Ensuring Information Security in the Intelligent Scientific and Technical Information Systems

Komil Kerimov
Department of System and Applied Programming
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
kamil@kerimov.uz

Zarina Azizova
Department of Information Security
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
z.i.azizova18@gmail.com

Fayzi Bekkamov
Department of Library information systems
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
bfayzi93@gmail.com

Mekhriddin Rakhimov
Department of Computer Systems
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
raximov022@gmail.com

Mannon Ochilov
Department of Computer Systems
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
ochilov.mannon@mail.ru

Abstract. **Scientific and scientific-technical information is a valuable tool for the development of education, technology and society as a whole. The increase in the volume of information and the development of information networks of data exchange requires special means to ensure information protection of data. Methods, means and systems for information security of scientific, technical and scientific-educational resources in intellectual information systems are of particular importance. The purpose of the research is to develop methods and software tools to ensure information security of valuable scientific and technical information resources in intelligent information systems.**

The proposed solution for intrusion detection in intelligent system is a web application firewall, which is used for enhanced security, detecting and preventing attacks before they reach the web application. It will protect the system from a whole range of attacks while allowing HTTP traffic monitoring and analyzing small changes or persistent state online. The Web Application Firewall (WAF) has the following features: logging of all HTTP protocol transactions, including request termination permissions and logging of the response; HTTP traffic can be examined in real time to detect attacks; preventing attacks before they reach the web application.

This work is performed within the framework of the project on creation of an integrated intelligent system “SMART TUIT”, which includes several subsystems (Information Retrieval, Voice Recognition, Pattern Recognition, Scientific Information Assessment, Geoinformation System).

Keywords: information security, intelligent system, scientific and technical information, information retrieval, behavioural analysis.

I. The Credibility of Scientific and Technical Information Sources

In recent times, significant scientific and technical discoveries stemming from diverse research endeavors have become increasingly valuable. Simultaneously, this body of knowledge plays a pivotal role as a primary resource for conducting further research. Hence, the ability to fully utilize such information during research endeavors is paramount. However, it is important to note that not all scientific and technical resources are readily accessible, and their utilization is governed by the owners or publishers of these resources. Presently, a plethora of international publishers and scientific databases, including Web of Science, Scopus, Ebsco, Springer, and ProQuest, alongside libraries, scientific laboratories, and research centers, offer access to these invaluable resources.

Various methods and tools have been proposed in previous studies for evaluating information, considering factors like accuracy, completeness, reliability, compatibility, usability, objectivity, and novelty. However, the complexity arises from factors such as the user is knowledge level, the research field, the information

Print ISSN 1691-5402
Online ISSN 2256-070X
https://doi.org/10.17770/etr2024vol2.8042

acquisition purpose, and the continuous evolution of science, leading to new information and changing the value of existing information over time. The evaluation of scientific and technical information sources is influenced by the evaluator is goals, interests, knowledge level, and timing of evaluation. Although evaluation often becomes subjective, there is a need for occasional objective assessments regarding the information is significance in advancing society and science [1].

In an era marked by the exponential growth of digital data, numerous search engines have emerged to cater to users' information requirements. These search systems are designed for various purposes, including accessing electronic libraries, archives, and educational databases, particularly for scientific and technical information retrieval. However, the majority of these search engines operate by retrieving information based on user queries. Consequently, it is crucial for scientific and technical database search systems to analyze user information needs before presenting relevant resources. Addressing this challenge, recommendation systems offer a viable solution, leveraging a knowledge base and machine learning algorithms. They have become integral to artificial intelligence and are widely employed across different organizations and services. For instance, Amazon has implemented a personalized recommendation system to suggest e-books to its customers effectively [2].

Research has commenced on implementing artificial intelligence within library systems [3]. This involves utilizing artificial intelligence in scientific and technical databases to analyze user data such as age, interests, expertise in specific fields, and past queries. By doing so, the system can suggest relevant information sources tailored to the user is needs, ultimately streamlining the librarian is tasks, enhancing the speed and precision of information retrieval, and improving overall management efficiency while better meeting the populace's informational requirements.

II. Security of the Integrated Intelligent System “SMART TUIT”

In integrated intelligent systems, user data stored in the database, such as personal data, their requests for information searches, preferences, search history and other types of confidential information may use by attackers implementing attacks such as URL interpretation, data entry validation attacks, SQL-injections and others for the purpose of stealing personal information, fraud or violating confidentiality.

Among the scientific works devoted to the protection of information systems from security threats the following ones: a comparative study of various web-application vulnerability tools and scanners to determine their effectiveness and accuracy in vulnerability detection [4], rationalization of security measures against SQL injections [5], increasing the security level of web-applications by using firewalls [6], study of various filtering techniques to prevent SQL injection attacks [7], and others.

Our solution for intrusion detection in an integrated intelligent system, “SMART TUIT”, is a web application firewall that used for enhanced security, detecting and preventing attacks before they reach the web application. It will protect your system from a whole range of attacks, while allowing you to monitor HTTP traffic and analyze small changes or persistent state online. The Web Application Firewall (WAF) has the following features:

- Logging of all HTTP protocol transactions, including permissions to terminate the request and log the response,
- HTTP traffic can be inspected in real time to detect attacks,
- Prevention of the attacks before they reach the web application.

A web application firewall can be part of a web server structure or a reverse proxy server on a network. Figure 1 illustrates detailed scheme of web-application firewall functionality. To create an effective protective tool, sets of rules have been developed that detect HTTP protocol violations, detect typical attacks on web application security, protect against automated activity (bots, worms, scanners and other malicious programs), detect access to Trojan horses, and distort error messages sent to the server.

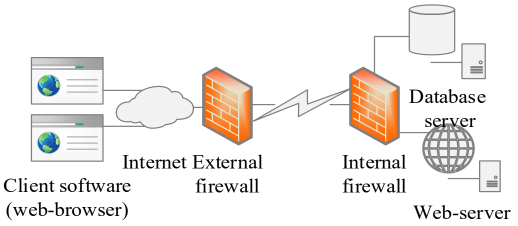


Fig. 1. Scheme of functioning of the web application firewall.

Consider in more detail the architecture of the proposed web application firewall shown in Figure 2.

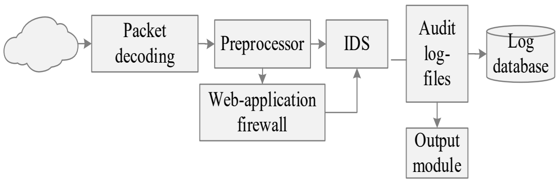


Fig. 2. Architecture of the proposed web application firewall.

a) Packet Decoding. The library used to capture packets transmitted over the network that contain the capture time, packet length, and link type (e.g., Ethernet, FDDI). A pointer created to point to each packet for performance analysis. The firewall in built-in mode has additional features such as packet transmission, packet modification, specific packet return, and packet deletion.

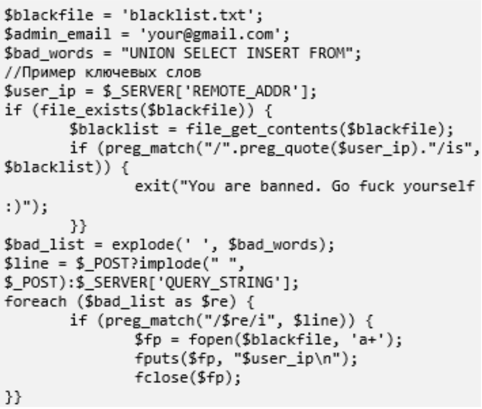
b) Preprocessor. After packets captured, they passed to the preprocessor for packet extraction and normalization using the formats of each protocol. The preprocessor also analyzes statistics of network traffic usage and identifies unwanted attacks such as worms.

c) IDS (Intrusion Detection System) is the core of the proposed system. With proper settings to detect network attacks effectively. If the captured packet contains any signature pattern, the system will alert about the attack and write the data to the log files.

d) Audit log-files. When the system recognizes attacks, log-file generated and message displayed that contains attack-related information for the administrator to remove the attack.

e) Output module. A module that allows to output errors in the desired format.

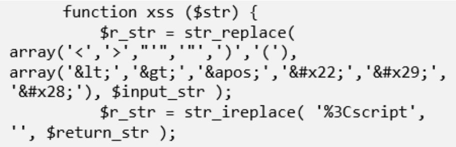
As we research XSS, SQL injection vulnerabilities, we wrote filters to protect against them. The following code used to prevent SQL injection attacks:



```
$blackfile = 'blacklist.txt';
$admin_email = 'your@gmail.com';
$bad_words = "UNION SELECT INSERT FROM";
//Пример ключевых слов
$user_ip = $_SERVER['REMOTE_ADDR'];
if (file_exists($blackfile)) {
       $blacklist = file_get_contents($blackfile);
       if (preg_match("/".preg_quote($user_ip)."/is",
$blacklist)) {
              exit("You are banned. Go fuck yourself
:)");
       }}
$bad_list = explode(' ', $bad_words);
$line = $_POST?implode(" ",
$_POST):$_SERVER['QUERY_STRING'];
foreach ($bad_list as $re) {
       if (preg_match("/$re/i", $line)) {
              $fp = fopen($blackfile, 'a+');
              fputs($fp, "$user_ip\n");
              fclose($fp);
}}
```

Fig. 3. Filtering rule to prevent SQL injection

The following code used to prevent XSS attacks:



```
       function xss ($str) {
            $r_str = str_replace(
array('<','>',"'",'"',')','('),
array('&lt;','&gt;','&apos;','&#x22;','&#x29;',
'&#x28;'), $input_str );
            $r_str = str_ireplace( '%3Cscript',
'', $return_str );
```

Fig. 4. Filtering rule to prevent XSS attacks.

In this section of the article, we presented an intrusion detection mechanism using web application firewall. The Web application firewall reduces the level of activity or blocks common attacks by using filters with the right set of rules. It should be noted, however, the need for the administrator to filter all incoming data from users.

III. PATTERN RECOGNITION SUBSYSTEM IN "SMART TUIT"

Recognition technologies applied in accordance with the specifics of the systems in which they are used. Among the most relevant today are bar codes, fingerprints, NFC and face recognition. Digital image processing allows devices to recognize an object, make a decision regarding the distribution and access control of the object in question, and act in accordance with the set access control rules in the system.

This allowed using the capabilities of computers in different spheres of activity, as devices today can see objects in the same representation as people do [8]. The development of deep learning technology has allowed bringing to a new level the process of image detection, by convolutional neural network (CNN) and R-CNN. The authors of [9] investigated the learning approaches for face recognition in integrated information systems, in addition, different neural network architectures such as convolutional neural networks and recurrent neural networks (RNN) and their applications are discussed. In turn, in [10], the authors present a hybrid computing platform based on a graphics processing unit (GPU) and a programmable gate array (FPGA) for processing machine learning tasks. They investigate the application of this platform in the context of pattern detection and recognition. The combination of the power of the GPU and the flexibility of the FPGA enables high performance and efficiency in complex pattern recognition tasks.

This section of the paper presents a model of image recognition, namely the faces of users of the integrated intelligent system "SMART TUIT". The integration of the pattern recognition module in "SMART TUIT" led to the automation of the process of image capture, its analysis, subsequent search, as well as the definition of tasks. The CNN detector used in the module is a method of recognizing facial images of the system users. The module consists of four main parts: image (face) formation module, training module, camera module, and image identity verification module.

The implementation of the pattern recognition module consists of initially recording information about each user (library visitor), then initiating the training process (Figure 5 illustrates the process of creating and training the neural network) and subsequent testing with basic functions in the CNN library.

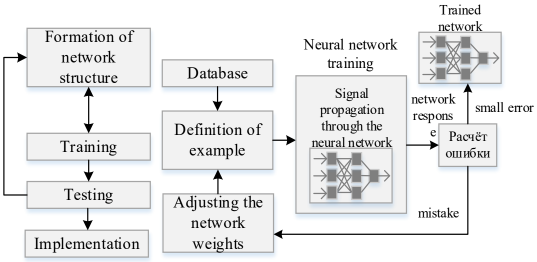
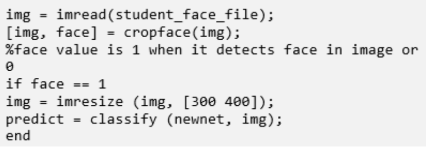


Fig. 5. Architecture of the proposed web application firewall.

Storage of recognition objects provides the primary process of recording the faces of the system users. All these objects are stored in a single library. The greater the number of images stored in the database, the more accurately the image recognition module in the "SMART TUIT" system will function. The next step is the learning process represented by the algorithm associated with the data stored in the library. The code compiled in the training module provides step-by-step instructions on the processes performed in the system.

One part of the presented image retrieval logic is shown in Figure 6.



```
img = imread(student_face_file);
[img, face] = cropface(img);
%face value is 1 when it detects face in image or
0
if face == 1
img = imresize (img, [300 400]);
predict = classify (newnet, img);
end
```

Fig. 6. Program code of the logic of functioning of the pattern recognition software module.

First, the image is stored in memory, image identification based on faces by cropping the desired object from the total image. Next, the camera module recognizes the image and compares the available image with the images stored in the database according to the previously trained sample. The main process of functioning of this module (Figure 7) is to recognize an object based on the shape of the face, after which the image captured for the purpose of subsequent comparison and determination of the identity of the images.

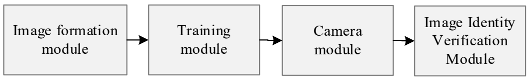


Fig. 7. Interrelation of pattern recognition modules in the "SMART TUIT" system.

The architecture of the image recognition module in the “SMART TUIT” system (shown in Figure 8) includes a logical level responsible for data maintenance and management, as well as a storage level including databases with visit information, a knowledge base and a database with user information. The interaction of these components allows the system to effectively recognize user facial images, manage information, and provide security and access control.

A. Logical level:

- Image Recognition.

At this level, the main process of recognizing user face images takes place. The image captured by the camera is stored in memory and is subjected to identification by cropping the desired object from the overall picture. The camera module then recognizes this image and compares it with the images stored in the database. The basic process of this module is to recognize an object based on the shape of the face and then capture the image to determine its identity.

- Data management.

This component is responsible for managing all necessary information in the system. This level contains data on users, staff, request management and statistics. Information about each user, visitor of the library, is recorded and stored in a database. The knowledge base contains information about the processes and rules of the system, and the user information database contains additional data about users, such as their personal data, visit history and other useful information.

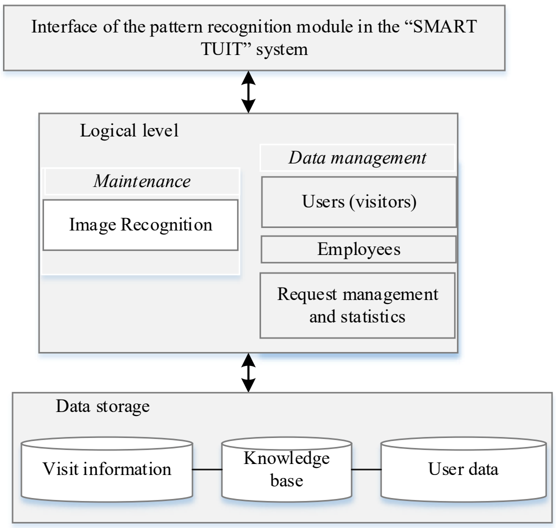


Fig. 8. Architecture of the pattern recognition module in the "SMART TUIT" system.

Each component of the architecture will be discussed in more detail:

B. Storage level:

- Attendance database.

This database contains information about each visit by users of the “SMART TUIT” system. It stores records of the date and time of the visit, user ID and other related data. This information can be used for analyzing and statistics of visits, as well as for security and access control purposes.

- Knowledge base.

The knowledge base contains information about the processes and rules of the “SMART TUIT” system. It includes instructions and manuals for using the system, pattern recognition algorithms, and other relevant knowledge necessary for the correct functioning of the pattern recognition module.

- Database with user information.

This database stores additional data about users of the “SMART TUIT” system. This may include personal data such as names, photos, contact information, as well as visit history, preferences, and other useful information that can be used to personalize and better serve users.

The architecture of the image recognition module in the “SMART TUIT” system includes a logical level responsible for data maintenance and management, as well as a storage level including databases with visit information, a knowledge base and a database with user information. The interaction of these components allows the system to effectively recognize user facial images, manage information, and provide security and access control.

Presented a pattern recognition model based on user face recognition in the integrated intelligent system “SMART TUIT”. The use of pattern recognition technologies, such as CNN and R-CNN, combined with digital image processing capabilities allowed to automate the process of capturing, analyzing and searching user faces, as well as determining the tasks. The process of implementing the module includes recording information about each user, training the neural network and testing the basic functions. One of the important components of the module is the storage of recognition objects in the database. The more images will be stored in the database, the more accurately will function the module of pattern recognition in the system “SMART TUIT”. The use of the module in the “SMART TUIT” system provides automation of the processes of image capture, analysis and search, as well as reliability and security in the access control system.

IV. CONCLUTION

In our article, we proposed the concept of protecting the integrated intelligent system “SMART TUIT”. One of the solutions we presented is a web application-based approach with a Web Application Firewall functionality, which enhances security by detecting and preventing attacks before they reach the web application. The primary focus is on using input data filtering to protect against common attacks such as SQL injections and cross-site scripting (XSS).

The image recognition subsystem in “SMART TUIT” is based on the application of CNN. The system automates the process of capturing, analyzing, and searching user faces. The facial recognition module consists of several key components: the image formation module, training module, camera module, and image authentication module. It enables automation of the processes of capturing, analyzing, and searching user faces, providing reliability, security, and access control in the system.

The architecture of the “SMART TUIT” system includes not only the intrusion detection module and image recognition module but also other components such as a knowledge base, user information database, and visit information database. The interaction among these components allows for effective information management, security, and access control. Overall, the scientific research and development in the field of intrusion detection and image recognition in the integrated intelligent system "SMART TUIT," along with the proposed solutions, contribute to enhancing system security and efficiency.

REFERENCES

[1] M. Rakhmatullaev, Sh. Normatov and F. Bekkamov, Fuzzy model for determining the information needs of library users, 14th international scientific and practical conference: Environment. Technology. Resources, June 15-16, 2023, Rezekne. Rezekne Academy Of Technologies.

[2] M. Bendechache, N. Limaye and R. Brennan, "Towards an Automatic Data Value Analysis Method for Relational Databases,", presented at the 22nd International Conference on Enterprise Information Systems, 2020.

[3] I. M. Omame and J. C. Alex-Nmecha, "Artificial Intelligence in Libraries," presented at the 22nd International Conference on Enterprise Information Systems, vol.2: ICEIS, 2020, pp. 833-840.

[4] J. J. Shahid, M.K. Hameed, I.T. Javed, K.N. Qureshi, M. Ali and N. Crespi, "A Comparative Study of Web Application Security Parameters: Current Trends and Future Directions, ", Applied Sciences, vol. 12, no. 8: 4077, 2022.

[5] J. Clarke, SQL Injection Attacks and Defense, 2nd ed., MA: Syngress, 2012. [E-book] Available: StudyLib, https://studylib.net/doc/26304630/sql-injection-attacks-and-defense.pdf---pdfdrive--.

[6] G. Betarte, E. Gimenez, R. Martinez and A. Pardo, "Improving Web Application Firewalls through Anomaly Detection," 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 779-784.

[7] R. Dubey and H. Gupta, "SQL filtering: An effective technique to prevent SQL injection attack," 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2016, pp. 312-317.

[8] K. Kim and S. Oh, "Edge Computing System applying Integrated Object Recognition based on Deep Learning," 24th International Conference on Advanced Communication Technology (ICACT), PyeongChang Kwangwoon_Do, Republic of Korea, 2022, pp. 415-419.

[9] G. Yolcu et al., "Deep learning-based facial expression recognition for monitoring neurological disorders," IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 2017, pp. 1652-1657.

[10] X. Liu, H. A. Ounifi, A. Gherbi, W. Li and M. Cheriet, " A hybrid GPU-FPGA based design methodology for enhancing machine learning applications performance ," Journal of Ambient Intelligence and Humanized Computing, vol.6, 2020, pp. 2309-2323.