Research of problems of computer networks expert systems

A L Zolkin¹, A N Losev², D V Gridina³ and T G Aygumov⁴

¹ Computer and Information Sciences Department, Povolzhskiy State University of Telecommunications and Informatics, L.Tolstogo Street 23, Samara 443010, Russia
² Department of Applied Informatics, Federal State Budgetary Educational Institution of Higher Education “Russian Timiryazev State Agrarian University”, Timiryazevskaya Street 49, Moscow 127550, Russia
³ Department of legal and customs regulation of foreign trade, Moscow Automobile and road construction State Technical University, 64 Leningradsky Prospekt, Moscow 125319, Russia
⁴ Department of computer engineering software and automated systems, Dagestan State Technical University, I. Shamyl Street 70, Makhachkala 367000, Russia

E-mail: alzolkin@list.ru

Abstract. The article is dedicated to the issues of development of expert systems in data-processing networks during web programming. The dependence of the processing, storage and transmission of information and web applications on the uniqueness and problematics of expert systems creating is observed. The process of development of a software control block of content (engine) as the own control system is considered in details; vulnerabilities and a set of necessary functions inherent in the specified engine are explored. Each new developed software engine can be exposed to unauthorized access and content theft, in this regard, sometimes it is worth considering the percentage of uniqueness of the development methodology and code. In this case, the control block (engine) must be written using its own unique variables and file names, which will help secure the site from third-party interference and will not require the purchase of more expensive hosting for its placement. In addition, the article investigates the problems associated with the threat posed by malicious programs in computer networks, such as direct DDOS attacks, identifies the features of such programs, and also provides ways to combat them.

1. Rationale
Currently, expert systems have reached a new level along with qualified specialists in the analysis and solution of complex issues and problems. Today they show very effective and high-quality results.
Expert systems include computer systems that have the ability to partially replace experts in problem situations resolution.

Expert systems are increasingly finding their use in the form of open-access online systems for various user groups. One of the most serious problems when building such systems is the complexity of design, since asynchronous interaction between the client and server sides is required. The difficulty lies in structuring and extraction of information, as well as in its storage. And the first issue is selection of place for information storage [1,2]. Of course, the expert system does not store confidential data, but it would be inappropriate to place the data on the client side every time due to the large amount of information.
The aim of the study is to analyze the method for solving the established problem, based on the complete set of a special package of tools that consists of standard development tools in the field of WEB technologies which have maximum performance, convenience and speed of development, as well as some specificity that makes it possible to solve the problem in non-traditional ways.

2. Comparative analysis of WEB-programming server technology

In order to identify the necessary qualities in development tools, a comparative analysis of the server technology of WEB-programming should be carried out. To do this, the following programming languages are taken:

- PHP.
- Perl.
- Phyton.
- Ruby.
- JavaScript.

The JavaScript language is classically considered to be client-side and interpreted strictly in the client's browser, but due to its specificity, the Node.JS framework was developed, which is a software platform based on the V8 engine (transfer of JavaScript into machine code), which turns JavaScript from a highly specialized language into a language general purpose [3].

Node.JS provides JavaScript with ability to interact with I / O devices through its API (written in C++), connect other external libraries written in different languages, providing refer to them from JavaScript code.

Node.JS is used mainly on the server, performing the function of a web server, while it is possible to develop desktop applications using Node.JS. Node.JS is more productive than other server languages.

Comparative analysis will be carried out for the following properties:

- Paradigms.
- Type designs.
- Memory management.
- Computing flow control.
- Functionality.
- Compiler / interpreter.

In terms of developer productivity, all languages have different parameters, however, based on the performance and consumption of computing resources, one optimal programming language can be identified. It is JavaScript (Node.JS), which is not inferior in development productivity to other languages.

Another reason why Node.JS is considered the most relevant technology to choose when developing ES is the fact that the server side will have the same language as the client, which will significantly affect the convenience of writing code and client interaction with the server.

Also, during use of Node.JS, the data transfer process is greatly simplified. This can also be called the advantage of server JavaScript, which makes it possible for user to write the own WEB server [4,5].

Considering the client side, the attention shall be paid to the fact that currently a number of browsers are rapidly developing the new HTML5 standard, which opens up new opportunities for the developer to work with the client side, in particular, the organization of a database based on the SQLite3 DBMS on the client side.

This approach makes it possible to provide logical caching and reduce traffic between the client and the server, thereby increasing the speed of the ES.

It shall be remembered that speaking of server-side languages, the user must request entry to the web page, the corresponding transition to the specified page will be made using a specific link in the
existing browser.

Further processing will be carried out on the server to execute software applications that have a connection with the page. After these actions, an answer will be given to the user of the computer network in the form of a file with various extensions, including Perl, SI, XHTML, HTML, etc.

The performance of software applications will already have a certain dependence on the server where the site itself will be located, as well as on the version of the programming language that supports it. A feature of the functioning of server programming languages is their ability to organize contacts with a database management system, a DBMS server.

The PHP programming language is open source and, in a number of its scripts, has the ability to integrate into conventionally ordinary html documents. However, the web server must have a PHP interpreter [6].

The specified programming language cannot be used for writing of system programs or their components. Web applications written in this language have difficulties with security, while there is no possibility of asynchronous and multi-flow work.

The Ruby programming language has in itself the ability to object-oriented functionality, independent of the operating system in terms of multithreading and RAM management. However, this language is especially difficult to learn, it is not very productive, it has little development.

The Python programming language is aimed at solving complex issues of a wide field in web development. This language generates a set of statistical html-pages maintained by servers. Python is very limited in the available tools for working with a DBMS, its functionality is not perfect enough, although it has a very good syntax, which makes the code readable.

3. Application of WEB-programming in the development of a software control unit

If the author's project is successful, the site will inevitably grow to several dozen pages. At the same time, for a multi-page site, the most standard techniques for a site maintenance (for example, the need to change the design or insert a new block, counter or banner at the same time on all pages of the site) lead to workout of all pages (including changing of code on each separate page). It is extremely ineffective. Also it is associated with significant time costs. At this moment, the idea of simplifying its maintenance is maturing.

Of course, the situation is not hopeless. To do this, first of all, we shall additionally turn to theory, namely to the PHP programming language, which will make it possible, in turn, to move to the next level - the creation of dynamic sites in PHP [7,8]. It is noteworthy that a special template developed in the PHP environment will significantly simplify not only the creation process itself, but also the support of multi-page sites. It will also be possible to easily change the design, add new pages, insert additional blocks, etc.

However, in this case, one slight inconvenience arises, namely, editing all pages in the "native" environment and then uploading them via ftp to the server. But this nuance is easily compensated by the advantages that are available for developer now.

In addition, the PHP environment provides the ability to use an MSQL database, create an administrative panel, with the ability to edit content and manage various add-ons and applications previously created by web pages, despite the presence of only the page content output line.

Further, in the process of automating the procedures for creating sites, the stage of developing and using the so-called "engine" follows, which is a software complex that manages content (software control block), with the help of which development, support, and management of absolutely all processes are implemented within the framework of the content.

It should be emphasized that the content management system of the site - CMS (site engine) is a special script application, or software for semi-automatic site content management [9]. It also includes the execution of tasks for processing requests, sending the generated website pages to the browser. The use of an information base is considered as the most successful option during use of the engine; otherwise, it is possible to use individual files, but there is no guarantee of increased load on server storage. Today CMS is the most popular system or method of website development, taking into account
the presence of some site engines with free add-ons.

In addition, a variety of audio and video materials are available to facilitate the process of learning to work with any CMS environment.

Today the most popular and widespread engine is wordpress. Most of the blogs and websites of the Russian Internet have been created using this engine. As a result, this has led to the fact that many sites have codes with the same files and variables and entire segments of code blocks, which, in turn, represents a wide field of activity for a competent cracker.

The creation of the own engine, control unit, software controller is explained by the desire of the owner of the future site for its uniqueness and ease of use, for the presence of an optimal interface, for an emphasis on information security, original aesthetics and design, which will contribute to giving the site a certain exclusivity [10,11].

The course of study of the technology of development of the own engine can be divided into several stages. At the first stage, the HTML / CSS environment is mastered, the second stage is PHP, which is devoted to mastering the development of dynamic sites, at the third stage it is necessary to deal with the tasks of creating and managing a website through the admin panels and learn the MSQQL language. At the penultimate stage, it will not be superfluous to replenish the knowledge by learning the language of interactive technologies JavaScript, which will undoubtedly come in handy in the future. And, finally, at the final stage, without making any special efforts, to write the own CMS and develop own sites based on it. It makes scenes to understand the structure of the engine in more details, since in the absence of knowledge about its structure, it is almost impossible to understand the principle of the engine's operation. Database: services of using MSQQL databases are available on all paid hosting services, which, as a rule, contain all the pages of the site and all additional information.

Control panel: this section, or module, directly solves problems that entail the maximum simplification of the maintenance of the entire resource as a whole. It is possible to change the design from the admin panel, as well as add or remove pages, edit, i.e. do whatever is necessary if such a function is programmed [12]. And this is not difficult to implement, it is enough to carry out the preliminary study of the necessary information. Any free project can serve as an editor for the admin panel, for example, the ckeditor can be used. The site engine itself is CMS, or management system (optional). This module manages the display of pages, all navigation and implements all the additional functions that are assigned to it. The essence of its work is as follows: using the .htaccess file, all requests to the site (without exception) are redirected to the main php script that directly generates the page upon request, taking into account the type of request: GET or POST.htaccess (hypertextaccess) is an additional configuration file for the Apache web server and similar servers. In this situation, it becomes possible to use a variety of additional configurations and updates for the efficient operation of the web server in the appropriate folders. In this case, these directories may not change key configurations, but control of access to directories and replacement of file typology will be carried out.

Htaccess is analogous to httpd.conf, the only difference is in the action exclusively on the directory in which it is located and on its child directories. The htaccess file can be placed in any directory. The directives of this file affect all files in the current directory and in all its subdirectories (unless these directives are overridden by the directives of the underlying .htaccess files).

In order to be able to use htaccess files, the main configuration file shall be appropriately set. As a rule, the vast majority of hosters allow the use of their own htaccess files.

Paths to files and directories shall be specified from the server root, for example, /var/www/domain.com/htdocs/. It does this simply by parsing the URL into its constituent parts, outputting the necessary information from the database. If the requested page is missing, the script displays the error text.

The template engine, receives data from the engine, displays it on the screen according to the design created for the HTML template. All the information in the database regarding the address typed in the browser is specific for each page and the number of times corresponds to the number of simultaneous or sequential requests.

The work process itself is not too complicated and includes two-way exchanges between the database
and the control panel, where manipulations with the site content take place. Next, the engine extracts information from the database and redirects it to the template engine, which redirects the already processed code to the browser, which shows the user what he sees on the screen.

The user has the ability to directly communicate with the engine using the site navigation, as well as by adding a comment. Here, the communication process goes a little bit different. The engine extracts information about the commenting system from the database and forces the templating engine to display an HTML form to the user where he can leave a comment [13]. This information first goes to the engine, which in its turn processes it, following a given algorithm. Than, the engine cleans and edits the comment code, and then sends it for storage to the database and the template engine, which loads it into the user's browser.

The CMS includes many add-on modules that provide the necessary management functions required by the user. The latter does not need to have programming skills to work with a CMS, but at the same time he needs to delve into the management environment itself in order to create the site authorly.

4. Research of problems associated with the threat of malicious software in computer networks

At the present stage of development of the market for information protection services, the problem of protecting a WEB server from so-called DDOS attacks is becoming increasingly important. DDOS means distributed denial of service. The purpose of this attack method is denial of the attacked computer from the normal operating state. As a rule, this type of attack is used as a tool for blackmail, liquidation of competitors, or as a fun activity for novice hackers. A DDOS attack today poses a large-scale threat to the World Wide Web and has already acquired the status of "cyber terrorism".

The reason for everything that happens is the easy accessibility of this type of attack, since in order to achieve the set goal, the attacker must redirect traffic from many infected computers to the attacked server. This is not so difficult to implement, user can limit himself to writing a script using the JavaScript language that downloads a WEB page from the attacked server as many times as he likes, and place this script on the visited page of some WEB site.

If this resource is absent, then the attacker will very easily make a fake site and carry out spam mailing containing a link to a page with malicious code. As a result, due to the congestion of users, a lot of requests will be created, from which the server is simply overloaded when trying to process them.

The above example demonstrates the ability to take out sites with an average level of processing power. To attack the most powerful servers (like it was with site of the Echo Moskvy radio station on May 4, 2012) the attackers use the so-called BOT-NET, which is a network of infected computers [14]. This attack used three zombie networks, a total of three hundred and fifteen thousand infected computers.

However, the usual increase in the processing power of the server to increase the security of the server in a large-scale DDOS attack will not be enough, since the server's Internet channel is often not enough to pass all requests coming to the server. As a result, ordinary users suffer. They are simply idle in the queue, waiting for the channel to become free and access to the server is obtained.

The issue of solving of this problem is touched by many scientists around the world. Hundreds of companies produce a variety of software and hardware protection methods. System administrators are constantly developing updated algorithms that can filter the traffic entering the server.

The main methods for solving the indicated problem that are used at the present stage will be considered below. A new method of protection of a server from large-scale attacks has been developed by American scientists from the University of Washington K. Dixon and T. Anderson. Their authorial method is based on creation of a cloud from several thousand computers configured to receive redirected filtered requests from infected machines, and the server, in its' turn, will process requests coming only from legitimate users. The specified complex was named Phalanx.

The position of the scientists is that the developed technology makes it possible to deter attacks of almost any force, it is enough only to increase the number of computers in a given cloud. Also, this complex can slow down the work of BOT-NET, since every computer trying to access the server has to solve a simple task, and this entails a significant slowdown in the work of BOT-NET.
System administrators and information security specialists provide protection from DDOS attacks, both at the hardware and software levels of access to the server. Analyzing the software level of protection, we will give an example of a study conducted by Ivan Ivanovich Slepovichev. The researcher proposed a method for solving the problem, implemented at the software level, using the mathematical apparatus of fuzzy logic and neural networks. The aim of the work was to create a program that allows detecting DDOS attacks in identifying illegitimate users. But the main effect of using this program is that it can learn and suppress any new methods of influencing the server by the attacker, since the program is able to replenish its knowledge base to identify new algorithms for generating a request by the attacker to present them to the system as requests from legitimate users. This method is efficient, but its implementation requires more powerful computing resources than the standard query filtering technique. As a result, it is unprofitable for ordinary non-commercial sites, as well as for small business sites, to use this method, since the cost of purchasing hardware to increase the resource of the computing power of the site is generally quite high. Summarizing the analytical information on the presented modern and most effective methods of protection against "distributed denial of service", we can make one general conclusion that there is no absolute solution for all types (classes) of projects for this type of attack. The main reason lies in the material component of all methods. The effectiveness of the protection technique directly depends on the amount of material costs by the owner of the resource. Consequently, the risk of a threat to low-budget sites is extremely high, and the creation of new, more effective protection methods is based only on increasing the material costs for their implementation.

Topical unexamined issues in DDOS attacks. Distributed denial of service attacks are becoming more and more cumbersome, and the bandwidth provided by ISPs is changing extremely slowly. The security policy of providers at the present stage, when the channel width is exceeded, leads to a complete blocking of the client and its complete disconnection from the network. In such a situation, providers attribute this to insufficient computing power resources to work with large amounts of incoming traffic. In this case, it makes sense to block the client until the received requests (traffic) decrease up to not more than 70-80% of the channel width provided to the client. This security policy is fully justified on the part of the provider, since there is no need for a huge amount of computing power to analyze requests coming from the network. And taking into account that the width of the incoming traffic during an attack is higher than the width of the server's channel, then, therefore, a certain mass of requests will remain on the side of the provider, waiting for their turn, occupying the resource capacity of its computing power and endangering all other clients. However, as applied to a WEB resource (site), this technique is destructive, since the computing power of the server, taking into account the software filtering of requests, is often sufficient to process requests received through communication channels and provided by its provider, and the channel width at this moment is fully loaded.

5. Findings
Thus, it shall be noted that for a successful and convenient design and writing of an expert system using a set of WEB programming languages, it is still optimal to use JavaScript on both the client and server sides. The specified language can only be limited to browser assignments, while it is suitable for solving complex tasks in different situations. Also this language has such qualities as simplicity and abstraction. In this case the browser for viewing of web pages executes orders of client applications in the computer network on the user's computer while the server applications are launched on the server.

The use of a publicly available engine is fraught with the fact that someone can take advantage of such public availability and find another weak link, with the help of which unauthorized access is possible with the subsequent management of the content of the site.

It must be taken into account and realised that the way out of this situation can only be the development of your own content management system. Such a system will be able to perform only the necessary set of functions, it will be written using its own unique variables and file names, which, in turn, although not completely, but to a greater extent will be able to secure the site from outside interference, and will not require the purchase of more expensive hosting for its placement.
Thus, the role of information protection at the provider level kills the whole essence of the protection technique at the WEB server level, leading to the server being idle in offline mode for its legitimate users.

Summarizing the above, it can be concluded that in order to protect WEB servers from “distributed denial of service” attacks, it is necessary to build not only a policy of analysis, filtering and additional computing power, but also providers shall develop a new policy for providing the Internet access. In addition, it is also necessary to create a new method for increasing the width of the communication channel provided to the WEB server by its provider.

References
[1] Ilin S Protection of USB flash-drives against viruses (In Rus.) http://www.windxp.com.ru/articles61.htm (Accessed: 22.02.2020)
[2] Viruses elimination Kaspersky lab fan-club forum (In Rus.) http://forum.kasperskyclub.ru/index.php?showtopic=8920 (Accessed: 22.02.2020)
[3] Bereza N V 2012 Current trends in the development of the global and Russian markets of information services *Inzhenernyi vestnik Dona* 2 URL: vdon.ru/ru/magazine/archive/n2y2012/758/ (accessed: 22.12.2019)
[4] Vasilenko K A, Zolkin A L, Abramov N V and Kurganov D O 2020 Enterprise data processing networks: authentication and identification procedures *Infocommunication technologies* 18(1) (Samara: Povolzhskiy State University of Telecommunications and Informatics) p 61-7
[5] Vasilenko K A, Zolkin A L, Abramov N V and Kurganov D O 2020 Features of flash drives information security *Infocommunication technologies* 18(2) (Samara: Povolzhskiy State University of Telecommunications and Informatics) p 200-6
[6] Vasilenko K A, Zolkin A L, Abramov N V and Kurganov D O 2020 Malicious software in computer networks *International Market Institute periodical* 2 (Samara: International market institute) p 108-11
[7] Morozova T, Akhmadeev R, Lehoux L, Yumashev A, Meshkova G and Lukiyanova M 2020 Crypto asset assessment models in financial reporting content typologies *Entrepreneurship and Sustainability Issues* 7(3) 2196-212 doi: 10.9770/jesi.2020.7.3(49)
[8] Yumashev A V, Koneva E S, Borodina M A, Lipson D U and Nedosugova A B 2019 Electronic apps in assessing risk and monitoring of patients with arterial hypertension *Prensa Medica Argentina* 105(4) 235-45
[9] Ling V V and Yumashev A V 2018 Estimation of worker encouragement system at industrial enterprise *Espacios* 39(28)
[10] Riakhovskii A N, Zheltov S I, Kniaz’ V A and Iumashchik A 2000 A hardware and software complex for producing 3D models of the teeth *Hardware and software complex for obtaining 3D models of teeth dentistry* 79(3) 41-5
[11] Dzhangarov A I, Suleymanova M A and Zolkin A L 2020 Face recognition methods *IOP Conference Series: Materials Science and Engineering. Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations* (Krasnoyarsk, Russia) 42046 DOI: org/10.1088/1757-899X/862/4/042046
[12] Tormozov V S, Zolkin A L and Vasilenko K A 2020 Setting, training and testing of the long-term memory neural network for pattern identification task *Industrial ACS and controllers* 3 (in Rus.) (M.: Scientific & Technical Literature Publishing House “Nauchtehlitizdat”) 52-7
[13] Faizullin R V and Hering Sh 2019 The model of data aggregation from clustered devices in the internet of things *Intelligent systems in production* 17(4) 156-62
[14] Zhukovskyy V, Zhukovska N, Vlasyuk A and Safonyk A 2019 Method of forensic analysis for compromising carrier-lock algorithm on 3G modem firmware *IEEE 2nd Ukraine Conference on Electrical and Computer Engineering, UKRCON 2019* – Proceedings 1179-82 8879941