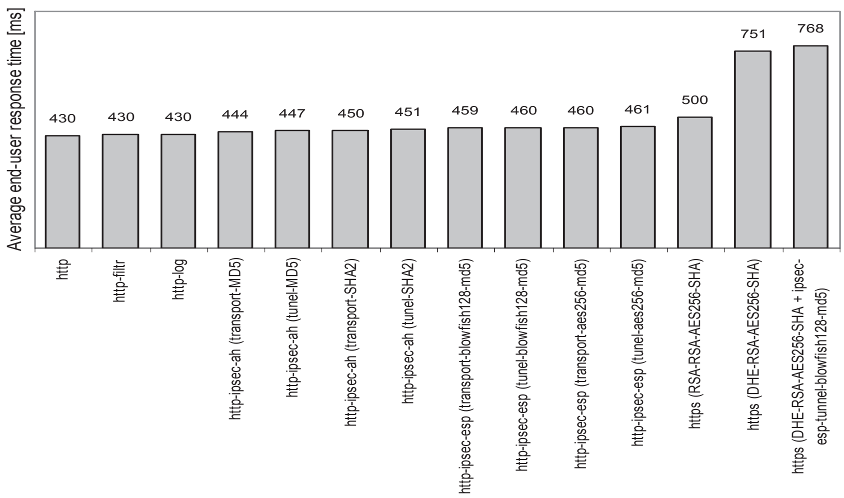


Fig. 4 - Average end-user response time

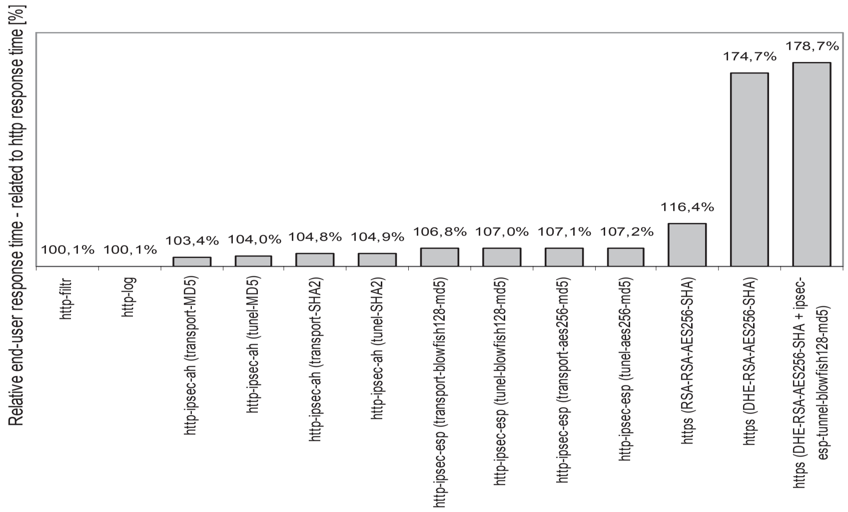


Fig. 5 - Relative response time – related to http response

Average server response times for each security measurement are shown on Fig. 6.

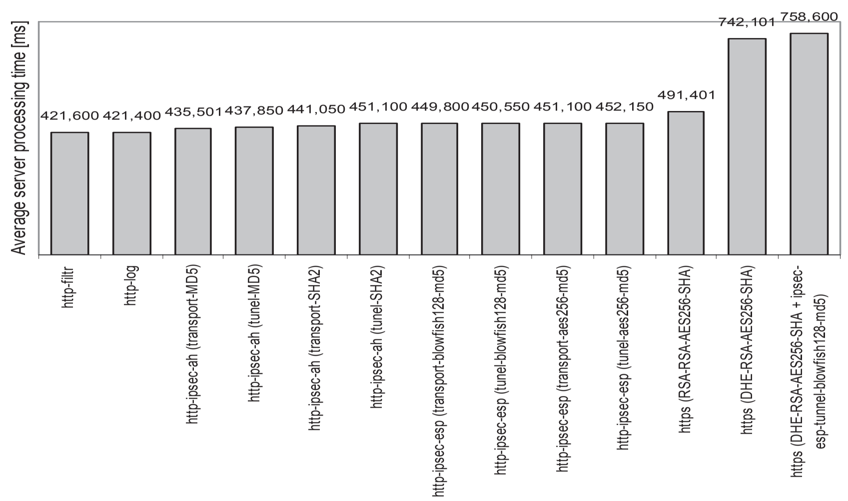


Fig. 6 - Average server response time

4.3 Usability Testing

As it was mentioned above, security and time behavior are not the only quality metrics of an information system. Usability of the application belongs to the significant metrics as well. Proposal of the set of typical tasks which should be performed by tested subjects during the usability testing belong to the most difficult tasks of the usability testing. The usability testing itself is time demanding and its duration can be significantly influenced by a properly defined set of typical tasks.

For the purpose of the first experiments with usability testing the following general set of tasks was proposed:

- Accessibility of Internet GIS applications (Finding a list of GIS applications on the web pages, Finding a given kind of map, Opening a map output, Return to the list of all available services (maps))
- Orientation within the map frame (Overview map, Menu and tools, Scale, Displayed layers and legend)
- Utilization of the map (Change of scale, Pan of the map by hand tool, Movement of the map by means of keyboard or map field arrows, Printing of the map, Adding and removing of data layers)
- Orientation in map (Finding the given feature (object) and its properties, Measurement of the distance between given features, Selection and cancelling the selection of features).

This set has to be refined in according to the tested GeoWeb applications but it covers all the functions needed by the users in general. A usability testing of one Web site providing geographic information by one tested subject took approximately 30 minutes when using the above proposed set of tasks.

5 Conclusion and future work

Today, Internet GIS are used as a regular information system for everyday work and even for critical applications like emergency systems. Their quality is still not considered although during the first part of the research devoted to the quality of Internet GIS it was found that many factors can influence quality of Internet GIS. An influence of selected security measurements on response time of UMN MapServer was studied. The factors which can influence a quality of GeoWeb have not been studied in deep yet so for the future work it is proposed to deal with more deep measurements and proposal of the whole set of metrics and their limitary values directly for Internet GIS because factors like data compression can be difficultly found in the other information systems.

ISO/IEC 9126 quality model can be taken as a framework for evaluation of the quality of GeoWeb applications which are used in a role of a regular information system. In any case, software quality metrics must be refined for the purpose of Internet GIS quality assessment to take into account the specific character of this kind of information technologies. Within the framework of the quality model a set of typical tasks for usability testing was proposed. The next step should be a heuristic analysis of the proposed tasks to exclude dependent tasks and shorten duration of the usability testing.

References:

[1] Alter, S.: *Information Systems: Foundation of eBusiness*. 4th edition. Prentice-Hall (2002)

[2] Azuma, M.: Software products evaluation system: quality models, metrics and processes - International Standards and Japanese Practice. *Information and Software Technology* 38 (1996) 145-154

[3] Canfora, G., Garcia, F., Piattini, M., Ruiz, F., and Visaggio, C.A.: A family of experiments to validate metrics for software process models. *Journal of Systems and Software* 77 (2005) 113-129

[4] Cukic, B.: The Virtues of Assessing Software Reliability Early. *IEEE Software* 22 (2005) 50-53

[5] Doyle, S., Dodge, M., Smith, A.: The potential of Web-based mapping and virtual reality technologies for modelling urban environments. *Computers, Environment and Urban Systems* 22 (1998) 137-155 [online], [cit. 2006-07-07]. Available from http://www.sciencedirect.com/

[6] Foxon, T., Garth, M., Harrison, P.: Capacity planning in client-server systems. Distrib. Syst. Engin. 3 (1996) 32-38

[7] Friedl, M., A., McGwire, K., C., and Star, J., L.: MAPWD: An interactive mapping tool for accessing geo-referenced data sets. *Computers & Geosciences* 15 (1989) 1203 – 1219

[8] Guyton, J., D., Schwartz, M., F.: *Locating Nearby Copies of Replicated Internet Servers*. Technical Report CU-CS-762-95. Department of Computer Science. University of Colorado – Boulder (1995)

[9] Hawick A., K., Coddington, P., D., James H., A.: Distributed frameworks and parallel algorithms for processing large-scale geographic data. *Parallel Computing* 29 (2003) 1297–1333

[10] Janakova, H.: Text categorization with feature dictionary problem of Czech language. *WSEAS TRANSACTIONS on INFORMATION SCIENCE AND APPLICATIONS* 1 (2004) 368 - 372

[11] Komarkova, J., Capek, J.: Czech Public Administration and Remote Access to Geoinformation. In *Politics and Information Systems: Technologies and Applications*. Proceedings of the International Conference. Orlando: Internation Institute of Informatics and Systemics (2003), p. 72 – 75, ISBN 980-6560-04-3

[12] Komarkova, J., Simonova, S., Dusek, V.: Geographic Information on the Web. *WSEAS TRANSACTIONS on INFORMATION SCIENCE AND APPLICATIONS* 1 (2004) 1185 – 1188

[13] Konecny, G.: Geoinformation: *Remote Sensing, Photogrammetry and Geographic Information Systems*. 1st publ. Taylor & Francis, London (2003)

[14] Krug, S.: *Don't Make Me Think: A Common Sense Approach to Web Usability*. 2nd edition. New Riders Press (2005)

[15] Leung, H., K., N.: Quality metrics for intranet applications. *Information & Management* 38 (2001) 137-152

[16] Longley, P., A.: *Geographic Information Systems and Science*. 1st edn. John Wiley & Sons, Chichester (2001)

[17] Losavio, F., Chirinos, L., Matteo, A., Lévy, N., and Ramdane-Cherif, A.: ISO quality standards for measuring architectures. *Journal of Systems and Software* 72 (2004) 209-223

[18] Moghal M., R., Hussain, M., Mirza, M., S., Jarral, M., W., Choudry, M., S.: Performance Evaluation and Modeling of Web Server Systems. *WSEAS TRANSACTIONS on INFORMATION SCIENCE AND APPLICATIONS* 1 (2004) 658 – 663

[19] Open Geospatial Consortium [online], [cit. 2006-07-07], Available from http://www.opengeospatial.org/about/

[20] Peng, Z.-R., Tsou, M.-H.: *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks.* John Wiley & Sons, Hoboken (2003)

[21] Peng, Z.-R., Zhang, Ch.: The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS). *J. Geograph. Syst*. 6 (2004) 95-116

[22] Plewe, B. S.: *GIS Online: Information Retrieval, Mapping, and the Internet*, OnWord Press, 1997

[23] Plewe, B., S.: What is DGI? [online], [cit. 2007-01-10], Available from http://www.geog.byu.edu/plewe/gisonline/dgi.htm

[24] Rubin, J.: *Handbook of usability testing: how to plan, design, and conduct effective tests*, John Wiley & Sons (1994)

[25] Schaller, J.: GIS on the Internet and environmental information and planning. In *13th ESRI European User Conference, 7. – 9. 10. 1998, Firenze, Italy,* [online], [cit. 2007-01-07], 2002, Available from http://gis.esri.com/library/userconf/europroc98/proc/idp27.html

[26] Sharma, D., K., Sharma, R., K., Ghosh, D.: A Spatial Decision Support System for Land Management. *Int. J. of Computers & Applications* 28 (2006) 50-58

[27] Veenendaal, E. Van, Hendriks, R., and Vonderen R. Van: Measuring software product quality. *Software Quality Professional* 5 (2002) 6-13

[28] Vlahavas, I., Stamelos, I., Refanidis, I., and Tsoukiàs, A.: ESSE: an expert system for software evaluation. *Knowledge-Based Systems* 12 (1999) 183-197