Scientific Research Publishing

Journal of Information Security, 2021, 12, 34-55
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

The Overview of Database Security Threats' Solutions: Traditional and Machine Learning

Yong Wang, Jinsong Xi, Tong Cheng

Guangxi Key Laboratory of Cryptography and Information Security, Guilin University of Electronic Technology, Guilin, China
Email: 56247452@qq.com, 1132231474@qq.com, 1132482532@qq.com

How to cite this paper: Wang, Y., Xi, J.S. and Cheng, T. (2021) The Overview of Database Security Threats' Solutions: Traditional and Machine Learning. *Journal of Information Security*, **12**, 34-55.
https://doi.org/10.4236/jis.2021.121002

Received: November 17, 2020
Accepted: January 8, 2021
Published: January 11, 2021

Abstract

As an information-rich collective, there are always some people who choose to take risks for some ulterior purpose and others are committed to finding ways to deal with database security threats. The purpose of database security research is to prevent the database from being illegally used or destroyed. This paper introduces the main literature in the field of database security research in recent years. First of all, we classify these papers, the classification criteria are the influencing factors of database security. Compared with the traditional and machine learning (ML) methods, some explanations of concepts are interspersed to make these methods easier to understand. Secondly, we find that the related research has achieved some gratifying results, but there are also some shortcomings, such as weak generalization, deviation from reality. Then, possible future work in this research is proposed. Finally, we summarize the main contribution.

Keywords

Database Security, Threat Agent, Traditional Approaches, Machine Learning

1. Introduction

Database has been widely used in production and life, but data pool has been under severe security threats. At present, due to the development of computer network, technical loopholes and other factors, the database is often attacked [1]. In January 2019, data from a Philippine financial services company, were leaked, over 900,000 customer data were stolen by unauthorized hackers; In September 2019, Facebook confirmed that 419 million user phone information was leaked. In 2018, the losses of various network security incidents reached $45 billion, and most of events were related to databases. The above instances show that the study of database security is urgent.

With the increasing complexity of data and database functions, the change of attackers' attacking methods and the improvement of technology, traditional methods cannot meet the reality. Machine learning (ML) can transform sequential scanning into calculation model and DBA (Database Administrator) experience into prediction model, which makes the intrusion detection more intelligent and dynamic to adapt to the rapid variety of workload changes [2], and now computing power can satisfy machine learning. Therefore, there are more and more articles applying machine learning in database security threat response, but few people sort out these coping methods, which reflects the advantages of machine learning over traditional methods in dealing with some threats types.

This paper first obtains the source of database security threats, as shown in Figure 1. Then we carefully sort out and review the papers dealing with these threats, and find that machine learning has its advantages. Finally, we point out the shortcomings of relevant research and possible research directions.

The organization of this paper proceeds as follows. Section 2 summarizes data security issues and solutions. Sections 3, 4, 5 and 6 elaborate database security threats' solutions from four aspects: ineffectively data protection, user exception, vulnerability of defense system, and external attacks. Section 7 carries out research prospects and briefly sums up the full text.

2. Database Security Issues and Solutions

With the development of IT, database security risks are manifold [3]. We comb the research on database security, and find these factors closely related to database security: data, role, defense system, external factors. Therefore, we mark off four main threat sources: ineffective data protection, abnormal users, fragile defense system and external attacks. Data can be further divided into three categories: data tampering, data exposure, data being monitored or collected. User exception is subdivided into: illegal behavior, unauthorized access, weak security

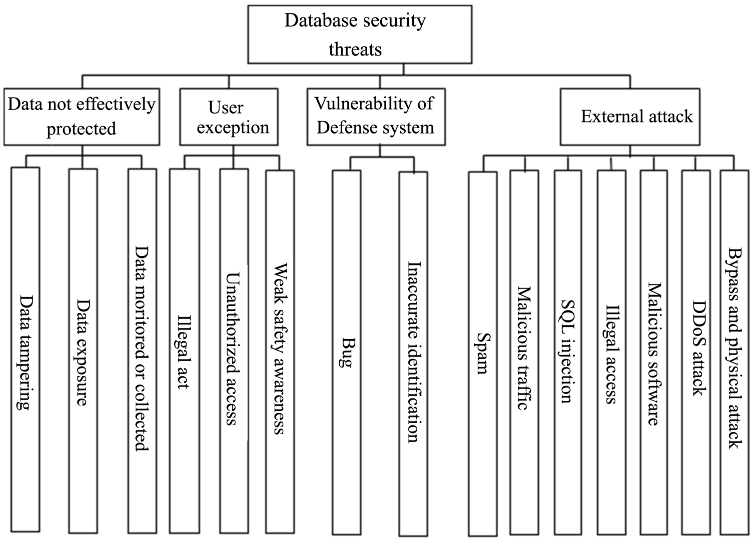

Figure 1. Sources of database security threats.

awareness. Weak defense system can also be divided by vulnerability, inaccurate identification. The external attacks are the main source of database security threats and they cause the most serious damage. Further, there are many secondary categories, including spam, malicious traffic, SQL injection, illegal access, malware, DDoS attacks, bypass and physical attacks. For the above-mentioned various security threats and their import, researchers use a series of methods to deal with these threats, as shown in Table 1. In the following four sections, the

Table 1. Solutions and damage to database security threats.

First level threats	Second level threats	Damage	Solutions
Data not effectively protected	Data tampering	Data distortion or invalid	Tamper detection, User authentication, data encryption, Tamper proof material
	Data exposure	Illegal use User' data	User authentication, data encryption, Audit, Construct machine learning model
	Data monitored or collected	Privacy disclosure	Establishment of special system, data encryption
User exception	Illegal act	Break the role code of conduct	Intrusion detection, Establishment of special system, User behavior analysis
	Unauthorized access	Illegal processing of data	Access control
	Weak safety awareness	Create a breakthrough for attackers	Empirical research
Vulnerability of Defense system	Bug	Used to destroy the database	Safety assessment, Empirical framework
	Inaccurate identification	Reject normal users and accept illegal users	User authentication
External attack	Spam	Occupy a lot of storage space and commit fraud	Access control
	Malicious traffic	Server works abnormally	Audit, Intrusion detection
	SQL injection	Embedded trojan horse and illegal right raising	Access control, Access control, User behavior analysis, System risk prediction
	Illegal access	Break system authentication mechanism and obtain others data	User authentication, Establishment of special system, Intrusion detection
	Malicious software	Illegal access to user secret data	Data encryption, Malware detection, Intrusion detection
	DDoS attack	System functions not available	Intrusion detection, Access control
	Bypass and physical attack	Hardware Damage and less preventable	Intrusion detection, Tamper proof material

above-mentioned four database security threats sources are expanded successively, various threats attack principles and response methods are analyzed in detail.

3. Data Ineffectively Protected Problems and Solutions

Data is the most watched factor among database security-related factors, since databases store large amounts of data. The data in the database is faced with serious threat. In January 2018, data from Indian citizenship database was leaked, including private information such as fingerprints and general personal information such as birthday. The main threat to data factor is ineffective data protection, such as data exposure, data tampering, data being monitored or collected. This section will focus on these threats.

3.1. Data Exposure Problems and Solutions

Data exposure means that data in a database is stored in clear text, and an attacker can easily get the data when he breaks through the defense system. In 2012, Rambler's database in Russia was leaked, and even more alarmingly, nearly 100 million user passwords were leaked and stored in plain text. Unfortunately, in order to the efficiency of access, much data is still stored in clear text recently.

Most researchers focus on data encryption. Ni *et al.* [4] proposed to encrypt sensitive data and the database, this method is only for specific systems and has poor scalability. Wang *et al.* [5] designed a general database encryption and decryption engine system, the system encrypted data on the application side, and utilized different user IDs to identify different transmission commands, and finally exploited the user's private key for encryption storage, but they should clarify the generation and distribution of keys. Hence, Huang *et al.* [6] adopted a weighted encryption scheme and related access control policies, however, the encryption and decryption process might be cumbersome excessively, leading to not so satisfactory application efficiency. Zhang *et al.* [7] firstly classified the users of web server: ordinary users, high-level users, and then encrypted the data of the high-level users. The method ensured the data security of high-level users, but might ignore ordinary users. Mei *et al.* [8] improved the AES algorithm and applied the encryption algorithm to the database management system, they converted the user name, password, database and user's activity with AES, this method had a wide range of applications and high reliability, but only processed binary files. Dandekar *et al.* [9] combined SHA-256 with ASCII control replacement technology to hide database data. They used the SHA-256 algorithm to make SOH replace binary information and generated hash values, and then compared the hash values with the encoded information, the efficiency of this approach was not so good. Andrey *et al.* [10] also embedded special code elements and representative data into a symmetric cryptographic algorithm, they replaced plain text elements with elements of the sequence associated with the key and then restored plain text through the key, the method effectively simplified the encryption operation, but had lower data security. Awais *et al.* [11] deployed parallel query execution techniques and AES on different data records, they used hash functions on metadata and multithreading on.NET applications, and then exploited AES encryption before inserting data into the data table. However, there would be conflicts when multiple technologies are used together. Uma *et al.* [12] utilized AES encryption and MD5 code conversion in the Medical Records Security System database. AES divides 128-bit medical data into four basic blocks for processing, while MD5 code divides any medical data into 512-bit data blocks and generates a fixed 128-bit length result, but the efficiency of this method is not high. He *et al.* [13] exploited quantum cipher to encrypt database data, they combined key dilution and auxiliary parameters, only a few quanta were sent in the quantum channel to generate the initial key, then the initial key was diluted by bitwise addition to several consecutive bits. The strategy's performance was high, but the quantum cipher was not yet mature enough. Fortunately, machine learning models were applied to data encryption. Shumeet [14] utilized DNN (Deep Neural Network) to hide image data from the database in the image. DNN is a neural network with a multilayer hidden layer. The basic structure is shown in **Figure 2** below. He exploited a large number of bits to embed RGB pixels of panchromatic images into another similar image of the same size, and then hid the decoding results and the appearance of the host image through a compression network of deep nerves. Experiments showed that the hiding effect was fine, but the hiding image required a lot of extra storage space. After an attacker detected a large number of hidden images, it was easy to recognize the image contents.

There are other ways to solve database data exposure issues. Wang *et al.* [15] firstly designed a signature scheme that could specify a verifier by using the authentication method. After signing the root node with this scheme, users need server participation to verify data using MHT tree. The experimental results showed that the verification speed was fast and the database data could be protected effectively, but the operability of the method was not strong. Jovan *et al.*

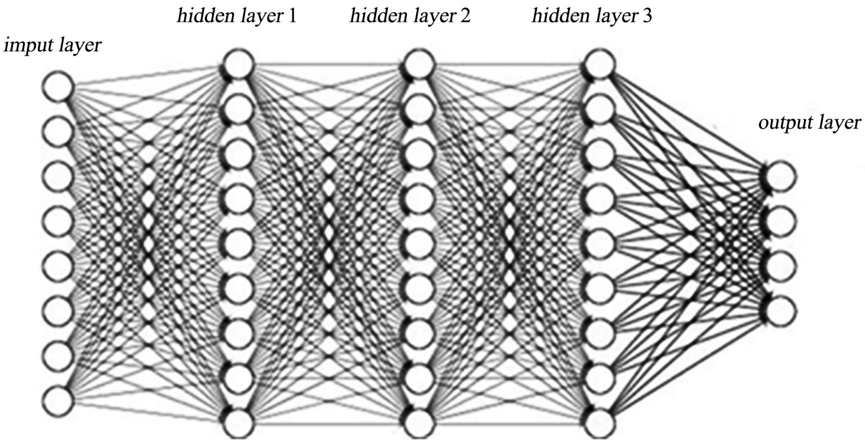

Figure 2. Basic structure of deep neural network.

[16] brought block chain technology to database security, their system sent different coded data blocks through separate channels, and exploited block chains to store encoding matrices for distributed storage systems. However, in some scenarios, each channel needs to be highly uncorrelated to avoid data interaction, so this method was limited. Auditing is also used by researchers, Vitthal *et al.* [17] proposed data auditing on the public cloud by third-party auditors. Auditors could read the data, but costs might be high unduly. Modeling methodologies are also considered. Minh *et al.* [18] attempted to build a common model for database data security using cloud services. They performed a feasibility analysis of information to create risk models. In machine learning, Boudheb *et al.* [19] exploited genetic algorithms and Naive Bayes to protect medical data. Genetic algorithm was a computational model that simulated the natural selection and genetic mechanism of Darwin's biological evolution. On the premise of independent and identical distribution of objects, Naive Bayesian obtains the posterior probability of objects from the prior probability of objects, and then uses the maximum posterior probability to determine the category of objects [20]. The specific calculation steps are as following: **Figure 3**. There are many sources of medical data and complex storage. The selection of safety features played a decisive role in the training model, the paper utilized the most representative safety features (patient identification, birthday, blood type, etc.).

3.2. Data Tampering Problems and Solutions

Data tampering means that the data in the database has been illegally altered, the situation causes the original data to be lost, replaced, or added or subtracted. In January 2010, the website of an educational examination center was invaded, and somebody logged into the database, he added a record of someone's exam passing information, such behavior seriously violated the fairness of the examination.

Some research is intended to prevent data tampering. Piggin *et al.* [21] exploited honeypot technology in common physical components of a database system to attract attackers to modify fake data, and then to protect truly valuable data, but there was a risk that the honeypot could be used to attack by attackers. Elena *et al.* [22] implemented data entry through spin current, they made use of

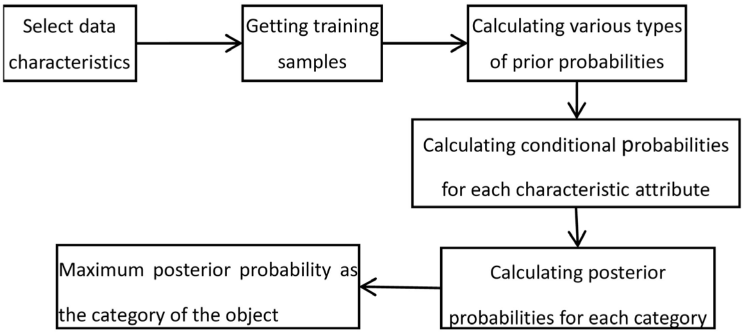
Select data characteristics
Getting training samples
Calculating various types of prior probabilities
Calculating conditional probabilities for each characteristic attribute
Calculating posterior probabilities for each category
Maximum posterior probability as the category of the object

Figure 3. Steps of simple Bayesian calculation.

the high variability that affected the resistance of magnetic tunnel junction devices and the special configuration of read operation reference units to make data physically non-cloning, the effectiveness of this method was proved in theory, but lack of practical verification. The development of machine learning also brings an opportunity to solve the issue. Some researchers have focused on ECG (electrocardiogram) data, which is physically non-cloning and can effectively combat data tampering. Yin *et al.* [23] learned and extracted different features before using the neural network training data to minimize overlap in the distribution of cosine/hamming distances between individual and inter-individual, but that needed large amount of calculation. Kiran [24] introduced a minimum absolute contraction selection operator to identify the most appropriate ECG features. This method effectively avoided random, correlated, and over-fitting features, and reduced the feature space, and improved the prediction speed, but the detection accuracy was reduced slightly. He also proposed an effective ECG feature extraction method [25], which extracted six optimal segments based on priority and normalizes positions, but there might be over fitting.

Some research focuses on the processing of data tampering when it has happened. Li *et al.* [26] hoped that the normal data query service would continue after the data was partially polluted. They utilized the data query service rules to determine whether they decided to return the user's partially legitimate data collection. This method improved the usability of the database, but could not determine the location of the data pollution. Yin *et al.* [27] designed a detection mechanism for database tampering, they exploited two signatures both horizontally and vertically to ensure that the data table could be detected by signature after tampering with the data sheet. However, the system cost a lot and the operation was cumbersome. Xian *et al.* [28] took a simpler approach. After the server responded to a query request, the servicer sent the verification value, the mask of the verification tree, and the signature of the mask and the number of root nodes of the verification tree to the query party to verify whether the data had been tampered with. This method effectively decreased the amount of computation, but reduced the safety. In machine learning method, Lai *et al.* [29] exploited K-means clustering algorithm. K-means' workflow is: randomly selecting k points as the initial centroid, and then assigning each point in the dataset to a cluster. In order to detect the web page data which had been tampered with, they grabbed information from the first page of some websites and established detection rules by classifying the data to determine whether the web page had been misrepresented. However, this method needed to adjust the detector, which required rich experience in dealing with hackers, and wrong adjustment would greatly reduce the recognition effect.

3.3. Data Monitored or Collection Problems and Solutions

Data is eavesdropped or collected by an attacker during transmission, and then they analyze the information about the target. Recently, social software has been exposed to monitor user chat records, the conduct seriously violates user privacy.

Data encryption is the most common way to solve the problem. Kushko *et al.* [30] proposed a new method to protect network data transmission, they hid the interaction between nodes in the network and utilized encryption, multicast and packet retransmit for traffic interaction, the operation was too complicated. Andrey *et al.* [31] introduced the homologous encryption and logistic regression model. Homologous encryption enabled people to perform certain forms of algebraic operations on cipher text and still encrypted it. The result of decryption was the same as that of plain text. They lessened the storage of encrypted databases by using an approximate homologous encryption method, and accelerated gradients by using logistic regression models to speed up computations. However, logistic regression was prone to the phenomenon of under fitting. In addition to data encryption, Li *et al.* [32] exploited a remote method to invoke the server to receive and parse network packets transmitted by the server-side proxy, and then to filter the address information securely, and finally to invoke the JDBC driver to connect to each database management system for data interaction and return the results, but the solution was costly to implement.

4. User Exceptions Problem and Solutions

User exceptions are the most difficult to guard against in database security threats. In March 2017, Tencent jointly with the Jingdong security team uncovered a case of self-theft. An insider in Jingdong stole more than 5 billion pieces of information. After that, they made profits by selling through various illegal ways, such action caused huge economic and reputation losses in Jingdong. Researchers subdivide user anomaly threats into illegal behavior, unauthorized access, and weak security awareness.

4.1. Illegal Acts Problems and Solutions

Illegal behavior refers to the user's behavior that violates the role positioning or behavior rules in the database, such as unauthorized access to the database, users' illegal operations in the database system, and so on.

Researchers want to detect such behaviors. Chen *et al.* [33] utilized C and C# to achieve real-time tracking and analysis of database operation information, database and server status, but the efficiency should be improved. In order to improve the processing speed, the machine learning model is introduced. Liu [34] exploited the naive Bayesian classification algorithm to build files for each database role, then trained the user behavior database, and finally classified the database transaction through the user behavior database, but it lacked experimental support. Andrey *et al.* [35] utilized a K-means clustering algorithm to process text log information. They converted the text log information into clustering vectors, calculated outliers, and sorted the output anomalies to get the clusters to which the user behavior log information most likely belonged, but the processing accuracy needed to be improved, and this method could only apply single structure text log.

4.2. Unauthorized Access Problem and Solutions

Unauthorized access refers to users illegally accessing data that does not conform to their privileges by means of delegation, etc. An average user can be an administrator, or even a super administrator by privilege promotion, and then he can acquire other user data.

Access control is a widely used solution. Xu *et al.* [36] gave the user a multilevel role name based on which to acquire internal roles before granting the user permissions, but this method could not resist hidden channel access effectively. He *et al.* [37] utilized the security baseline to evaluate the database access control, and took measures to improve the control effect after quantifying the score, however, there was no specific method to improve the effect of access control. An *et al.* [38] exploited the history of multi-connection pool and different configurations to achieve strict and dynamic access control. Yang *et al.* [39] proposed a method to refine database access control through permission extension. They split the primary key in the permission table into corresponding storage structure and saved permission information with built-in key values to achieve more refined access control, but the application scenarios were limited.

4.3. Weak Safety Awareness Problem and Solutions

Weak security awareness means that database users create attack points that may be exploited by attackers for the sake of saving trouble, such as setting weak password and not modifying the default password of database, the consciousness can improve security through educational means. Therefore, there are a few related technological research papers. Yung *et al.* [40] investigated the impact of security awareness on bank security performance management and the use of information technology through a questionnaire, and concluded that compliance had a significant impact on information security management performance and information technology capabilities.

5. Vulnerability of Defense System Problem and Solutions

The vulnerability of database defense system is reflected in two layers: the operating system layer and the database layer. The former refers to that the user's host is easy to be controlled by hackers and then attacked, while the latter refers to the unclear division of storage authority and the incorrect configuration by DBA. There was fragility in SQL server, the default password of SA, the super administrator, was empty. Attackers could log in to SQL server directly through SA account without password. There are two reasons for the vulnerability of database defense system: firstly, there are defects in initial configuration, secondly, the system's identification is not accurate.

5.1. Bug Problems and Solutions

Bug refers to the design defects of database defense system. In May 2011, hackers used the user of Oracle database to invade the database of Korea Convention and

Exhibition Center. The reason why the system was broken was that the DBSNMP user used the default password.

Most researchers adopt the strategy of defense in advance. Gao [41] designed a database security evaluation model for SQL server, Sybase and Oracle, the method could evaluate the overall security of the database. Kozlov *et al.* [42] utilized fuzzy logic to evaluate the security of enterprise information management system. They expressed all possible threats of the system into a function, and each value of the function represented a possible threat. An attack tree was constructed to deal with each threat. Zhang *et al.* [43] presented intelligent security assessment for system software. They exploited crawlers to obtain natural language evaluation data, and then utilized various machine learning methods to obtain safety evaluation indicators to build a security assessment model, but the method was not easy to implement.

5.2. Inaccurate Identification Problems and Solutions

Inaccurate identification means that the illegal users are wrongly identified as normal users or normal users are identified as illegal users when the database conducts identification. As a good identification method, biometric identification method is fast, safe and rapid development, but biometrics verification also acts out some problems.

Prabu *et al.* [44] exploited the effective linear binary pattern and scaled invariant Fourier transform to process and store the biometrics of hand type and iris into database, and then utilized neural network and Bayesian network classifier to detect, due to mix the two biological features together, the recognition is inefficient. Musab *et al.* [45] exploited CNN (Convolutional Neural Network) to improve the recognition effect of face recognition. CNN is a feed forward neural network with deep structure including convolution calculation. The basic structure is shown in **Figure 4** below. The author improved CNN by adding standard operation between input layer and output layer, the improvement could accelerate network standardization, but there was a problem of over-fitting in face recognition. Aishwarya *et al.* [46] utilized aggregation and RF (Random Forest) to improve face recognition rate. RF introduces random attribute selection in the training process of decision tree [47]. They exploited local aggregation to store the features of the detected face images, and then use RF to train and classified face images, this method consumed a lot of storage space. However, Csaba [48]

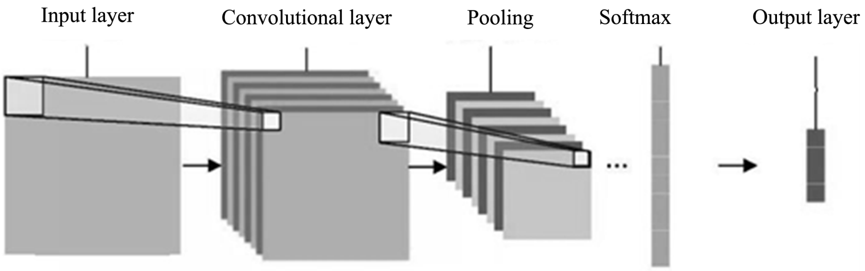

Figure 4. Basic structure of convolutional neural network.

pointed out the inherent problems of biometrics: lacking relevant features, high spending, and privacy issues.

6. External Attack Problems and Solutions

Overt attack refers to external attacker directly threatening database security through some ways. The blackmail virus appeared in the first half of 2019, which encrypted the important data in the user system. This virus had caused great damage to the social service infrastructure on a global scale. As the main threat to database security, external attacks can be roughly divided into seven categories: spam, malicious traffic, SQL injection, illegal access, malware, DDoS attacks, bypass and physical attacks.

6.1. Spam Problems and Solutions

Spam refers to a large number of emails sent by attackers to users with phishing, advertisements, viruses, etc., junk mail will occupy a large amount of storage space in the e-mail database, and users may suffer economic losses after clicking on such e-mails. The operation of the mail system will involve multiple databases and protocols, and the specific process is shown as Figure 5.

Researchers use machine learning method to improve the ability of spam detection. He *et al.* [49] utilized language decision tree to improve the performance of spam detection based on semantic features. Language decision tree classifies samples with different linguistic attributes through tree structure. They extracted feature information from spam information and decomposed junk mail into several feature subsets in the light of the meaning of attributes. Then they processed, classified and trained these feature subsets by using language decision tree to get the spam classification model, however, they did not consider that the machine learning model was attacked.

6.2. Malicious Traffic Problems and Solutions

Malevolent traffic refers to a large number of requests forged by external attackers through some tools to prevent normal users from accessing the database. In August 2019, snapex platform was attacked by malicious traffic saturation by hackers, which made the platform users temporarily unable to access, and some users suffered economic losses because they were unable to trade virtual currency.

There are many ways for researchers to deal with malicious traffic. Zhang *et al.* [50] exploited the method of security audit. They firstly captured the user's access operation data to the database, then submitted the data to the auditors for analysis, and finally fed back the results. The method relied too much on the

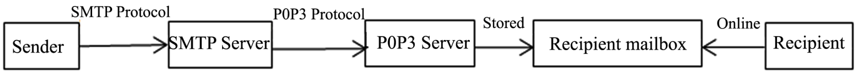

Figure 5. Workflow of mail system.

expertise of auditors and was inefficient to handle large-scale data situations. Yu [51] introduced CNN (convolutional neural network) for intrusion detection. He numerically processed the traffic data set of KDD99 network, and then set the learning rate, iteration times, sample size and other parameters of the convolutional neural network, and finally trained the traffic data to generate the traffic classification model. During the training process, the random gradient descent method was used to accelerate the convergence speed of the model, however, the accuracy of other types of malicious traffic classification was not high, and it was only experimented on classical datasets and not detected in practical applications.

6.3. SQL Injection Problems and Solutions

SQL injection means that external attackers submit query code to the database and get the desired data according to the feedback from the database. The attacker can enter select * from users where (username: "1" or "1" = "1") and (password: "1" or "1" = "1"), the operation can bypass the user name and password input and directly log into the database management system.

There are many ways to solve SQL injection problems. Li *et al.* [52] prevented SQL injection through the integrity evaluation of user behavior policy. The administrator stored the mandatory access policy and behavior constraint code summary value into the database management system. After the user submitted the transaction, TSB (the trusted software base) calculated the behavior constraint code summary value and compared the result with the previously stored value, thus finding an exception, but this could not handle logical operations, concurrent transactions. Ma *et al.* [53] defined a set of ternary strings in the flow of accessing the database, including user name, password and SQL injection attack detection results, so as to describe the probability of database intrusion, and then they executed the pattern matching algorithm. If the user name and password were the identical and the result of SQL injection attack detection is normal, users would be allowed to access, detection of SQL injection attack took too much time to detect, which resulted in long processing time and could not meet the requirements of real-time detection. Xing *et al.* [54] proposed a real-time detection model of SQL injection attack, which included two parts: real-time detection module and model training module, the real-time detection module detected the access packets in the actual network environment. The model training module improved the convolutional neural network, and normalized the SQL injection samples between 0 - 1. In the training process, the model training module used ReLu activation function and ADM algorithm, and exploited dropout strategy to prevent over fitting, this method was slow to detect.

Machine learning specific methods are also used to resist SQL attacks. Hu *et al.* [55] utilized vulnerability mining methods to solve the problem of SQL injection. They labelled the PHP class SQL injection code and transformed the code using the bag of words model, and then exploited SVMs (support vector machines) to classify them. Finally, they put new PHP files into the model for classification. Only SQL injection attacks limited to PHP classes were handled. Solomon *et al.* [56] utilized pattern driven corpus to reduce the harm of SQL injection attack on back-end database. They extracted SQL injection code and exploited regex constraint analytic learning, and then hashed the features to the matrix to meet the needs of classifiers, and divided the feature matrix values into training set and test set, and finally trained the feature matrix value with support vector machine to obtain support vector machine classifier, this approach was specific to specific types of SQL injection attacks.

6.4. Illegal Access Problems and Solutions

Illegal access refers to the external attacker for some illegal purposes to access the database. The deed will cause damage to the database or obtaining the desired data. Attackers write codes to bypass the database management system and its authorization mechanism, and directly access and modify the data in the database through the operating system.

A small number of researchers adopt unique methods. Xiang *et al.* [57] utilized SSL protocol to build a secure channel between server and database users, and adopted double authentication. Fu [58] proposed the strategy of building database security defense system. He exploited antivirus software and data mining technology to deeply analyze database information. Seok *et al.* [59] combined convolutional neural network and learning classifier system to improve the detection effect. They extracted features from logs to obtain feature vectors, then selected features by rules, trained with convolutional neural network, and obtained chromosome model with better characteristics by genetic operation, and extracted feature classification for new transactions to find anomalies, but it was difficult to ensure the stability of classification results because of the poor genetic operation.

Most researchers use intrusion detection methods. Wang *et al.* [60] monitored, counted and analyzed the log and illegal access behavior on SQL Server 2000 database. Li *et al.* [61] utilized SQL statement structure, statement operation data and system behavior to detect whether the transaction submitted by normal users, this method was prone to single-task misses. Tang *et al.* [62] were concerned about the illegal scanning of the server port. They extracted several features in the log by Naive Bayesian algorithm and limited the threshold value. If the threshold value was exceeded, Naive Bayesian model would determine the deed as the scanning behavior and blocked. This method relied on the log information and couldn't handle semi open scanning. Zhang [63] combined support vector machine and ant colony algorithm to build an intrusion detection classifier. The specific process is shown in **Figure 6**. The recognition accuracy of the model was more than 95%, and the recognition speed was fast. But the ant colony algorithm was easy to fall into the local optimal solution, which made the parameters not optimal, and the parameters affected the classification accuracy. Jong *et al.* [64] solved the problem of illegal connection through traffic analysis. They developed an abnormal link detection system, which could detect real-time

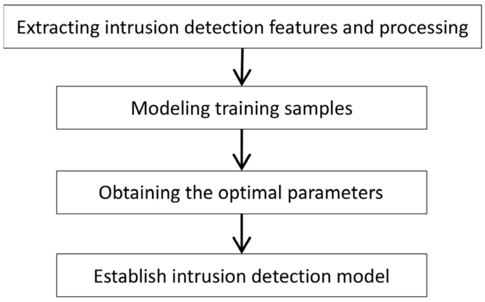
Extracting intrusion detection features and processing
Modeling training samples
Obtaining the optimal parameters
Establish intrusion detection model

Figure 6. Network intrusion detection classifier construction process.

data flow of MySQL database, but the method only had a narrow application range. Salimov *et al.* [65] detected and countered attack IP addresses that exceeded the specified threshold, they formulated a white list of database access and exploited honeypot technology to attract attackers to visit specific databases and analyzed attackers. Finally, a model was established according to the threat degree. Sharmila *et al.* [66] utilized density based clustering technology and supervised learning for database intrusion detection. They constructed normal behavior clusters of users, and the transactions submitted by users were either located in normal behavior clusters, or were classified into abnormal behavior clusters because of local outlier factors, so as to identify intrusion documents, but this method had a high false positive rate.

6.5. Malware Problems and Solutions

Malicious software refers to a kind of special software that attackers secretly install on the victim's terminal. They are used to monitor, collect and even destroy resources, including database data. The attacker can obtain the database administrator's authority through some rootkit tool, so as to further implement the database data replication, modification, deletion and other destruction actions.

While mobile phones bring convenience and entertainment to people, there are also some malicious applications, which steal data, monitor calls, and even damage the storage area of mobile phones. Liu *et al.* [67] exploited K-Nearest clustering to detect Android malware. K-Nearest clustering finds K records closest to the new data from the training set, and uses the main classification to determine the category of new data. They utilized reverse engineering to obtain application feature information, and trained K-Nearest clustering algorithm to obtain malware detection model, Chen *et al.* [68] resisted the repackaging attack of Android applications through encryption technology. They encrypted the code of the application. If the application was repackaged, the decrypted code would change, thus people could discover that the application had become malware, but the same problem existed.

Some researchers use common methods to resist all kinds of malware. Liu *et al.* [69] collected common malicious code into malicious images and processed the image into the same size. After that, the image samples were input into convolutional neural network for training to obtain a classification model. This method had good detection effect, but the key feature information of malicious images might be lost if the malicious images were transformed into the same size. Ajay *et al.* [70] exploited virtual machine monitoring technology and machine learning to build a malware detection system. They reconstructed executable files, and then detected and classified these files through advanced client assisted automatic multi-level malware detection system, but the use of virtual machine monitoring technology could lead to a steep increase in economic input and complex configuration.

6.6. DDoS Attack Problems and Solutions

DDoS is one of the most common malicious traffic. Large scale DDoS requests will occupy the network traffic, and constantly submit query requests to the database, thus preventing real users from accessing the actual services [71]. In May 2016, anonymity, the world's largest hacker organization, launched a short-term DDoS attack on bank websites around the world. This attack led to the network system of many central banks in a state of paralysis. DDoS attacks can also bring about website failure loading, unavailable software, and the state of failure logging in game accounts.

Bashar *et al.* [72] had carefully sorted out the application of artificial intelligence and statistical methods in resisting DDoS attacks in recent years. According to the types of vulnerabilities, the degree of automation and the degree of dynamics, they divided DDoS attacks. After summarizing the achievements of predecessors, they put forward better methods to deal with DDoS attacks, they also pointed out that the existing classifications of DDoS attacks were not detailed enough to block certain specific DDoS attacks. Liu *et al.* [73] resisted capacity attacks through traffic control. They exploited a router dependent access control list to eliminate traffic that did not pass through all MBOX, thus avoiding a large number of unknown traffic blocking normal database query requests and other database operations. Shan *et al.* [74] utilized feature learning, multi-kernel learning and automatic encoder to predict DDoS attacks. Multi-kernel learning refers to fusing multiple kernel functions when using support vector machines. Automatic encoder is a special neural network, which transforms the input into features, and then reconstructs the original input from the new features. They firstly learnt multi-level automatic encoder, then exploited the encoder to train data to get features, and finally combined multi-core learning algorithm to obtain a unified detection model, but multi-core learning was sensitive to large sample data, which made the computation time-consuming and space-consuming, resulting in poor detection performance.

6.7. Bypass and Physical Attack Problems and Solutions

Bypass and physical attacks refer to attacking on database hardware. If the laser fault injector is used to inject faults into the chip supporting the database, the system will work abnormally.

In recent years, there are a few related research papers. In order to resist bypass attacks, Mohammad *et al.* [75] exploited Intel cache monitoring technology and hardware performance counter to provide hardware fine-grained information, and utilized Gaussian anomaly detection method to detect bypass attacks based on virtual machine cache, but this method could only be used on devices that installed Intel processors. It also targeted specific bypass attacks and had poor scalability. Jeyavijayan *et al.* [76] exploited emerging materials to make electronic devices, such as silicon nanowire field-effect transistors and Nano electromechanical switches. Experimental results showed that Nano materials had security advantages over traditional materials, but there were also many challenges, including equipment stability, new protocols, etc.

7. Conclusions

According to the above contents, in the light of data, users, protection system and external attackers, researchers study database security. Some efficient and high recognition methods have been produced in the field of data tampering, data exposure and illegal access. However, some methods to deal with database security threats are only feasible in theory. Therefore, they should improve to adapt the actual needs.

1) The generalization ability of the method is improvement required. In reality, database security is threatened in many ways. However, most of the research only focuses on one aspect or even a specific threat to database security. Although some methods may perform well in well-designed experiments, the protection effect may not be ideal in real environment. People hope that a solution can be extended to all aspects of database security as far as possible. It has become a difficult problem to be studied.

2) Some methods need more practical requirements. There are conditions when any idea or scheme is put into practice: safety, ease of implementation and understanding, cost, etc. Previous studies paid more attention to security and accuracy, some of which were too complex and costly. In future research, we should pay attention to convenience, simplicity, efficiency, user experience, cost and other factors while ensuring security.

After discussing the recent research on database security and the existing problems, we analyze the possible research directions in the future. The method of combining similar scenarios can combine the DNN method of image data processing proposed by Shumeet [11] with the Naive Bayesian method of text data processing by Boudheb *et al.* [16] to satisfy the effective protection of data in the hospital DBMS. According to the use frequency and privacy level, different types of data can be encrypted discriminatingly. Using K-means clustering algorithm model based on user behavior information in database to log information is a possible research point. After extracting the feature information of the vulnerability, various machine learning methods can be used to detect the vulnerability quickly and automatically, the same processing method can be introduced to other database threat processing. When using machine learning to deal with database security threats, the idea of ensemble learning is reasonable, better classification results can be obtained by using AdaBoost algorithm.

In this paper, we find that database threats include data, users, protection system and external attacks, then their harms and the existing literatures' solutions are elaborated. Then, we sum up the current shortcoming of related research, and finally give several possible research directions.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Meng, J. (2018) Security Threats and Countermeasures of Computer Network Database. *PC Fan*, **23**, 37-37.

[2] Meng, X.F., Ma, C.H. and Yang, C. (2019) Review of Machine Learning Database System. *Computer Research and Development*, **56**, 1803-1820.

[3] Zhan, S.Q. (2017) Security Threats and Measures of Computer Network Database. *Telecom World*, **3**, 39-40.

[4] Ni, Q. and Mao, Y.G. (2014) Database Security in Measurement Management Information System. *Computer and Modernization*, **1**, 182-185+191.

[5] Wang, X., Zhu, Z.X., Shi, C.Y., *et al.* (2014) Encryption and Decryption Engine System for Database Security Protection. *Computer Research and Development*, **24**, 143-146.

[6] Huang, B.H., Jia, F.W. and Wang, T.J. (2016) Attribute Based Database Access Control Policy under Cloud Storage Platform. *Computer Science*, **43**, 167-173.

[7] Zhang, J.X., Chen, M.L. and Wang, Q. (2010) Intrusion Threat and Protection Strategy of Web Database. *Science and Technology Bulletin*, **26**, 769-773.

[8] Mei, H.W., Li, C.H. and Zhang, M.Q. (2010) Enterprise Database Security Strategy Based on Improved AES Algorithm. *Microcomputer Information*, **26**, 19-21.

[9] Dandekar, S.C., Ahire, P.G. and Rao, J. (2018) Improved Secret Information Hiding Using SHA-256 and Invisible ASCII Character Replacement Technology. 2018 *Fourth International Conference on Computing Communication Control and Automation*, Pune, 16-18 August 2018, 1-4. https://doi.org/10.1109/ICCUBEA.2018.8697764

[10] Andrey, A. and Natalya, Z. (2018) Mathematical Model of Symmetric Cryptoalgorithm Based on Representing Mumbers as Sums of Special Code Elements. 2018 *Global Smart Industry Conference*, Chelyabinsk, 13-15 November 2018, 1-6. https://doi.org/10.1109/GloSIC.2018.8570078

[11] Ahmad, A., Ahmad, M., Habib, M.A., Sarwar, S., Chaudhry, J., Latif, M.A., *et al.* (2019) Parallel Query Execution over Encrypted Data in Database-as-a-Service (DaaS). *The Journal of Supercomputing*, **75**, 2269-2288. https://doi.org/10.1007/s11227-019-02831-8

[12] Uma Maheswari, S. and Vasanthanayaki, C. (2019) Secure Medical Health Care Content Protection System (SMCPS) with Watermark Detection for Multi Cloud Computing Environment. *Multimedia Tools and Applications*, **79**, 4075-4097. https://doi.org/10.1007/s11042-019-7724-z

[13] He, X.Y., Pei, C.X. and Yi, Y.H. (2015) A Low Complexity Quantum Private Information Retrieval Protocol. *Journal of Xi'an University of Electronic Science and Technology* (*Natural Science*), *No.* 5, 33-37, 74. http://dx.chinadoi.cn/10.3969/j.issn.1001-2400.2015.05.006

[14] Shumeet, B. (2019) Hiding Images within Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 1685-1697. https://doi.org/10.1109/TPAMI.2019.2901877

[15] Wang, X.M. and Yuan, D.B. (2010) Privacy-Protecting Outsourcing Database Query Verification Technology. *Journal of Beijing University of Technology*, **36**, 703-709.

[16] Karamacoski, J., Paunkoska, N., Marina, N. and Punčeva, M. (2019) Blockchain for Reliable and Secure Distributed Communication Channel. 2019 *IEEE International Conference on Industry* 4.0, *Artificial Intelligence*, *and Communications Technology*, Bali, 1-3 July 2019, 91-97. https://doi.org/10.1109/ICIAICT.2019.8784853

[17] Gutte, V.S. and Deshpande, P. (2015) Cost and Communication Efficient Auditing over Public Cloud. 2015 *International Conference on Computational Intelligence and Communication Networks*, Jabalpur, 12-14 December 2015, 807-810. https://doi.org/10.1109/CICN.2015.164

[18] Nguyen Minh, T. and Khorev, P.B. (2019) Information Risks in the Cloud Environment and Cloud-Based Secure Information System Model. 2019 *International Youth Conference on Radio Electronics, Electrical and Power Engineering*, Moscow, 14-15 March 2019, 1-6. https://doi.org/10.1109/REEPE.2019.8708845

[19] Tarik, B. and Zakaria, E. (2019) Privacy Preserving Feature Selection for Vertically Distributed Medical Data based on Genetic Algorithms and Naïve Bayes. *International Journal of Information System Modeling and Design*, **9**, 1-22. https://doi.org/10.4018/IJISMD.2018070101

[20] Wang, S.Y. (2015) Research on Intrusion Detection Method Based on Machine Learning. *Journal of Chaohu College*, **17**, 25-27. http://dx.chinadoi.cn/10.3969/j.issn.1672-2868.2015.06.006

[21] Piggin, R. and Buffey, I. (2016) Active Defense Using an Operational Technology Honeypot. 11*th International Conference on System Safety and Cyber-Security*, London, 11-13 October 2016, 1-6. https://doi.org/10.1049/cp.2016.0860

[22] Ioana Vatajelu, E., Di Natale, G., Torres, L. and Paolo Prinetto (2015) STT-MRAM-Based Strong PUF Architecture. 2015 *IEEE Computer Society Annual Symposium on VLSI*, Montpellier, 8-10 July 2015, 467-472. https://doi.org/10.1109/ISVLSI.2015.128

[23] Yin, S.H., Bae, C.S., Kim, S.J. and Seo, J.-S. (2017) Designing ECG-Based Physical Unclonable Function for Security of Wearable Devices. 2017 39*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Jeju Island, 11-15 July 2017, 3509-3512. https://doi.org/10.1109/EMBC.2017.8037613

[24] Kumar Patro, K., Prakasa Rao Reddi, S., Ebraheem Khalelulla, S.K., Rajesh Kumar, P. and Shankar, K. (2019) ECG Data Optimization for Biometric Human Recognition Using Statistical Distributed Machine Learning Algorithm. *Journal of Supercomputing*, **76**, 858-875. https://doi.org/10.1007/s11227-019-03022-1

[25] Kumar Patro, K., Rajesh Kumar, P. (2017) Machine Learning Classification Approaches for Biometric Recognition System Using ECG Signals. *Journal of Engineering Science and Technology Review*, **10**, 1-8. https://doi.org/10.25103/jestr.106.01

[26] Li, L., Qin, X.L. and Dai, H. (2013) Damage Tolerance Data Query Degradation Service Mechanism. *Computer Science*, **40**, 90-93.

[27] Yin, T.F., Xie, X.L. and Mei, X.L. (2014) Detection Mechanism of Database Tampering Based on Digital Signature and HSM. *Journal of East China University of Technology* (*Natural Science*), **40**, 376-380. http://dx.chinadoi.cn/10.3969/j.issn.1006-3080.2014.03.018

[28] Xian, H.Q. and Feng, D.G. (2010) Integrity Detection Scheme in Qutsourced Database Model. *Computer Research and Development*, **47**, 1107-1115.

[29] Lai, Q.N., Chen, S.Y., Ma, H., *et al.* (2016) Batch Web Page Tampering Detection Method Based on Machine Learning. *Journal of Central China University of Science and Technology* (*Natural Science Edition*), **44**, 16-20. http://dx.chinadoi.cn/10.13245/j.hust.161104

[30] Kushko, E.A. and Parotkin, N.Y. (2019) Efficiency Evaluation of Secure Data Communication Protocols Stack Based on Dynamic Network Topology. 2019 I*nternational Russian Automation Conference*, Sochi, 8-14 September 2019, 1-6. https://doi.org/10.1109/RUSAUTOCON.2019.8867782

[31] Kim, A., Song, Y., Kim, M. and Cheon, J.H. (2018) Logistic Regression Model Training Based on the Approximate Homomorphic Encryption. *BMC Medical Genomics*, **11**, Article No. 83. https://doi.org/10.1186/s12920-018-0401-7

[32] Li, B., Chen, M.Y. and Gu, F.Q. (2010) Research and Implementation of Database Cross-Network Access. *Power System Automation*, **34**, 103-105.

[33] Chen, L. and Yuan, X.P. (2010) Analysis and Improvement of Information Acquisition Technology for Database Server. *Computer and Modernization*, **181**, 94-96.

[34] Liu, D. (2017) Database Intrusion Detection System Based on Naïve Bayesian Classification Algorithm. *Network Space Security*, **8**, 32-34.

[35] Sapegin, A., Gawron, M., Jaeger, D., Cheng, F. and Meinel, C. (2017) Evaluation of In-Memory Storage Engine for Machine Learning Analysis of Security Events. *Concurrency and Computation*: *Practice and Experience*, **29**, e3800. https://doi.org/10.1002/cpe.3800

[36] Xu, P.J., Zheng, J. and Xu, M.J. (2015) Role-based Multilevel Security Database Model. *Computer Engineering*, **41**, 135-138.

[37] He, B.Y. and Liu, R. (2013) Oracel and SQL Server Database Security Baseline Review. *Journal of Yunnan University* (*Natural Science Edition*), **35**, 63-68.

[38] An, H.X. and Xu, Y.S. (2010) Dynamic Multi-Connection Pool for Efficient Database Access. *Microcomputer Application*, **31**, 34-41. http://dx.chinadoi.cn/10.3969/j.issn.2095-347X.2010.12.006

[39] Yang, X.Y. and Gan, L.M. (2018) Non-Relational Database Security Research of NOSQL Based on Hadoop. *Microcomputer Application*, **34**, 43-45. http://dx.chinadoi.cn/10.3969/j.issn.1007-757X.2018.12.015

[40] Chang, Y., Wu, L.F. and Shiann, M.W. (2017) A Study on the Impact of Regulatory Compliance Awareness on Security Management Performance and Information Technology Capabilities. 2017 13*th International Conference on Natural Computation*, *Fuzzy Systems and Knowledge Discovery*, Guilin, 29-31 July 2017, 2866-2871.

[41] Gao, H.T. (2015) Design of Database Security Evaluation Software. *Electronic Test*, **1**, 31-32.

[42] Kozlov, D. and Noga, N.L. (2018) Risk Management for Information Security of Corporate Information Systems Using Cloud Technology. 2018 11*th International Conference* "*Management of Large-Scale System Development*", Moscow, 1-3 October 2018, 1-5. https://doi.org/10.1109/MLSD.2018.8551947

[43] Zhang, Y.F., Tang, E.Y., Su, D.Z., Kuang H.-Y. and Chen, X. (2018) Intelligent Software Security Evaluation Method Driven by Natural Language Data. *Journal of Software*, **29**, 2336-2349. http://dx.chinadoi.cn/10.13328/j.cnki.jos.005526

[44] Prabu, S., Lakshmanan, M. and Noor Mohammed V. (2019) A Multimodal Authentication for Biometric Recognition System using Intelligent Hybrid Fusion Techniques. *Journal of Medical Systems*, **43**, Article No. 249. https://doi.org/10.1007/s10916-019-1391-5

[45] Coúkun, M., Uçar, A.l, Yildinm, O. and Demir, Y. (2017) Face Recognition Based on Convolutional Neural Network. 2017 I*nternational Conference on Modern Electrical and Energy Systems*, Kremenchuk, 15-17 November 2017, 376-379. https://doi.org/10.1109/MEES.2017.8248937

[46] Aishwarya, K., Suresh Kumar, B., Viswanadha Raju, S. (2019) Facial Recognition Using Aggregation and Random Forest Classification Method. *Journal of Physics*: *Conference Series*, **1362**, 1-9. https://doi.org/10.1088/1742-6596/1362/1/012078

[47] Zhou, Z.H. (2016) Machine Learning. Tsinghua University Press, Beijing.

[48] Otti, C. (2016) Comparison of Biometric Identification Methods. 2016 11*th IEEE International Symposium on Applied Computational Intelligence and Informatics*, Timisoara, 12-14 May 2016, 339-344. https://doi.org/10.1109/SACI.2016.7507397

[49] He, H.M., Tim, W., Carsten, M., Mehnen, J. and Tiwari, A. (2017) A New Semantic Attribute Deep Learning with a Linguistic Attribute Hierarchy for Spam Detection. 2017 *International Joint Conference on Neural Networks*, Anchorage, 14-19 May 2017, 3862-3869. https://doi.org/10.1109/IJCNN.2017.7966343

[50] Zhang, W.W., Zheng, F.G. and Zhang, Q.W. (2015) Design and Implementation of Auditing System Based on Database Security. *Journal of Zhengzhou Light Industry Institute* (*Natural Science Edition*), **30**, 69-74. http://dx.chinadoi.cn/10.3969/j.issn.2095-476X.2015.3/4.015

[51] Yu, J. (2019) Malicious Traffic Detection Based on TensorFlow and Convolution Neural Network. 12*th Academic Meeting of China Electrical Engineering Society Electric Power Communication Professional Committee*, Beijing, 12-15 November 2019, 406-410.

[52] Li, Y.F., Gong, B., Xu, D.W. and Le, J.J. (2018) Study on Database Compulsory Behavior Control in a Feasible Computing Environment. *Computer Application and Software*, **35**, 66-72.

[53] Ma, Z.C., Zhang, L., Yang, M.J., *et al.* (2015) Study on the Pattern Matching Technology and Its Application in Network Security. 2015 8*th International Conference on Intelligent Computation Technology and Automation*, Nanchang, 14-15 June 2015, 888-891.

[54] Xing, G.S. and Luo, H.W. (2019) Research on Web Injection Behavior Real-time Detection Technology Based on Deep Learning. *Network Security Technology and Application*, No. 7, 39-40. http://dx.chinadoi.cn/10.3969/j.issn.1009-6833.2019.07.024

[55] Hu, J.W., Zhao, W., Yan, Y. and Zhang, R. (2019) Analysis and Implementation of SQL Injection Vulnerability Mining Technology Based on Machine Learning. *Information Network Security*, **19**, 36-42. http://dx.chinadoi.cn/10.3969/j.issn.1671-1122.2019.11.005

[56] Ogbomon Uwagbole, S., Buchanan, W.J. and Fan, L. (2017) An Applied Pattern-Driven Corpus to Predictive Analytics in Mitigating SQL Injection Attack. 2017 7*th International Conference on Emerging Security Technologies*, Canterbury, 6-8 September 2017, 12-17. https://doi.org/10.1109/EST.2017.8090392

[57] Xiang, C.Z. and Li, X.M. (2011) Design of Database Security Agent Based on SSL. *Coal Technology*, **30**, 135-137.

[58] Fu, S.G. (2012) Strategies for the Construction of Database Security Defense System. *Coal Technology*, **31**, 174-175. http://dx.chinadoi.cn/10.3969/j.issn.1008-8725.2012.05.077

[59] Seok, J.B. and Sung, B.C. (2017) A Hybrid System of Deep Learning and Learning Classifier System for Database Intrusion Detection. 2017 *International Conference on Hybrid Artificial Intelligence Systems*, La Rioja, 21-23 June 2017, 615-625. https://doi.org/10.1007/978-3-319-59650-1_52

[60] Wang, Z.P. and Liu, C.G. (2011) Research and Implementation of Database Security Monitoring Technology. *Coal Technology*, **30**, 116-118.

[61] Li, R., Zhao, L. and Yang, J.W. (2010) An Application-Level Database Intrusion Detection Method. *Computer Applications and Software*, **27**, 280-283.

[62] Tang, Q.B., Yang, B. and Pan, L.M. (2019) Research on Anti-Scan Technology Based on Machine Learning. *Information Security Research*, **5**, 303-308.

[63] Zhang, X. (2018) Network Intrusion Detection Based on Machine Learning Algorithm. *Modern Electronic Technology*, **41**, 124-127.

[64] Lee, J.-H., Kim, I.K. and Han, K.J. (2015) An Abnormal Connection Detection System based on Network Flow Analysis. 2015 *IEEE* 5*th International Conference on Consumer Electronics Berlin*, Berlin, 6-9 September 2015, 71-75. https://doi.org/10.1109/ICCE-Berlin.2015.7391336

[65] Salimov, A.S., Dolgopolov, N.M., Sukhov, A.M. and Sagatov, E.S. (2018) Application of SDN Technologies to Protect Against Network Intrusions. 2018 *International Scientific and Technical Conference Modern Computer Network Technologies*, Moscow, 25-26 October 2018, 1-9. https://doi.org/10.1109/MoNeTeC.2018.8572127

[66] Subudhi, S., Kumar Behera, T. and Panigrahi, S. (2017) Use of OPTICS and Supervised Learning Methods for Database Intrusion Detection. 2017 3*rd International Conference on Computational Intelligence and Networks*, Odisha, 28 October 2017, 78-82. https://doi.org/10.1109/CINE.2017.10

[67] Liu, W. and Li, S.Y. (2019) Android Mobile Application Detection Research. *Computer Applications and Software*, **36**, 322-326.

[68] Chen, K., Zhang, Y.J. and Liu, P. (2018) Leveraging Information Asymmetry to Transform Android Apps into Self-Defending Code Against Repackaging Attacks. *IEEE Transactions on Mobile Computing*, **17**, 1879-1893. https://doi.org/10.1109/TMC.2017.2782249

[69] Lin, X.B., Lin, Y.P., Li, H. and Zhang, J.L. (2020) A Novel Method for Malware Detection on ML-based Visualization Technique. *Computers & Security*, **89**, Article ID: 101682. https://doi.org/10.1016/j.cose.2019.101682

[70] Ajay Kumara, M.A. and Jaidhar, C.D. (2018) Automated Multi-Level Malware Detection System Based on Reconstructed Semantic View of Executables Using Machine Learning Techniques at VMM. *Future Generation Computer Systems*, **79**, 431-446. https://doi.org/10.1016/j.future.2017.06.002

[71] Chandel, S., Ni, T.-Y. and Yang, G. (2018) Enterprise Cloud: Its Growth & Security Challenges in China. 2018 5*th IEEE International Conference on Cyber Security and Cloud Computing*/2018 4*th IEEE International Conference on Edge Computing and Scalable Cloud*, Shanghai, 22-24 June 2018, 144-152. https://doi.org/10.1109/CSCloud/EdgeCom.2018.00034

[72] Ahmed Khalaf, B. and Mostafa, S.A., Mustapha, A. Abed Mohammed, M. and Mustafa, A.W. (2019) Comprehensive Review of Artificial Intelligence and Statistical Approaches in Distributed Denial of Service Attack and Defense Methods. *IEEE Access*, **7**, 51691-51713. https://doi.org/10.1109/ACCESS.2019.2908998

[73] Liu, Z.T., Jin, H., Hu, Y.-C. and Bailey, M. (2018) Practical Proactive DDoS-Attack Mitigation via Endpoint-Driven In-Network Traffic Control. *IEEE/ACM Transactions on Networking*, **26**, 1948-1961. https://doi.org/10.1109/TNET.2018.2854795

[74] Ali, S. and Li, A. (2019) Learning Multilevel Auto-Encoders for DDoS Attack Detection in Smart Grid Network. *IEEE Access*, **7**, 108647-108659. https://doi.org/10.1109/ACCESS.2019.2933304

[75] Mohammad-Mahdi, B., Thibaut, S., Marc, L., Südholt, M. and Menaud, J.-M.(2018) Cache-Based Side-Channel Attacks Detection through Intel Cache Monitoring Technology and Hardware Performance Counters. 2018 3*rd International Conference on Fog and Mobile Edge Computing*, Barcelona, 23-26 April 2018, 1-6.

[76] Rajendran, J., Karri, R., Wendt, J.B., Potkonjak, M., McDonald, N., Rose, G.S. and Wysocki, B. (2015) Nano Meets Security: Exploring Nano Electronic Devices for Security Applications. *Proceedings of the IEEE*, **103**, 829-849. https://doi.org/10.1109/JPROC.2014.2387353