ABOUT DEVELOPMENT OF KNOWLEDGE BASE OF NATIONAL DOMAIN NAME SYSTEM

Rena Gasimova

Institute of Information Technology of ANAS, Baku, Azerbaijan
depart1@iit.ab.az

Introduction. Internet is one of the most rapidly developing fields in Information Technologies (IT) field. Every computer connected to internet is identified with a unique code that contains a set of numbers. Based on its technical essence, this code compiles the IP-address of the computer. Due to complexity for users to operate by signifying their addresses by an IP- address of a computer and remember the numbers, created Domain Name System (DNS) is one of the most important components of internet infrastructure. Domain (domain derives from a Latin word dominium meaning to own) is a logical level of internet. Domain names consist of character fields divided with points [1].

The rapid development of IT has resulted in occurrence of a large number of problems in DNS field. There are juridical, technical, safety and other approaches for solution of the problem. The conducted researches regarding juridical problems of domain names, demonstrate that conflicts related to domain name occur upon infringement of trade mark and brand mark rights [2-5]. The majority of current scientific researches are performed in order to detect DNS safety problems [6- 9].

Although primary appropriation of DNS names in the Internet is the addressing attribute and carriage of trademark of firms, companies and other organizations, according to analyses, today it has become an instrument of unfair and unjust competition. As a result of researches conducted with this purpose it is determined that domain names that reflect geographical names, historical, cultural and other values belonging to Azerbaijan are registered by foreign citizens in different countries (For example, www.azerbaijan.tv, www.azer.info, www.baku.net, www.baku.su, www.nakhchivan.net, www.sumqait.net, www.aghdam.com and etc.)

There are several reasons of registration of domain names by foreign citizens. Firstly, there are almost no restrictions for purchasing and registration of domains with open country codes (*.com, ***.**info, *.org, *biz, *.net, etc). Therefore, the domains are created by using the names of authoritative people, historical-moral values, geographical names and etc. because of the information war (IM), in order to gain benefit and etc. (for example, www.baku.tv, www.karabakh.us, www.baku.ru and etc). The second reason is the gap in open registration of high level general domains (*.com, ***.**info, *.org, *biz, *.net and etc). Therefore, some people own the titles indentifying geographical locations, historical-cultural values and realities, and register them as domain names (For example, www.azerbaijan.com, www.karabakh.com, www.karabakh.info, www.lachin.com, www.sumgait.info and etc) for different purposes. This is a specific and important issue for not only Azerbaijan, but also for world countries. Consequently, compliance to statements about prevention of use of information resources and technologies for criminal and terroristic purposes, protection of human rights, and immunity of personal information and freedom of speech are the main problems arising in Internet environment.

Considering the problems noted above, today it is necessary to scientifically analyze information resources about domain names collected in DNS servers. During registration of domain names, registration data is located in DNS registry. The data located in the DNS registry is the main information source of registrants (owner of the domain name or registering person). Consequently, registration data of domain name added to DNS database contains all information about registrants. Conducted researches provide grounds that deficiencies in domain names field create necessity to obtain more accurate and detailed information (knowledge) than registration

data stored in DNS. The following can be indicated as prime deficiencies in domain names field:

– Intellectual analysis of registration data about domain names stored in DNS is not conducted;

– no methods are applied for automatic classification of domain name registrants in accordance with characteristics such as IP-address, electronic-mail (e-mail), address, name, surname, age, nationality, country, NS server, organization and etc.;

– absence of program software tools allowing accurate analysis of domain name registrants;

– deficiencies in regulation of domain names registration;

– deficiencies in transparency of domain names registration;

– violation of registration terms by domain name users;

– irresponsible registration and use of domain name;

– no common policy against domain name invaders (cyber squatting, fishing and etc.);

– discrepancy between domain name and content of site and etc.

This problem is characteristic for the entire internet network. It is necessary to take complex actions for their solution. Alongside with creating new opportunities for solution of issues in domain names field, the current level of IT development sets new requirements for solution methods of these issues. Scientific-technological development of society depends on system operation of information collection, storing and transferring. Collection of information is not sufficient; it's necessary to process and analyze information and obtain new knowledge by detecting the consistent patterns.

Considering the abovementioned deficiencies, one of the actual problems consist of development of analysis methods and algorithms for automatic classification of registration data, grouping and identification of domain name owners for increasing the number of domain names reflecting Azerbaijan's realities, national and moral values, their improvement, prevention of their leakage to foreign countries, for protection of domain names against cyber-squatters, regulation, administration, control, web-monitoring and etc. of domain market, processing.

The solution of these problems is reflected in "Electronic Azerbaijan" government program that consider development of IS in our Republic. In virtual space considered in the program, issues such as large application and protection of Azerbaijani language, alphabet, terminology, protection of national values (historical, territorial, cultural and etc), formation of national content and etc., justify the creation of national knowledge base of domain names system.

Main principles of creation of national domain names system. Rapid growth of number of domain names causes automatic accumulation of large volume information mass. DNS is the largest database (DB) that processes billions of requests on a daily basis [10]. DNS-servers unite activity and administration of domain names belonging to different zones. They possess unique information resources about domain names. It must be noted that, this information is regularly updated, administered and accessible. This data can be in different types.

In all domain zones, necessary information for domain registration (included in DNS registry) are as following: domain, registrant identifier(*registrar*), complete name of person (*person*), contact address *(address*), identifier of domain administrator *(admin-o),* name of organization (*organization*), domain registration date *(created*), refreshment date of domain (*updated*), free date of domain (free-date), telephone numbers with city and international codes (*phone)*, electronic mail (*e-mail*), list of DNS-servers supporting domain (*nserver*), state of object *(state*), type of domain *(type*), information source *(source*), payment date for domain registration (*paid-till)* and etc. Naturally, much of the useful information stored in different DNS-servers remain unused. It is necessary to collect and process this information. Different analysis methods (statistic, dynamic, interactive, multidimensional, data mining) can be used for

a complete and efficient use of information that is collected and prepared for analysis. For this purpose, 80 – 90% strategic data are obtained from the analysis of open sources [11].

Currently, there is no software enabling an accurate analysis of domain owners. In this direction, acquisition of required information from registration data stored in DNS requires processing and application of new methods. Existence of problems related to domain names belonging to Azerbaijan in virtual space, requires the creation of new mechanisms for their solution. It is necessary to form a systemized knowledge base of domains related to interests of the Republic of Azerbaijan and conduct relevant scientific-analytical analyses based on it. Domain names investigation, provision of analytical information about registrants and web-sites, performance of monitoring operations can be carried out by using registration data accumulated in DNS.

Currently, application of scientific methods and IT is very important to simplify the acquisition problems of new knowledge from large-volume information masses. But complicated object fields require individual approach, it is difficult to form (mathematically describe) them. Modern analysis and description methods of knowledge are used for solution of these problems. Considering these, it is advisable to prepare and develop special program complex of data mining systems and decision helping systems. Special knowledge base containing necessary information about domain names (object field) form the basis of special program complex.

Considering the sharp growth in the quantity of domain names, there is a large demand for these software products. It is possible to process thousands of registration data, analyze, and obtain new information and make the necessary decisions by detecting consistent patterns. Therefore, efficient use of information collected in DNS-servers is required for these purposes. Intellectual systems enable the collection of registration data about domain names belonging to different zones in DNS-servers in database, provision of ample opportunities for conduction of scientific analyses based on this information, and acquisition of operative knowledge about required resources. Therefore, information about domain names belonging to Azerbaijan from different DNS-servers should be stored in national domain names data base. The purpose is the creation of a general purpose intellectual national domain names system (INDNS), which efficiently interprets the registration data, performs qualitative searching, possesses vast capabilities and increases functional capabilities. Collection of national domain names to unique DB in INDNS and creation of knowledge base (KB) can provide the solution for these problems.

Creation issues of KB of national domain names form the basis of INDNS conception. The source of initial information is registration data about domain names stored in DB of DNS-servers. The purpose of creation of software for domain names analysis is the detection of knowledge concealed from the KB created on basis of registration data of domain names. Also creation of this analysis system can help to efficiently analyze the registration data. This feature of the system can solve any logical problems. The system can group domain owners by analyzing registration data collected in domain names and analyze the purpose of the registrant to purchase the domain name by analyzing its content.

On the other side, these systems can simplify the analysis of the state of domain names belonging to Azerbaijan, relations among domain owners, current state and tendency of domain market in the Internet. Today, it is important for detection of topical problems, forecasting and making correct decisions.

It is important to note that, knowledge about object field (domain name) is important for making decisions. Decision making should be complemented with scientific methods and extended with their assistance. The creation of INDNS is impossible without scientific - theoretic analyses. Considering the set forth tasks, different analysis methods can be used for complete and efficient use of data collected and prepared for analysis in the DB: data warehouse, online Analytical Processing - OLAP, data mining. These concepts are most efficient methods for conduction of analysis of the large volume of information resources. In

order to obtain and administer data, through intellectual analysis of data (IAD - Data Mining),it is possible to detect in unknown, nontrivial, practically useful knowledge from raw data required to make decision in different fields of human activity. Data Mining is a process of making a decision based on detection of hidden consistent patterns in data [12].

Conclusion. Rapid growth of Domain names, collection of registration data in DNS-servers require development of complete architectural principals of intellectual systems for detection of new knowledge. Requirements set forth to development of scientific-theoretical basics of INDNS, require the development of analysis systems based on intellectual methods. It must be noted that, intellectual analysis of registration data of domain names has not been conducted. It is possible to automatically classify the data in DNS-servers using the methods of knowledge discovery in databases (KDD). It is possible to obtain necessary results but using the methods of Data Mining (associative rules, decision tree, clustering, classification etc) for automatic detection of new consistent patterns of large volume of data.

References

1. А.Г. Серго. Доменные имена. М:. «Бестселлер», 2006, 368 с.
2. В. Кравченко. Домен - не адрес, а средство индивидуализации // Интеллектуальная собственность, 2001, № 2, с. 66-70.
3. B.G. Davis. The New Thing, Uniform Domain-Name Dispute-Resolution Policy of the Internet Corporation for Assigned Names and Numbers // The Journal of World Intellectual Property, vol. 3, no. 4, 2000, pp. 525-554.
4. D. Cabell. Foreign Domain Name Disputes 2000 // The Computer & Internet Lawyer, vol. 17, no. 10, 2000, pp. 5-17.
5. В.Б. Наумов. Споры, связанные с нарушением прав на объекты интеллектуальной собственности в сети Интернет // Арбитражные споры, 2001, №1(13), с. 81-88.
6. M. Daniel, E.D. Dorothy. Guest Editors' Introduction: Securing the Domain Name System // IEEE Security and Privacy, vol. 7, no. 5, 2009, pp. 11-13.
7. S. Edelman. Cybersquatting Claims Take Center Stage // Computer & Internet Lawyer, no. 1, 2001, p. 1-6.
8. O. Eric, Z. Lixia. Inter-Administrative Challenges in Managing DNSKEYs // IEEE Security and Privacy, 2009, pp. 44-51.
9. C. Ramaswamy, R. Scott. Open Issues in Secure DNS Deployment // IEEE Security and Privacy, vol. 7, no. 5, 2009, pp. 29-35.
10. А.А. Венедрюхин. Доменные войны. СПб.: Питер, 2009, 224 с.
11. R.M. Əliquliyev, R.T. Qasımova, İ.Y. Ələkbərova. Qərarların qəbul edilməsini dəstəkləyən müasir konsepsiyalar haqqında // AMEA-nın Xəbərləri, fizika-riyaziyyat və texnika elmləri seriyası, №2, 2005, s. 70-75.
12. R.M. Əliquliyev, R.T. Qasımova, İ.Y. Ələkbərova. Çoxölçülü analitik sistemlərdə mürəkkəb sorğuların realizasiyası məsələlərinin analizi / PCİ-2006 "Kibernetika və informatika problemləri", Beynəlxalq elmi konfransı, cild 2, 2006, s. 71-72.