Expert System for Security Audit Using Fuzzy Logic

K. Kozhakhmet, G. Bortsova, A. Inoue, L. Atymtayeva

Kazakh-British Technical University
Tole bi st., 59
Almaty, Kazakhstan
kanik85@gmail.com, gerdabortsova@gmail.com, inoueatsushij@gmail.com, l.atymtayeva@gmail.com

Abstract

Information security auditing plays key role in providing any organization's good security level. Because of the high cost, time and human resource intensiveness, audit expenses optimization becomes an actual issue. One of the solutions could be development of software that allows reducing cost, speeding up and facilitating Information Security audit process. We suggest that fuzzy expert systems techniques can offer significant benefits when applied to this area. This paper presents a novel expert systems application, an Expert System in Information Security Audit (ESISA).

Today organizations, facing with a wide range of potential threats to their information security (IS), are increasingly interested in high level of it. One of the best ways to estimate, achieve and maintain security of information is an Information Security auditing. Audit of security (broadly-scoped) is a complex, many-stage and labor-intensive process involving high-qualified specialists (experts) in IS, what makes it a quite expensive service. There are many types of audit, including certain security standards (e.g. ISO 27K) compliance audits.

Typically, information security audit is conducted in the following steps [1]:

1. **Scoping and pre-audit survey**: determining the main area of focus; establishing audit objectives.
2. **Planning and preparation**: usually generating an audit workplan/checklist.
3. **Fieldwork**: gathering evidence by interviewing staff and managers, reviewing documents, printouts and data, observing processes in action, etc.
4. **Analysis**: sorting out, reviewing and examining of the accumulated evidence in relation to the objectives.
5. **Reporting**: reviewing all previous stages, finding relations in the collected information and composing a written report.
6. **Closure.**

Each of the stages is accompanied with a large amount of information, which needs to be recorded, organized and, finally, analyzed.

One of the efforts taken in reducing expenses and facilitating audit is using of helping tools for identifying the gaps that exist between certain security standard and an organization's security practices, like checklists and questionnaires. For example, ISO 17799 Checklist ([2]) provides number of audit questions (like "Whether responsibilities for the protection of individual assets and for carrying out specific security processes were clearly defined."), each corresponding to particular section of the standard (4.1.3 for the previous example). ISO IEC 27002 2005 (17799) Information Security Audit Tool, described in [3], offers several hundred audit questions, stated in yes-no form (e.g. "Have you reduced the risk of theft, fraud, or misuse of facilities by making sure that all prospective employees understand their responsibilities before you hire them?"), pointing to security practices that need to be implemented and actions that should be taken (in case of "no" answer to question). So, the auditing may be viewed as a process of asking questions and analyzing answers to produce recommendations.

Of course, these tools are very useful to auditors and security related staff. But the questionnaires don't give an overall impression of organization IS level, entries of the checklists are too general (not concrete, not related to particular organization's actual policies, procedures, etc.). Such kind of disadvantages doesn't allow them to be used independently, without any additional security measurements.

Another step forward developing effective tools for audit is a knowledge base for Chief Information Security Officers (CISOs) assisting them in justifying their information security management decisions, presented in [4]. Key components of the base are: "Asset", "Source" (standard), "Vulnerability", "Step" (a refinement of part of a "Guideline" in particular standard) and others. Every "Step" is linked with an asset it protects, type of vulnerability it is against and also cross-references to other stored guidelines. The proposed tool provides a search in the knowledge base for guidelines in standards using their components.

That is, a sort of meta-model of security standards' recommendations could be constructed.

We think that considerable expenses accompany regular security audit of companies could be significantly reduced by intellectual software, capable of substituting human specialists in performing IS audit. This is a good field for application of artificial intelligence techniques, like expert systems.

Expert System in Information Security Audit

An expert system (ES) is a computer system that emulates the decision-making ability of a human expert. (Jackson 1998)

The knowledge in expert systems, commonly represented in form of IF-THEN type-rules, may be either expertise or knowledge that is generally available from written sources. [5] We think that in IS field, along with human knowledge, security standards' (ISO/IEC, COBIT and ITIL, in particular) recommendations can also serve as a source of expertise and may be translated into rules.

We consider implementing question-answer interaction between user and system, similar to checklist and questionnaire principle: ES will take user's answers on auditing questions, analyze them, and output a result in form of recommendations.

A little more detailed procedure of audit, performed by the expert system:

1. Company information acquisition: defining assets to be protected (equipment, data, etc.). Depending on this, the system will prepare some general questions to start from.
2. Process of obtaining information by the system from personnel by asking appropriate (possible in particular situation of the organization, described in stage 1) questions.
3. Expert system's logical inference.
4. The system produces the output as a list of recommendations.

In comparison with audit process described in previous section, looks much easier. I.e., our idea is to automate some stages of the audit. In our opinion, expert systems technique has much to offer in information security.

Some of advantages of the use of expert systems (according to [5]), particularly in IS field are:

- **Reduced cost.** Development of an expert system is relatively inexpensive. Taking into consideration an opportunity of repeated use by multiple organizations, the cost of the service per client is greatly lowered.
- **Increased availability.** Expert knowledge becomes available using any suitable device at any time of the day. Web-based expert systems open up ability to access expertise from any Internet connected device. In some sense, "expert system is the mass production of expertise." (Giarratano & Riley 1998)
- **Multiple expertise.** Using knowledge from multiple sources increases total level of expertise of the system. In case of IS, a combination of number of security standards' recommendations and knowledge of several independent specialists could be used.
- **Time saving.** IS auditing is a time consuming process. Expert systems at some phases of audit (analysis of gathered evidence, reporting) can save days (or weeks) by faster responding (in comparison with a human expert) and reducing amount of paper work.
- **Steady, unemotional, and complete response at all times.** By the use of programs, human factor influence decreases.

We believe that developing web-based Expert System in Information Security Audit (ESISA), from the first, practical side, will save time and money of companies-clients, and, from the second, theoretical side, it will be a good fundamental experience for further development of methodologies for applying Artificial Intelligence techniques in IS field.

Previously expert systems approach in security area was applied in computer security auditing. An Expert System in Security Audit (AudES), designed for automating some audit procedures, like identifying potential security violations by scrutinizing system logs, described in [6]. But the field of expert systems methodology application in information security audit in its broader sense, i.e. not only IT, (what actually we would like to implement) remains largely untouched.

Information security usually divides on administrative, physical and computer security. We're planning to involve each of those types in our system. If to be based on ISO 27K, some of issues, those will be included, are: asset management (corresponding chapter 7 of ISO), human resource security (8), communications and operations management (10), access control (11), incident management (13), etc.

But we decided to go further in increasing of our ES human thinking pattern emulation accuracy by adding uncertainty management ability, i.e. developing *fuzzy* expert system (expert systems using fuzzy sets and logic for inference) in IS auditing. The exploitation of the tolerance for uncertainty underlies the remarkable human ability to make rational decisions in an environment of imprecision. [7]

Handling Uncertainties

We think that the task of developing ES in broad scale audit requires methods, more sophisticated than classical expert systems. They don't capture all the aspects of complex procedures, such as security estimation, which involves so many factors.

In real life, people do not often think about problems in terms of crisp names and numbers, they constantly deal with wide range of uncertainties. This is also applicable to professionals, when they are solve problems. [8] The subjective judgment of experts produces better results than objective manipulation of inexact data [9].

Experts in information security usually operate with fuzzy terms, such as "sensitiveness" (e.g. when applied to information), "completeness" (job applicant's CV) and so on. To handle uncertainties like this we consider applying another Artificial Intelligence technique – fuzzy sets and logic – those are effective tools for approximate reasoning, in comparison with traditional methods.

Fuzzy inference method was already used in risk assessment field (it is described in [10] and will be discussed further), which itself contains great value of uncertainties. We can use it for information security risk management, which is necessary in audit.

In information security, likely in every field, where humans are involved, things like perception take place. For example, auditor asks from user: "How frequently do you change your password?" He doesn't expect answers like "often", "rarely", because usually people's perceptions differ; furthermore, user may have distorted concept about information security. Here auditor perception is more adequate than auditees' perception. A numerical value (e.g. password changes per month) would be absolute, independent and, therefore, more sufficient answer. Fuzzyfication is performed on expert's side (he decide, if it is often, rare, etc.).
Of course, fuzzy logic and sets approach is advantageous here. The need of fuzzy logic is going to be proved in this paper.

System Modeling (Framework)

The following terms play a key role in organization's information security assessment [1, 4, 11]:

- **vulnerabilities**: any weaknesses in the system of controls that might be exploited by threats;
- **threats**: generally people, things or situations that could potentially cause loss;
- **impacts**: what would be the (worst case) effects if some of those threats actually materialized.

In order to perform qualified security estimation, an auditor should think carefully about each of those things.

We decided to follow their thinking pattern and define assets of the organization, vulnerabilities that may exist, threats, particular harm could be inflicted to them, and also consider the impacts of those threats.

On the scheme below (Figure 1) you may see these categories organized in 3 layers, several samples for each are given (they are going to be discussed further).

There are two asset types: physical assets (for example, computers, servers, etc.) and information (e.g. employees', clients' data stored in databases), which are to be protected. (ISO/IEC 27002, 7.1.1, "Inventory of assets")

Each of the assets matches one or more vulnerability it may have, each of the vulnerabilities is influenced by several factors (white boxes in 2nd layer), and may cause particular threat(s) materialization with some possibility. For example, physical security weakness, like poor physical entry controls (ISO/IEC 27002, 9.1.2, "Physical entry controls") depends on proper use of authentication controls and good monitoring, monitoring in turn depends on turnover rate on guard's position and their background checks; this weakness may become a cause of physical assets damage, or sensible data theft, or both.

Because the possibility of something to happen, especially an IS event, is very hard to evaluate precisely, it should be represented as fuzzy term. In order to calculate overall possibility, all factor's impacts should be taken into account.
Impact of vulnerability on the particular threat is reflected in rules, which have the following pattern:

```
IF vulnerability is very serious, THEN threat execution possibility is low/moderate/high (fuzzy value).
```

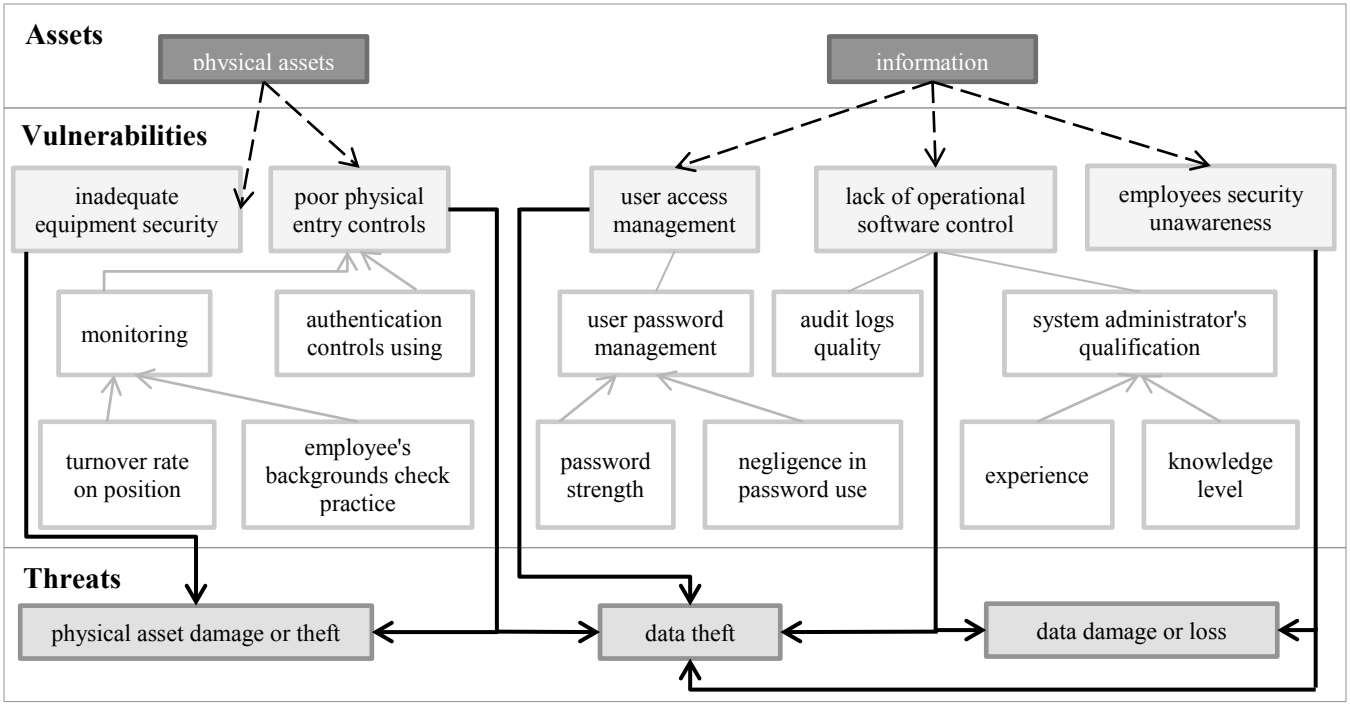

Figure 1: Audit scheme

For example:

IF an equipment security is *inadequate*, THEN physical asset damage or theft *slightly* increases.
IF a physical entry control is *poor*, THEN physical asset damage or theft *greatly* increases.

(ISO/IEC 27002, 9.2, "Equipment security", 9.1, "Secure areas")
According to this principle, there will be as many rules with same consequence (differentiating by degree of severity), as many vulnerabilities influence this threat. In order to produce one value, corresponding fuzzy numbers are summed and divided by maximal numbers * quantity (e.g. *high* fuzzy number).
We can also consider impacts of materializing of these threats in money equivalent, in order to perform some risk assessment.
There is no such thing as an "exact" value of risk. Risk assessment is based on imprecisely defined inputs: the likelihood of the threat being exercised against the vulnerability and the resulting impact from a successful compromise. For example, in [10], Security Management System robustness (with values inadequate, good and excellent) and severity of consequence (category of health harm from 1 up to 5) of incident on the industry are taken as the input, the value of the risk (negligible, low, moderate, high and unacceptable) is the output.
In our system risk could be calculated in the same way, as a function of likelihood of the threat, found as summation of vulnerabilities impact rates, and size of possible impact in money equivalent. According to these risks, factors, those lower security, may be sorted and recommendations are given labeled with a requirement level.
According to Brander [12], we can use keywords in our recommendation reports of expert system like "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY" and "OPTIONAL", which could be implemented to fuzzy variables. These keywords can deeply and clearly show recommendation's priority and notification manner.
Once common methodology is defined, one of the issues that arise is how standards' recommendations will be translated to the rules and what kind of inputs the system will gather.

Knowledge Acquisition and Inputs

The simplest case for inputs is numerical values: this is, for example, turnover rate (per cent of employees substituted on particular position during a year), employee's experience (years). Let's look at example.
According to ISO 27002 "Control of operational software" (12.4.1) "a" step, "the updating of the operational software, applications, and program libraries should only be performed by trained administrators upon appropriate management authorization".
It cannot be directly figured out if employee is *trained* or not trained enough. Admin's qualification depends on his/her experience and knowledge. Experience may be retrieved as simply a numerical value. Of course, knowledge of human, even in restricted field, cannot be assessed by asking one question.
Not only direct values asking need may appear. For example, some test could be provided:
1. When setting permissions in NTFS for an individual's network drive, which option(s) of the following levels do you give a default user?
Answers: Full Control / Modify / Read & Execute / Read / Write.
2. What do administrative shared folder names always end with?
Answers: # / $ / @ / % / ~
3. Which one of the following is equal to 1 kilobyte (KB)?
Answers: 512 bytes / 1000 bytes / 1024 bytes / 1028 bytes / 2048 bytes.
4. etc.
The score, expressed in %, is also a fuzzy variable.
Sample rule, displaying system administrator's experience, knowledge level and qualification:

IF employee is sufficiently experienced AND score is very high THEN employee is well qualified.

Use of tests exists in many aspects, like User security awareness (8.2.2, "Information security awareness, education, and training") estimation. Each member of particular user group may be offered to answer some questions like (multiple choice test, one answer is correct):
1. What is true?
- Leave terminal logged in is a bad security practice; (correct)
- Frequent logging in and logging out leads to computer's hardware faster deprecation;
- Logging out when leaving a work place is a good corporate culture indicator;
- Constantly logging in and out is time consuming.

2. Do you use your personal laptop at work? If no, do you want to?
- No, I think it's reasonable expense; (correct)
- No, I don't want to buy my own;
- Yes, it is convenient;
- Yes, personal laptop is a secure decision.

All scores of the group could be combined in one value (average score), expressed either in %, or a number from 0 to 1.
But some of variables that not explicitly expressed in numbers could be still obtained using 1 question. It refers to a situation when a particular quantity consists of several

simple (true-false valued) weighted components, it could be calculated as a checklist.

Let's consider one of the aspects of user access management issue, password managing, as an example.

ISO/IEC 27002 "Password use" (11.3.1):

"All users should be advised to:

a) keep passwords confidential;
b) avoid keeping a record (e.g. paper, software file or hand-held device) of passwords, unless this can be stored securely and the method of storing has been approved;
c) change passwords whenever there is an y indication of possible system or password compromise;
d) select quality passwords with sufficient minimum length which are:
e) easy to remember;
f) not based on anything somebody else could easily guess or obtain using person related information, e.g. names, telephone numbers, and dates of birth etc.;
g) not vulnerable to dictionary attacks (i.e. do not consist of words included in dictionaries);
h) free of consecutive identical, all-numeric or all-alphabetic characters;
i) change passwords at regular intervals or based on the number of accesses (passwords for privileged accounts should be changed more frequently than normal passwords), and avoid re-using or cycling old passwords;
j) change temporary passwords at the first log-on;
k) not include passwords in any automated log-on process, e.g. stored in a macro or function key;
l) not share individual user passwords;
m) not use the same password for business and non-business purposes."

These guidelines could be clearly divided into two parts: concerning user's negligence in password managing and password strength. Password security variable (which is going to be computed) represents a possibility of password to be stolen (number from 0 to 1), which can take values, say, high, moderate, or low and depends on two parameters, mentioned in previous sentence.

At first, we will try to compose some questions for user about how he/she manages his/her passwords: one question for one variable to be retrieved.

Question 1. Mark points you think are true for you:

- My colleagues/family members/friends or somebody else know my password. :0.2
- I consider writing down my logins and passwords on paper, storing them in files, or let my browser remember them very convenient way not to forget my passwords. :0.15
- If something suspicious happens, I don't think it is necessary to immediately change my password. :0.25
- I don't change my password without any serious reason, my memory is not so good to remember all this stuff.
- I use a default password, I think it is strong enough. :0.25
- I advocate a use of same password in multiple services. :0.15

The exact value of negligence level in password using is computed as a sum of coefficients for all points that were matched as true (value from 0 to 1).

Question 2. My password normally:

- is difficult to remember
- is a default password, like password, default, admin, guest, etc. :0.2
- contains dictionary words, like chameleon, RedSox, sandbags, bunnyhop!, IntenseCrabtree, etc. :0.1
- consists of words with numbers appended: password1, deer2000, john1234, etc. :0.15
- is one of common sequences from a keyboard row: qwerty, 12345, asdfgh, fred, etc. :0.3
- contains personal information, like name, birthday, phone number or address. :0.15
- contains symbols such as (mark each):
  - Lowercase letters (26)
  - Uppercase letters (26)
  - Numbers (10)
  - Punctuation marks (5)
- has average length: (specify number of characters)
- (not an option: using 2 previous options number of possible combinations of characters is calculated as (summary number of symbols)^(length); coefficient for this question is 0.1 if combination is bigger than 10^12, and combinations number / 10^12 * 0.1 else)

Value for the question is calculated as sum of coefficients of all entries.

These two values (value of negligence level and for the password strength) are subjects to fuzzyfication into fuzzy subsets like weak, good, strong for password's strength and low, moderate, high for negligence (a sample of fuzzy sets you can see at Figure 2).

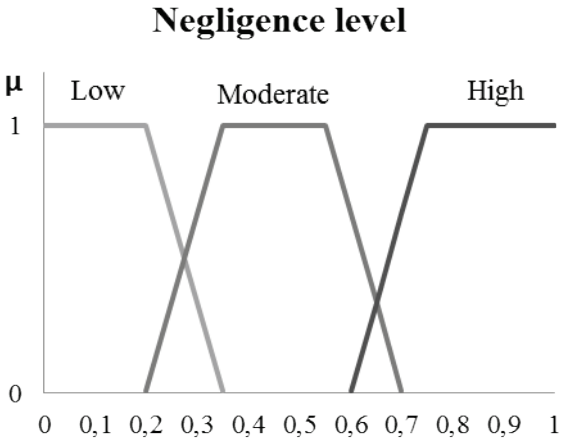

Figure 2: Negligence level fuzzy set sample

The password security also could be high (H), low (L) and moderate (M). We outline it in table on the next page.

Negligence \ Strength	weak	good	strong
low	M	H	H
moderate	L	M	H
high	L	L	M

Table 1: User account's break likelihood

On this table basis fuzzy rules could be composed as following:

```
IF negligence IS low AND password IS
strong, THEN password security IS high.
IF negligence IS low AND password IS good,
THEN password security IS high.
IF negligence IS low AND password IS weak,
THEN password security IS moderate.
IF negligence IS moderate AND password IS
strong, THEN password security IS high.
IF negligence IS moderate AND password IS
good, THEN password security IS moderate.
IF negligence IS moderate AND password IS
weak, THEN password security IS low.
IF negligence IS low AND password IS
strong, THEN password security IS moderate.
IF negligence IS low AND password IS good,
THEN password security IS low.
IF negligence IS low AND password IS weak,
THEN password security IS low.
```

We performed a baseline audit [13] before the password policy changes and two follow-up password audits in the course of implementation. The results are following: The base line audit cracked results were 91%, and after recommendations it is decreased to 57%.

To summarize, data could be retrieved from user in various ways, including directly asking for an exact value, using "checklists" and test. We think that in some cases the use of fuzzy sets as the input would be also efficient.

Size of the possible loss of an organization in case of materializing of particular threat may serve as a good example. In qualitative risk analysis, the impact in money equivalent is usually treated as low, moderate and high. It could be retrieved as user constructed fuzzy set, like *about 10000$* (i.e. more precise than those 3 levels, but less exact than numeric value).

Conclusion

Fuzzy logic methodology provides a way to characterize the imprecisely defined variables, define relationships between variables based on expert human knowledge and use them to compute results. Fuzzy expert system applied to information security field is sufficient technique for emulating specialist's decision-making ability.

Also, one of the advantages of fuzzy expert systems is that the rules can be written in natural language, rather than in computer jargon. As a consequence, communication between domain expert and knowledge engineer is greatly simplified.

In conclusion, we claim that there are enough unexplored areas and bright intersections in implementing expert systems in security auditing, development of fuzzy based knowledge base for expert systems, integration of fuzzy coefficients in development of recommendations for security auditing, and etc. Theoretical significance of researches above has been presented in authors' publications [14]. This paper is actually a part of whole scientific research, touched approaches and several issues of implementing fuzzy logic in problems of information security auditing and development of fuzzy expert systems. It is obvious that this kind of research directions could be a good scientific fundamental in artificial intelligence area.

References

1. Hinson, G. 2008. Frequently Avoided Questions about IT Auditing - http://www.isect.com/html/ca_faq.html
2. Val Thiagarajan, B.E. 2002. BS 7799 Audit Checklist. - www.sans.org/score/checklists/ISO_17799_checklist.pdf
3. ISO IEC 27002 2005 Information Security Audit Tool - http://www.praxiom.com/iso-17799-audit.htm
4. Stepanova, D., Parkin, S. and Moorsel, A. 2009. A knowledge Base For Justified Information Security Decision-Making. In 4th International Conference on Software and Data Technologies (ICSOFT 2009), 326–311.
5. Giarratano, J., and Riley, G. eds. 2002. Expert Systems: Principles and Programming. Reading, Mass.: PWS Publishing Company.
6. Tsudik, G. and Summers, R. 1990. AudES - an Expert System for Security Auditing. IBM Los Angeles Scientific Center.
7. Zadeh, L. 1994. Fuzzy Logic, Neural Networks, and Soft Computing.
8. Siler, W., Buckley, J. eds. 2005. Fuzzy Expert Systems and Fuzzy Reasoning. Reading, Mass.: Wiley-interscience.
9. Borjadziev, G., Borjadziev, M. eds. 1997. Fuzzy Logic for Business, Finance, and Management. Reading, Mass: World Scientific.
10. Mahant, N. 2004. Risk Assessment is Fuzzy Business—Fuzzy Logic Provides the Way to Assess Off-site Risk from Industrial Installations. Bechtel Corporation.
11. Elky, S. 2006. An Introduction to Information Security Risk Management. SANS Institute.
12. Bradner, S. 1997. Key words for use in RFCs to Indicate Requirement Levels. IETF RFC Repository. http://www.ietf.org/rfc/rfc2119.txt
13. Williams, M. 2003. Adventures in implementing a strong password policy. SANS Institute.
14. Atymtayeva, L., Akzhalova, A., Kozhakhmet, K., Naizabayeva, L. 2011. Development of Intelligent Systems for Information Security Auditing and Management: Review and Assumptions Analysis. In Proceedings of the 5th International Conference on Application of Information and Communication Technologies, Baku, Azerbaijan, pp.87-91.