Abstract—This article describes problems connected with web-application user identification. There are discussed some specific conditions and issues in this kind of user identification due to features of front-end applications and application protocols used in background connection. Some tips and techniques are accentuated to use for undesirable multi-user issue detection and avoidance. The usability is discussed across common web-application in comparison to special application such as public administration systems or geographic information systems.

Keywords—GeoWeb, Identification methods, Information system, User identification, Web-application

I. INTRODUCTION

The usability of information systems depends on many features. One of these features is a user-friendly environment with a simple user interface. In present, modern systems use thin clients realized as web page or web-site presented via browser, and the robust system backend running on powerful server hardware. This configuration forms the “web-application” as a common information system design.

Most of the information systems need a basic user access control subsystem at least. The access control subsystem can be designed for different security levels. It depends on the target application purpose and security measurements suitable for the owner of the information system.

The information system designed as the web-application deals with specific environment of the worldwide web rules and conditions. In general, the web is unsecured. Many issues and threats can encounter the information system in the form of the web-application. Designers are responsible for sufficient solution of such situations.

Each one solution depends on the security demands of the information provided by the system. Some systems need to distinguish the separate identities of their users, the user’s accounts uniqueness is important for others. In general, each one needs to assign an access rights to provided information by the user or user-group identity. In the web environment, the process of user identification is more difficult.

II. WEB-APPLICATION ARCHITECTURE

A web-application consists of two basic part overall. As a network application it has a server part and one or more client parts (client/server architecture). “If we look at the client-server model in detail, we see that two processes are involved, one on the client machine and one on the server machine. [1]”

The server part is commonly found as a running service of a dedicated network server. The client part can be found on workstations as an application used to browse through the worldwide web service (web-browser), see Fig. 1.

Client-to-server connection is realized by unsecured (HTTP) or secured (HTTPS) protocol going over transport protocol (TCP). The common data format transmitted between server and client part is hypertext. Server part also provides other file types transmission including graphics and multimedia.

Other file type transports from client part to server is unhandy and only way to realize it is the interactive upload service provided by form HTML element. By this reason, any file transport from client to server is problematic, and it requires user’s effort. [2]

A specific attention must be focused on the application security. Fig. 2 illustrates various possible attacks on the web-application. The points of the possible vulnerability also are demonstrated, so it can be seen where they happen.

III. IDENTIFICATION METHODS OVERVIEW

The usage of user identification is various, but two main purposes are commonly used. The first one is used for monitoring of individual user’s activity, usually without user’s
Three groups of identification method exist. First, user can be identified by some unique knowledge. Typical case is using of a password. Next group includes methods based on ownership of some unique item, typically a token or a thing. The methods of the third group are based on some specific feature (mainly body feature) of given user as some unique biometrics parameters.

A. Knowledge Oriented Methods

These methods are based on user knowledge of given fact, i.e. username and the corresponding password. For user identification by this method a unique username is required. But, the username and password can be accidentally disclosed, and the security of the identification fails.

Some form of shared or pre-shared password or pass-phrase without the username can’t be used for sufficient user identification, because more users could use the shared password to be accepted by a system. Such form of identification can be used for access into shared environment, i.e. a course of e-learning system, where is sufficient to distinguish between given groups of users only.

B. Ownership Oriented Methods

All the methods based on ownership of some item are inappropriate for secure user identification, when the owned item can be passed to other user or can be stolen. The improper user can be identified in this case, because of ownership of foreign identification item. Additional measures must be realized to avoid this situation.

C. Biometrics Methods

Specific parameters of human body or user’s specific behavior can be used for user identification. Many biometrics methods depend on special techniques or complicated equipment is needed for such body or behavior features extraction. However, some methods can be realized using common computer accessories or peripheral.

IV. IDENTIFICATION METHODS USABILITY

The usability of identification method depends on several basic conditions or features of the target system. Each method requires other type of features extraction, and each system requires other form of feature transmission across its infrastructure.

A. Network System Specific Conditions

The common feature of the network architecture designed systems is user remote access. An identification feature transport across the network can be slightly difficult there.

First, the amount of the information used for the identification can be bigger in many cases. A transport capability problem originates by this reason. The second problem springs from security issue, because to keep the identification feature safe is necessary. Third, the question of technological capability to gain the information by the client
part from the user can be fundamental at several methods. A qualified decision about whether to extract basic features from identification information by server or by client part must be done. It depends on used transport technology and on network application type.

The next problem originates from technical requirements. Some methods require sophisticated reading equipment that can be more expensive than others. It makes some methods unusable for network environment because each one client has to have its own expensive reader. It’s the disadvantage of some ownership-oriented and biometrics methods especially, so these can be used only by systems where the client count is limited.

**B. Web-application Specific Conditions**

The influence of conditions described above is stronger when using web-application environment. The transport is very limited by the protocol specification that is why only a small amount of data can be transmitted from the client to server part. The need of a quick identification is evident, so the transport speed has to be sufficient.

The basic application protocol used to hypertext transport (HTTP) is unsecured that is why it is not well usable for identification features transmission. The secured form (HTTPS) provides secured sockets layer (SSL) capability to encapsulate that transport into an encrypted stream. It increases the communication security and enables to keep the secret data safe. But, the encapsulation consumes some system and connection resources, such as computing time or transfer capacity.

Other security issue is connected with processing of the identification data obtaining or feature extraction. Basic technologies providing a scripting capability of the server or client part aren’t designed to be usable securely. The main problem rising from the insufficient security of scripting is on the client side.

The client part of the web-application is in general realized by web-browsers. Scripting support is limited to a few scripting languages only (JavaScript, VBScript) with many incompatibility issues between browsers. Along with them some security questions are solved, because the source code of the web page and other resources linked in are viewable and can be inspect by the browser additions as a plain text. Some compiled additions (i.e. Java-applets) can be decompiled easily by free utilities, so the security is out of the question. [3]

**C. Identification Usage Specific Conditions**

More conditions lead from the usage of user identification. If the identification is used for statistical monitoring of user behavior, the system needs to distinguish between users, but the knowledge of their identity is not necessary. The access control management system usually needs the user identity knowledge.

In the web-application environment the tracking cookie method can be used for the individual user identification. This cookie-based user identification requires the web server’s and web-application access log analyzes. “For tracking the users’ behavior the log files are extended with cookies and some other fields as well. Cookies are the most common way of client side data storing, ... [4]”

**V. Examples of Web-application**

The common web-application characteristics are described above. Special features are needed by application focused on private or secret information providing. These two examples are used for demonstration and better explanation:

**A. Public Administration Systems**

This group of information systems includes various types of applications used to provide public or private citizen or property data. A part of information can be provided as public data (some registrars, business identification data, etc.). The access to other part of information needs to be regulated by some rules (private or sensitive data, criminal record, penal register, etc.).

The grant of the access to provided data is based on either user identity or user’s membership to a group. The user’s behavior monitoring can be used for a better usability of the web-application. The system can predicate the next step of the user’s way across the web site, and offer direct links to related data. Some system can provide direct or indirect web-application personalization.

**B. Geographic Information Systems**

The geographic information systems often provide the information with web-based interface as GeoWeb sites. The management of user access uses the same principles as described in previous paragraph. But, the GeoWeb often displays data from different servers with different access rules. The access management needs to be solved not only for the entire web-application, but for the individual layers of the map service.

**VI. Motivation for GeoWeb User Identification**

As it was described above, the GeoWeb sites can use data from more backend systems and servers, with various types of security level and access rules. It is the motivation for user identification. The GeoWeb site has to distinguish individual users or users group to allow or deny access to some specific map-layers or geo-information. It is important on systems used for tasks as Emergency action plan, Evacuation plan and procedure during floods, or Critical infrastructure and key resources.

**A. Geographic Information Systems Architecture**

If the GIS provide a GeoWeb interface, it has the form of the typical client/server application with the three basic layers:

**Data layer** can be provided by the given server, but some data can be dynamically loaded or linked from another server or public services across the local network or Internet.
Application layer has to process the user requests and provide appropriate responses. This task is typically divided to be processed partly by server and by client side of the system.

Presentation layer is the user interface used to collect the request and to visualize the response. [5]

B. Geographic Information Systems User Groups

The GIS provided as web-application have some user groups with various relations to the presented data. Each one user group needs a different access to have granted. Because the GIS application can be able to analyze some information on demand, the user access has to be solved not only as the "data access", but the solution has to be able to manage access to run these analyses. There are these user groups:

High-end users – some specialists with access to run a high time or system resources consumption tasks, i.e. spatial analyses, etc.

Regular end-users – employees or regular customers, etc., with access to only several functions or data.

Casual end-users – citizens or common guests of the web-application with access to only some analyses results or public data. [6]

VII. USER IDENTIFICATION METHOD COSTS

If we need to add the user identification functionality to web-application, the costs of this addition will be the matter of the user identification method choice. For this purpose the costs of addition needs to be evaluated. The cardinal evaluation is more difficult and can be used to express the absolute financial costs. In many cases the ordinal evaluation is sufficient. It provides only the pseudo-values usable for the methods implementation costs comparison. These additional system upgrade costs can be expressed as a total of these values:

\[ \sigma \] – server side addition implementation costs
\[ \kappa \] – client side addition implementation costs
\[ \tau \] – data transmission addition implementation costs

Each one of these values consists of three parts. First, there are hardware costs. It means the value of the hardware components that are added to given system part (server, client or transmission circuit) for the user identification process needing. Most of the methods do not need additional server side hardware, but some client side hardware can be necessary.

Second part of the value represents the software costs, that is the value of additional software needed to user identification feature of the system. The maintenance of the software is included in this value.

Last value depends on the computing time when system processes the user identification. The costs of this time can be most important in some critical applications.

A. Server Side Costs

\[ \sigma_{SW} \] – server side hardware costs are low or none for most user identification methods.

\[ \sigma_{SW} \] – server side software costs are very dependent on type of method; if the data processes the server side, there is only one instance of the server software, but more data needs to be transmitted from clients.

\[ \sigma_{I} \] – server side identification costs depends on the identification process.

B. Client Side Costs

\[ \kappa_{HW} \] – client side hardware costs are low or none for some methods, but for others can be very high.

\[ \kappa_{SW} \] – client side software costs depends on the element processing collected data; if this process is provided by server part, no more client software addition needed, and this costs go to zero, but the client side data processing increases the costs with each one new client requesting to add some software instance in most cases licensed and expensive.

\[ \kappa_{I} \] – client side identification costs depends on the used method and the way of processing.

C. Data Transmission Costs

\[ \tau_{HW} \] – data transmission hardware costs are none or very low, if the Internet is used as the transmission media; the system designer can not modify the infrastructure, excluding of usage in private network environment.

\[ \tau_{SW} \] – data transmission software costs explanation is similar to above described \[ \tau_{HW} \].

\[ \tau_{I} \] – data transmission identification costs depend on the amount of data the system transmits from client to server side; if there are used some method with identification processing located in the server part, the amount of data to transmit can be high.

VIII. IDENTIFICATION METHOD COSTS COMPARISON

The comparison by the costs is one of the important parts of the decision what user identification method use. In this point it is necessary to say that the weight of this criterion may be very different in dependency on the web-application usage. For the example of comparison see Table I. Some identification method implementation costs are compared by pseudo-values none, low and high, because the real costs are not evaluable directly in general case.

If we set a weight for these criteria, the result will be expressed as the sum of pseudo-costs. The client side hardware
and software costs have to be multiplicatively by the count of clients. All the per-identification costs must be multiplicatively by expected identification processes in the calculated period.

This example displays the password method as the low costs way. There are none client side and data transmission costs. The value of $k_{HW}$ is important for web-application, because of many clients in such system expected. Other method with none client side hardware costs is speaker recognition. Each one new personal computer includes sound interface with or without microphone (if none, it is cheap to buy it) today. It is sufficient hardware for speaker recognition processing. But, if the client side processing of the audio signal is needed, the client side software will be expensive. If the audio signal is transmitter to be processed server sides, more costs for the transmission will be expected.

**IX. CONCLUSION AND FUTURE WORK**

The usability of user identification in the web-application environment is the complex problem. The common task of user identification by information system is limited by the specific conditions and rules of the web environment and communication across the Internet.

The special focus of GeoWeb applications leads to more specific rules and needing given by various tasks of such systems. More users and user groups access levels and the security reasons of the provided information make the solution more difficult.

The decision what identification method select for concrete usage depends on the costs of that solution. At the other hand the usability of identification method depends on more characteristics such reliability and success ability. This problem can be eliminated by use of combination of two or more methods, but it may lead to more additional costs.

The future work is to elaborate more precisely the evaluation and comparison of identification methods. Inclusion of the reliability in described evaluation model is necessary, because the costs can not be the only criterion for the decision.

Some identification methods need to be tested in this specific environment with the focus given to the client side implementation. The server side can be modified or upgraded easily, but the thousands of user’s browsers used as client sides can not be modified above the options of the web and hypertext standards.

**ACKNOWLEDGMENT**

This work was supported in part by the Grant Agency of the Czech Republic under Grant project no. 402/09/0219.

**REFERENCES**


