Research on Anomaly Detection in Massive Multimedia Data Transmission Network Based on Improved PSO Algorithm

LIQIANG GUO
School of Educational Science, Luoyang Normal University, Luoyang 471934, China
e-mail: hongzhi3610z@163.com

This work was supported by the training plan for the outstanding young college teachers of Henan Province "The research on barriers and mechanism for the effective learning of online courses" (Grant No. 2019GGJS201).

ABSTRACT With the development of computer network technology and the expansion of network system, sensitive data is facing the threat of hacker attack. Intrusion detection is an active network security defense measure, which is an attempt to invade, an ongoing intrusion or an intrusion that has occurred to identify the process. At present, the detection rate of intrusion detection method is low, the false alarm rate and false alarm rate is high, and the real-time performance is poor. It needs a large number of or complete data to achieve better detection performance. In this paper, the concept, characteristics, classification, research contents and difficulties of traditional intrusion detection for mass multimedia data transmission network are described. Then, the basic principle of neural network and particle swarm optimization (PSO) algorithm and the basic idea of particle swarm optimization algorithm with quantum (QPSO) behaviour are introduced. It is emphasized that QPSO has better convergence performance than PSO algorithm in global optimization problems. In this paper, the concept, characteristics and structure of neural network are described, and the algorithm and classification of wavelet neural network are introduced. Then taking wavelet neural network (WNN) as the object, using the QPSO algorithm as the training algorithm, the concrete operation process is given. The research work in this paper shows that the performance of neural network trained by QPSO algorithm and improved QPSO algorithm is better than that of other intelligent algorithms such as PSO algorithm and genetic algorithm, and the convergence speed is faster than that of PSO algorithm or GA algorithm. QPSO is a high performance neural network training algorithm, which can play a good role in neural network anomaly detection.

INDEX TERMS Anomaly detection, particle swarm optimization, data transmission, wavelet neural network.

I. INTRODUCTION
Computer technology has made a breakthrough in all aspects since the birth of the computer, people are no longer satisfied with the command line, text and other boring information representation, Instead, it must be the multimedia technology that integrates text, graphics, animation, sound and image, which is consistent with the way people perceive the natural world [1]. Especially with the popularization of network and the development of communication technology in recent years, multimedia technology has been widely used in various levels of practical applications [2]. Due to the characteristics of multimedia information itself, the effective management of mass multimedia data has been a hot spot in the development of computer technology in recent years [3]. This paper will systematically introduce a massive multimedia data management system based on distributed communication network. Classic network architecture is shown in Figure 1 [4], [5].

With the rapid development of network communication technology, the combination of multimedia technology and network technology is the inevitable trend of the development of multimedia information management technology [6]. Using distributed network technology to separate the functions of multimedia information management from
geographical location will help to expand the application scope of multimedia data management system, spread the load of data processing, and improve the data processing ability of the system [7]. It can increase the ability of parallel computing and improve the effectiveness and stability of the system. Distributed multimedia mass data management system model but in the distributed environment, the design of the system will have to consider the key issues such as data compression and decompression, network load balancing, information synchronization and effective management of multimedia information [8]. Packet transmission rates are getting faster and faster. This requires that the response speed of intrusion detection system must be improved, and the accuracy should be improved accordingly [8]. But due to the misuse or abnormal intrusion detection technology, it is necessary to compare each data with a large number of patterns in the signature database, so the efficiency of both is not high. Therefore, how to distinguish the abnormal state from the normal state and improve the accuracy and rapid response of intrusion detection system has become a practical problem [9].

II. IMPROVED PSO ALGORITHM

A. PARTICLE SWARM OPTIMIZATION

Particle swarm optimization (PSO) was proposed by American social psychologist James Kennedy and electrical engineer Russell Eberhart in 1995 [10]. The basic idea was inspired by their early findings on the behavior of bird populations and the use of biologist Frank Heppner’s biological population model. PSO is another swarm intelligence algorithm in the field of computational intelligence besides ant colony optimization. It is similar to genetic algorithms [11]. Global search is realized by cooperation and competition of individual questions [12]. The system is initialized as a set of random solutions called particles. Through the particle flight in the search house, it does not have the crossover and mutation operator of genetic algorithm, but the particle in the solution space to follow the optimal particle search [13].

Since PSO was put forward, due to its computational speed and easy realization of the algorithm itself, it has attracted the attention and research of many scholars in the related fields in the world. The research can be divided into three parts: algorithm itself, parameter selection, topology, fusion with other evolutionary techniques and applications [14]. Initially, PSO was designed for real value problems. Later, the algorithm is extended to binary PSO and discrete problem, which provides a useful way to compare the performance of PSO algorithm and genetic algorithm. This method can be used in the structure optimization of neural network [15].

PSO algorithm is derived from the simulation of predation behavior of birds. Imagine a situation where a group of birds are randomly searching for food, and there is only one food in this area, and all birds don’t know where the food is, but they know how far away they are from the food [22]. So what is the best strategy for finding food? The simplest and most effective is to search the surrounding area of the bird that is currently closest to the food [23]. PSO draws inspiration from this model and uses it to solve optimization problems.

In the PSO model, the solution of each optimization problem is the state of a “particle” in the search space. Each particle has a fitness value (fitness value), determined by the optimized function and a velocity that determines the direction and distance of their flight [24]. Particles can be dynamically adjusted according to their own flight experience and that of their peers, which can be said to update themselves by tracking two positions [25]. One is the optimal solution p-best found by the particle itself. That is, the best position of the individual; the other is the optimal solution ghost, which is the global best position, which is found by the whole population [26].

Particle swarm optimization (PSO) is a general heuristic search technique. The basic idea of PSO is to find the optimal solution through cooperation and information sharing among individuals in the population [27]. The search process of PSO is to initialize a group of random particles first. Each particle has its own position and initial velocity (which determines the direction and distance of flight), assuming that in a D-dimensional target search space, m particles are randomly generated to form a community. Then the position of the f particle in the D dimensional search space is vector [28].

\[ X_i^j = [x_{i1}^j, \ldots, x_{iD}^j], \quad (i = 1, 2, \ldots, m) \quad (1) \]

In which, \( T \) is the number of iterations, and its individual extremum is recorded as follows,

\[ P_i = [p_{i1}, \ldots, p_{iD}]^T \quad (2) \]

The global extremum is recorded as the following expression,

\[ P_g = [p_{g1}, \ldots, p_{gD}]^T \quad (3) \]

During the reception, the \( t \) particle flies in the search space with a velocity \( V_i \), and the velocity and position of each particle are modified according to the following formula [29]:

\[ v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 [P_{id} - x_{id}^t], \quad 1 \leq i \leq m, \quad 1 \leq d \leq D \quad (4) \]
In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_{ij}^{t+1} = \begin{cases} 
  v_{ij}, & \text{if } \text{randm} \leq C \text{ or } j = \text{randn}(i) \\
  x_{ij}, & \text{if } \text{randm} \geq C \text{ or } j \neq \text{randn}(i) 
\end{cases} 
\]  

(7)

In which, subscript \( i \) represents the first individual, and subscript \( j \) represents the \( j \) dimensional random number \( m(j) \in 0 \) of the intermediate generation vector \( \vec{v} \) (or population vector \( \vec{x} \)) as a uniformly distributed random number, \( \text{randn} \) is a random selection exponent, which ensures that \( n \) can obtain at least one parameter from \( \vec{v} \), and \( C_i \in 0 \) is a crossover probability.

In the basic differential evolution algorithm, the selection operation adopts greedy strategy, that is, only when the pre-generational individual is better than the parent individual, otherwise the parent individual will be retained to the next generation. Differential evolution algorithm flow chart is shown in Figure 3.

Differential crossover operations are designed to increase the diversity of new populations by exchanging parts of code between old and new individuals according to the crossover strategy, thus forming new individuals. There are usually two crossover operators. The first is the binomial crossover operator, the second is the exponential crossover operator. In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_i(t + 1) = x_{i3}(t) + F(x_{i1}(t) - x_{i2}(t)) 
\]  

(6)

In which, subscript I represents the first individual, subscript \( t \) represents the algebra of evolution, \( x_{i1}, x_{i2}, x_{i3} \) is a random selection of three different individuals from the evolutionary population, and \( F \) is a scaling factor [33]. It is used to control the influence of the difference vector \( F \in [0, 1] \).

Differential evolution algorithm (DEA) is a stochastic parallel direct search algorithm. Its basic idea is that starting from a random initial population, we can get the intermediate population by using the difference between the individuals of the current population [31]. Then a generation of population is obtained by the competition of fitness value between father and son, and the global search of the solution space is realized. There are three operations running through the whole process of the algorithm, namely, mutation, crossover and selection. The differential mutation operation can generate a new individual from the parent population to the intermediate generation by the following formula [32]:

\[
x_id^{t+1} = x_{id}^t + v_id^t, \quad 1 \leq i \leq m, \ 1 \leq d \leq D 
\]  

(5)

In which, \( v_t \) denotes the velocity after the \( t \)-th iteration, \( c1 \) and \( c2 \) are two constants, \( r1 \) and \( r2 \) are random numbers between \([0, 1]\) and \( 0.2 \), \( w \) is the inertial factor, and generally between 0.1 and 0.9. Flow chart of particle swarm optimization algorithm is given in Figure 2 [30].

**FIGURE 2.** Flow chart of particle swarm optimization algorithm.

**B. APPLICATION OF PARTICLE SWARM OPTIMIZATION IN INTRUSION DETECTION**

In recent years, although the intrusion detection system (IDS) has been a relatively large development, the overall situation is not satisfactory [16]. The biggest problem is that the current intrusion detection product detection accuracy is relatively low, false positives and false reports are more [17]. In order to improve the monitoring accuracy and speed of IDS, some other concepts and methods have been introduced into the IDS. Anomaly detection based on intelligent algorithm has become a hot research topic in the near future [18], [19]. Some intelligent detection techniques such as neural network (ANN), support vector machine (SVM), artificial genetic, multi-agent, hidden Markov chain model (HMM) and so on have been developed. Optimization theory and method with high convergence and stability has become a new research field and development direction [20]. In this case, some scholars consider the shortcomings of these algorithms, such as complex computation, slow convergence and so on, and consider using particle swarm optimization to overcome these shortcomings. Not only is it getting more and more attention. And it has begun to play a key role in different environments [21].

**C. IMPROVED PSO ALGORITHM DESIGN**

In the basic differential evolution algorithm, the selection operation adopts greedy strategy, that is, only when the pre-generational individual is better than the parent individual, otherwise the parent individual will be retained to the next generation. Differential evolution algorithm flow chart is shown in Figure 3.

Differential crossover operations are designed to increase the diversity of new populations by exchanging parts of code between old and new individuals according to the crossover strategy, thus forming new individuals. There are usually two crossover operators. The first is the binomial crossover operator, the second is the exponential crossover operator. In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_i(t + 1) = x_{i3}(t) + F(x_{i1}(t) - x_{i2}(t)) 
\]  

(6)

In which, subscript I represents the first individual, subscript \( t \) represents the algebra of evolution, \( x_{i1}, x_{i2}, x_{i3} \) is a random selection of three different individuals from the evolutionary population, and \( F \) is a scaling factor [33]. It is used to control the influence of the difference vector \( F \in [0, 1] \).

Differential crossover operations are designed to increase the diversity of new populations by exchanging parts of code between old and new individuals according to the crossover strategy, thus forming new individuals. There are usually two crossover operators. The first is the binomial crossover operator, the second is the exponential crossover operator. In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_{ij}^{t+1} = \begin{cases} 
  v_{ij}, & \text{if } \text{randm} \leq C \text{ or } j = \text{randn}(i) \\
  x_{ij}, & \text{if } \text{randm} \geq C \text{ or } j \neq \text{randn}(i) 
\end{cases} 
\]  

(7)

In which, subscript \( i \) represents the first individual, and subscript \( j \) represents the \( j \) dimensional random number \( m(j) \in 0 \) of the intermediate generation vector \( \vec{v} \) (or population vector \( \vec{x} \)) as a uniformly distributed random number, \( \text{randn} \) is a random selection exponent, which ensures that \( n \) can obtain at least one parameter from \( \vec{v} \), and \( C_i \in 0 \) is a crossover probability.

In the basic differential evolution algorithm, the selection operation adopts greedy strategy, that is, only when the pre-generational individual is better than the parent individual, otherwise the parent individual will be retained to the next generation. Differential evolution algorithm flow chart is shown in Figure 3.

Differential crossover operations are designed to increase the diversity of new populations by exchanging parts of code between old and new individuals according to the crossover strategy, thus forming new individuals. There are usually two crossover operators. The first is the binomial crossover operator, the second is the exponential crossover operator. In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_i(t + 1) = x_{i3}(t) + F(x_{i1}(t) - x_{i2}(t)) 
\]  

(6)

In which, subscript I represents the first individual, and subscript \( t \) represents the algebra of evolution, \( x_{i1}, x_{i2}, x_{i3} \) is a random selection of three different individuals from the evolutionary population, and \( F \) is a scaling factor [33]. It is used to control the influence of the difference vector \( F \in [0, 1] \).

Differential crossover operations are designed to increase the diversity of new populations by exchanging parts of code between old and new individuals according to the crossover strategy, thus forming new individuals. There are usually two crossover operators. The first is the binomial crossover operator, the second is the exponential crossover operator. In this paper, the binomial crossover operator is chosen. The crossover strategy can be expressed by the following formula:

\[
x_{ij}^{t+1} = \begin{cases} 
  v_{ij}, & \text{if } \text{randm} \leq C \text{ or } j = \text{randn}(i) \\
  x_{ij}, & \text{if } \text{randm} \geq C \text{ or } j \neq \text{randn}(i) 
\end{cases} 
\]  

(7)

In which, subscript \( i \) represents the first individual, and subscript \( j \) represents the \( j \) dimensional random number \( m(j) \in 0 \) of the intermediate generation vector \( \vec{v} \) (or population vector \( \vec{x} \)) as a uniformly distributed random number, \( \text{randn} \) is a random selection exponent, which ensures that \( n \) can obtain at least one parameter from \( \vec{v} \), and \( C_i \in 0 \) is a crossover probability.

In the basic differential evolution algorithm, the selection operation adopts greedy strategy, that is, only when the pre-generational individual is better than the parent individual, otherwise the parent individual will be retained to the next generation. Differential evolution algorithm flow chart is shown in Figure 3.
Fitness function:
\[
\max F = \sum_{i=1}^{12} N (t) \cdot \Delta t
\]  
\[
N (t) = A Q(t) H(t)
\]
\[
Q(t) = \frac{(V(t) - V(t - 1))}{\Delta t} - q(t)
\]
\[
H(t) = S(t) - S(t - 1)
\]

Constraint condition:
\[
V(t + 1) = V(t) + F(t) - Q(t) - S(t)
\]
\[
V_{\text{min}}(t) \leq V(t) \leq V_{\text{max}}(t)
\]
\[
Q_{\text{min}}(t) \leq Q(t) \leq Q_{\text{max}}(t), \quad S(t) \geq 0
\]
\[
N_{\text{min}}(t) \leq N(t) \leq N_{\text{max}}(t)
\]

In the previous formula, \(P_i\) represents the first particle, \(m\) represents the size of the particle swarm, that is, the number of particles in the swarm, \(V_{\text{(1)}}\) represents the storage capacity of the \(t\) scheduling period, \(t\) is the dimension of the solution vector, that is, the scheduling period (12 months). \(N\) denotes the number of network nodes. \(A\) represents the throughput coefficient of the network node, \(Q_{\text{(1)}}\) represents the network traffic in the \(t\) scheduling period, \(H_{\text{(1)}}\) represents the network data. \(Q_{\text{(1)}}\) for the \(t\) scheduling period and represents the network traffic for the \(t\) scheduling period. \(S_{\text{(1)}}\) represents the network capacity of the \(t\) scheduling period, \(F(t), Q_{\text{(1)}}\), and \(S_{\text{(1)}}\) respectively representing the number of uploads, downloads and the number of packets lost in the network during the \(t\) period. \(Q_{\text{(1)}}\) represents the traffic, \(V_{\text{min}}(t)\) of the \(f\)th-scheduled network. \(V_{\text{max}}(t), Q_{\text{min}}(t), Q_{\text{max}}(t), N_{\text{min}}(t)\) and \(N_{\text{max}}(t)\) represent the lower upper limit of the storage capacity, download and network capacity of the network in the \(t\)-th scheduling period, respectively.

III. ANOMALY DETECTION IN MASSIVE MULTIMEDIA DATA TRANSMISSION NETWORK

A. REQUIREMENTS FOR MULTIMEDIA DATA

Mass multimedia data processing includes the analysis, conversion, calculation, storage, fast downloading and accessing of massive multimedia data. Usually, the data scale can reach TB or even PB level. The main contents are as follows: (1) high resource consumption, (2) high cost of system construction, (3) high requirement of QOS, (4) read-only multimedia data. Once media such as image, audio, video and animation enter the storage system, they are seldom modified.

Due to the large amount of information, unstructured and dynamic characteristics of multimedia data itself, traditional database systems (including relational, hierarchical, network) cannot meet the requirements of effective management of multimedia data. There are two important development directions in multimedia data management. The security and transaction integrity of relational database is the basis of multimedia extension. Security is expressed in the form of domain-level, table-level and database-level access control; transactional integrity refers to the synchronization of the affected parts of the database when the transaction is backed up.

Cloud storage system has the characteristics of massive storage and computing resources, high reliability, high expandability and low cost. It consists of hundreds or thousands of cheap storage machines. Continuous monitoring, error detection, fault tolerance and automatic recovery are essential components of the system. Each storage machine is an independent storage node, and the addition and separation of nodes do not affect the normal operation of the system. The amount of data in the file is very large in the traditional sense, and most of the file changes are the operation of adding new data instead of overwriting the existing data. These characteristics show that cloud storage system has a unique advantage in processing massive multimedia data.

In the distributed file system of Google, the data file is divided into blocks of fixed size. Each block is marked by a globally unique and invariant block handle. The block handle is allocated by the system when the block is created. The block of files is stored in the local disk of each data node. Read and write blocks through block handles and byte ranges. To ensure the reliability of the data, each block is stored on different data nodes in the form of multiple replicas. Table 1 is an architecture of a distributed file system in a Google cloud storage system.

This storage feature of files is very suitable for large-scale parallelization of massive multimedia data. The task of data processing is assigned to the node where the data is stored, and the advantage of transferring computing to storage in cloud computing is brought into play. If we want to process multimedia data files in parallel on a large scale, we need to store the files in a specific format to eliminate the dependencies between them, so that we can process the data in a file.
TABLE 1. Sorting NoSQL with persistent storage.

| Database       | Durable design       |
|----------------|----------------------|
| Cassandra      | Memtable / SSTable   |
| CouchDB        | Append B tree        |
| HBase          | Memtable / SSTable on HDFS |
| MongoDB        | B tree               |
| Neo4j          | Disk link list       |
| Redis          | RAM                  |
| Riak           | uncertain            |
| calars         | RAM                  |
| Tokyo Cabinet  | Hash or B-tree       |
| Voldemort     | Pluggable (mainly BDB MySQL) |

TABLE 2. Dynamo use technology and benefits.

| Problem                        | Solution            | Advantage                        |
|--------------------------------|---------------------|----------------------------------|
| Partition                      | Consistency hash    | Improve scalability              |
| High-availability write operations | Vector Clock       | Version decoupling               |
| Handling temporary failure     | Loose election mechanism | High availability and durability |
| Recover from persistent failure | Anti-entropy        | Background consistency replication |
| Failure detection              | Heartbeat protocol  | Avoid centralized registration   |

block. There is no need to read another file block of data. The development of cloud computing brings both opportunities and challenges to large-scale parallel processing of massive multimedia data. And each file block and its copy are stored in different nodes in the cloud system, so the optimal scheduling of tasks is required on a large number of nodes. In order to achieve the goal of load balancing and increasing the utilization rate of system resources, the current mechanism of task scheduling and resource allocation in cloud systems is a static local scheduling mechanism. Considering only the current resource state of the nearest node to the data, although it has the advantage of transferring computation to storage, the optimal allocation of resources has not been made from the global point of view of the system. Dynamo use technology and benefits are given in Table 2.

B. JOB MODEL FOR MASSIVE MULTIMEDIA DATA PROCESSING

User requirements in multimedia applications are submitted in the form of job description files to the client of the cloud storage system. When a job is submitted, the client splits the job into several tasks for parallel execution based on the file block information. Each task is associated with a file block, so that different tasks in the same job can be distributed to different nodes for execution. It can avoid the bottleneck caused by resource competition among tasks and adapt to the characteristics of high resource consumption in the process of massive multimedia data processing. The file block information is described as follows.

$$T_{tree} = (C, R)$$  \hspace{1cm} (18)

In which, $C$ is a collection of data elements where the elements represent a block of files; $R$ is a binary relational set in which the elements represent whether the contents of the two chunks are the same. The element $c_i$ in the data element set $C$ can be described as a binary group vector:

$$c_i = \langle \text{chunk}_i, a_i \rangle$$ \hspace{1cm} (19)

In which,.chunk_id represents the number of block I and $a_i$ indicates the number of nodes where block I. The job stream submitted to the client can be described as a random time when the job is submitted to the client of the cloud storage system

$$J = \{J_0, J_1, J_2, \ldots, J_n\}$$ \hspace{1cm} (20)

In which, $J_i$ can be described by a six-tuple vector:

$$j_i = \langle \text{job}_{id}, \text{dec}_{id}, \text{file}_{id}, \text{level}_{i}, \text{rcv}_t, \text{end}_t \rangle$$ \hspace{1cm} (21)

In which, job_id is the job number, dec_id is the job description file number, file_id is the multimedia data file number that the job needs to process, level is the priority of the job, and rcv_t is the job submission time. end_t is the deadline for completing the job. The submitted job J_i is split into several tasks by the client. The split task can be described as a directed acyclic graph DAG:

$$DAG = \langle V, E \rangle$$ \hspace{1cm} (22)

For CPU resource consumption, multiple computationally intensive tasks run on one CPU core at the same time, resulting in a sharp reduction in execution speed. Therefore, you need to limit the number of computation-intensive tasks on a processor. Therefore, CPU resources are also abstracted into quantities to measure. Abstract a processor into four CPU processor. Chink_id is the file block number to be processed by the task. Each element in the edge set $E$ represents a dependency between the two tasks. $E_{ij}$ indicates that the task $vi$ is the successor of the task $vj$. Task $v_j$ is the successor of task $v_i$. Each task can only be executed if all precursor tasks have been completed.

C. TASK SCHEDULING AND RESOURCE ALLOCATION MODEL FOR MASSIVE MULTIMEDIA DATA PROCESSING

Combined with the characteristics of the job model of massive multimedia data processing. In our proposed task scheduling and resource allocation model, each data node has a scheduling process. These scheduling processes know the current state of the virtual machine of the data node in which they are located, and the scheduling processes of different nodes are able to communicate with the client’s scheduling process. The client scheduling process divides the
A. Clustering Scheduling Algorithm to Reduce the Selection Range of Resources and Improve the Efficiency of Task Scheduling

The clustering scheduling algorithm is designed to cluster nodes according to the queried multimedia file block information. For each query, the system first clusters the data nodes according to the binary relation set $R$, the clustering rules are as follows:

$$a_i \in D_m, \quad D_m \subseteq D$$

(23)

All the data nodes that store the same file block $m$ form a cluster set $D_m$ which is the set of all nodes in the system. $A_i$ is a node in the cluster set $D_m$. The scheduling process of the client selects one of the best nodes in $D_m$ to perform the task from the point of view of improving resource utilization and load balancing. The scheduling process is only within a limited range of local nodes. Within a task for resource allocation and task scheduling, it is the goal of clustering scheduling algorithm to reduce the selection range of resources and improve the efficiency of task scheduling. The whole framework of trust routing algorithm based on QoS is given in Figure 4.

FIGURE 4. The whole framework of trust routing algorithm based on QoS.

IV. EXPERIMENTAL SIMULATION

In the intrusion detection based on network behaviour, the data set collected by DARPA in 2018 for Mr Linco In Labs intrusion detection evaluation is used as experimental data. The data was also used in the third International Competition on knowledge Discovery and data mining tools, which was co-sponsored by the Fifth International Conference on knowledge Discovery and data mining. Therefore, these data are also referred to as KDDCUP99 data in many literatures. The original data set collected by DARPA in 2018 for Mr Linco In Labs imitated the Air Force LAN. The data was collected nine weeks after MIT Lincoln Labs.

B. DATA SET DESCRIPTION

The training data in KDDCUP99 contains seven weeks of network traffic and about 40 gigabytes of compressed binary data. After processing, the data is restored to about 50000 connection records, each of which is a network packet at a certain time. A transmission drama 551 via some protocol from the source IP address to the target IP address. In the experiment, 10% KDDCUP99 is used as the data set. The dataset contains 22 different attack methods, 494021 connection records, including 97278 normal connection records, each connection record has 41 different attributes. Among them, 34 are digital and 7 are character types. The specific intrusion data are described in the table 3 as follows. The test sample data contains 311029 connection records and 37 different attack methods, 17 of which have no specific

A. PREPROCESSING OF NETWORK CONNECTION INFORMATION IN RAW DATA

The data collection and data processing of KDDCUP99 are mainly completed by Lee et al. They designed an intrusion detection model for network connection information. This model reconstructs the packets that the system listens to through the network interface to form a network connection information set. Then the intrusion detection model is built directly on the network connection information set according to the method of machine learning. The original audit records, such as the network packets collected by Tcpcdmp, are first converted into packet information in the form of ASCII characters, and such packet information is then processed to recover the connection session records contained therein. Finally, the feature set and classification model suitable for intrusion detection are established on the connected session set.

The data obtained from the network interface is the network traffic recorded by network listening software such as Tcpcdmp. The traditional intrusion detection method uses rule-based pattern matching to detect attacks directly on this data. Lee et al. restored network traffic data to network connection session information through network connection information pre-processing. This process is divided into three steps, which can be represented by graph:

Step 1: collect all network traffic data at network interface with Tcpcdmp and other tools.

Step 2: the binary data collected by Tcpcdmp is transformed into packet level information represented by ASCII characters, each line of which represents the information of one packet.

Step 3: processing the packet information and generating the connection record set. In this method, each TCP session is treated as a connection, and each UDP packet is treated as a session, in which the records of each connection contain the following basic properties: time stamp, duration, source address, destination address, service, etc. These basic properties represent the basic information for each session.

The software used in the first and second steps of connection information pre-processing is Tcpcdmp. The third step is a segment of processing program written by Bro, the script language of intrusion detection system. This program scans the record set in the form of characters and reorganizes the session information. And statistics some statistics, such as the number of rejected connections, the number of repeated ACK packets, error packet size ratio, and so on. The resulting connection information has some basic properties such as duration, service, and connection flag and so on.

B. DATA SET DESCRIPTION

The training data in KDDCUP99 contains seven weeks of network traffic and about 40 gigabytes of compressed binary data. After processing, the data is restored to about 500000 connection records, each of which is a network packet at a certain time. A transmission drama 551 via some protocol from the source IP address to the target IP address. In the experiment, 10% KDDCUP99 is used as the data set. The dataset contains 22 different attack methods, 494021 connection records, including 97278 normal connection records, each connection record has 41 different attributes. Among them, 34 are digital and 7 are character types. The specific intrusion data are described in the table 3 as follows. The test sample data contains 311029 connection records and 37 different attack methods, 17 of which have no specific
intrusion data description in the training sample as shown in table 4 below. These attacks can be classified into four categories:

DoS: denial-of-service attacks, such as smurf, teardrop attacks;
R2L: access to remote permissions, such as password guessing;
U2R: privilege escalation, such as various buffer overflow attacks;
PROBE: IP address, port and vulnerability scan.

C. EXPERIMENTAL DATA PREPROCESSING

In the application of neural networks, variables must be scaled to the range allowed by the transfer function of the processing unit, so the training sample data and the test data of the input network must be converted in the same proportion. The dimensions of each attribute of the KDDCUP99 dataset are not exactly the same, and the range of values of attribute items varies greatly. For example, the range of values for src-bytes and dst-bytes is [0, 2194619] and [0, 51555465], respectively. The range of values of wrong-fragment and urgent is [0, 3] and [0, 3] respectively. Therefore, it is necessary to normalize all the values of data on the same attribute before the wavelet neural network is input.

The most commonly used linear conversion method is generally used. If the maximum and minimum values of variables are $X_{\text{max}}$ and $X_{\text{min}}$, respectively. And the actual limits of the network are $A_{\text{max}}$ and $A_{\text{min}}$, the variables $X$ can be transformed by the following formula:

$$ A = r \left( X - X_{\text{min}} \right) + A_{\text{min}} $$

(24)

In which,

$$ r = \frac{A_{\text{max}} - A_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} $$

(25)

For the output value $A$ of the network, the following can be converted to a variable $X$:

$$ X = \frac{A - A_{\text{min}}}{r} + X_{\text{min}} $$

(26)

In this experiment, genetic algorithm, QPSO algorithm and standard PSO algorithm are used as the training algorithms of wavelet neural network respectively. Finally, the established model is used to detect the network anomalies. Now the parameters of each model are set: set the crossover probability $P_c = 0.75$, mutation probability of GA, such as the acceleration constant $C(1)$ $C(2)$ of the GA algorithm, and the inertial weight $CO$ decreases linearly from 0.8 to 0.4. The contraction expansion coefficient of QPSO algorithm decreases linearly from 0.7 to 0.2. The population size of GA, PSO algorithm and QPSO algorithm are all 50, the iterations are all 1500, the weight value of $v_{jk}$ is defined in the range of ($-1, 0$), and the range of scaling coefficient $Q$ and translation coefficient is (1, 100). Statistical results of PSO algorithm performance are given in Table 5.

The range of the scaling factor $Q$ and translation coefficient of the data is the statistical result shown in Figure 5.

In addition, the performance and computation time of different methods for detecting anomalies in massive network data are shown in Table 6.

It can be seen from Figure 5 that as the cost increases, the performance of the three anomaly detection algorithms will become better and better. The improved PSO algorithm proposed in this paper has achieved the best results in anomaly detection in mass multimedia data transmission networks. This is because this paper designs an effective algorithm for abnormal phenomena in mass multimedia data transmission network, and improves the PSO algorithm to further improve the detection accuracy. Finally, the improved PSO algorithm proposed in this paper achieves the best performance.

In order to test the application performance of this method in information security risk assessment, the simulation
experiment is carried out, and the algorithm is designed and implemented based on Matlab Simulink simulation software. In the information security risk assessment algorithm, the information security risk collection is first carried out. There are 30 data acquisition nodes in the data acquisition server, the number of neural network nodes for information security assessment is 200, and the number of access nodes is 10 clusters. Each cluster has 5 nodes with a central frequency of 15 KHZ and a time parameter. According to the above simulation and parameter setting, the information security evaluation party is analyzed, and the information collection results are shown in Figure 7.

Evaluate the information security of the information collection sample in figure 4, the correlation compensation and the adaptive control are taken according to the information security risk data spectrum characteristic extraction result, the gray neural network is used to carry on the adaptive learning, the information security evaluation is obtained. The convergence curve of the safety assessment is shown in Figure 8.

Figure 8 shows that the proposed method can converge to zero in a short time, and test the accuracy of different
methods for information security evaluation. The comparison results are shown in Figure 9. It concludes that the accuracy of information security assessment by this method is high.

V. CONCLUSION

As an important part of network security research, intrusion detection has attracted the attention of scholars. At present, the detection rate of intrusion detection method is low, the false alarm rate and false alarm rate are high, and the real-time performance is poor. It needs a large number of or complete data to achieve better detection performance. In this paper, the concept, characteristics, classification, research contents and difficulties of traditional intrusion detection for mass multimedia data transmission network are described. Then, the basic principle of neural network and particle swarm optimization algorithm and the basic idea of particle swarm optimization algorithm with quantum behaviour are introduced. It is emphasized that QPSO has better convergence performance than PSO algorithm in global optimization problems. In this paper, the concept, characteristics and structure of neural network are described, and the algorithm and classification of wavelet neural network are introduced. Then taking wavelet neural network as the object, using the QPSO algorithm as the training algorithm, the concrete operation process is given. The research work in this paper shows that the performance of neural network trained by QPSO algorithm and improved QPSO algorithm is better than that of other intelligent algorithms such as PSO algorithm and genetic algorithm, and the convergence speed is faster than that of PSO algorithm or GA algorithm. QPSO is a high performance neural network training algorithm, which can play a good role in neural network anomaly detection.

REFERENCES

[1] Z. Zheng, A. K. Sangaiah, and T. Wang, “Adaptive communication protocols in flying ad hoc network,” IEEE Commun. Mag., vol. 56, no. 1, pp. 136–142, Jan. 2018.
[2] Z. Zheng, T. Huang, H. Zhang, S. Sun, J. Wen, and P. Wang, “Towards a resource migration method in cloud computing based on node failure rule,” J. Intell. Fuzzy Syst., vol. 31, no. 5, pp. 2611–2618, Oct. 2016.
[3] Z. Zheng, H.-Y. Jeong, T. Huang, and J. Shu, “KDE based outlier detection on distributed data streams in multimedia network,” Multimedia Tools Appl., vol. 76, no. 17, pp. 18027–18045, Sep. 2017, doi: 10.1007/s11454-016-3681-y.
[4] S.-X. Pan, W.-J. Sun, and Z. Zheng, “Video segmentation algorithm based on superpixel link weight model,” Multimedia Tools Appl., vol. 76, no. 19, pp. 19741–19760, 2017, doi: 10.1007/s11454-016-3439-6.
[5] M. Chahal, S. Harit, K. K. Mishra, A. K. Sangaiah, and Z. Zheng, “A survey on software-defined networking in vehicular ad hoc networks: Challenges, applications and use cases,” Sustain. Cities Soc., vol. 35, pp. 830–840, Nov. 2017, doi: 10.1016/j.scs.2017.07.007.
[6] Z. Zheng, P. Wang, S. Sun, and J. Liu, “Real-time big data processing framework: Challenges and solutions,” Appl. Math. Inf. Sci., vol. 9, no. 6, p. 3169, 2015.
[7] S. Heintz et al., “Psychometric comparisons of benevolent and corrective humor across 22 countries: The virtue gap in humor goes international,” Frontiers Psychol., vol. 9, p. 92, Feb. 2018.
[8] R. K. Sahu, S. Panda, and G. T. C. Sekhar, “A novel hybrid PSO-PS optimized fuzzy PI controller for AGC in multi area interconnected power systems,” Int. J. Electr. Power Energy Syst., vol. 64, pp. 880–893, Jan. 2015.
[9] M. Sharafi and T. Y. ELMekawy, “Multi-objective optimal design of hybrid renewable energy systems using PSO-simulation based approach,” Renew. Energy, vol. 68, pp. 67–79, Aug. 2014.
[10] E. T. Mohamad, D. J. Armaghani, E. Momeni, and S. V. A. N. K. Abad, “Prediction of the unconfined compressive strength of soft rocks: A PSO-based ANN approach,” Bull. Eng. Geol. Environ., vol. 74, no. 3, pp. 745–757, Aug. 2015.
[11] F. Fan, Z. Liu, X. Mu, S. Gao, and Y. Wang, “Research and analysis of transient process of locomotive passing neutral section based on Hasedabak arc model,” in Proc. Int. Conf. Electr. Infotech. Rail Transp. (EITRT), vol. 1, no. 14, Berlin, Germany: Springer, Jan. 2014, pp. 233–241, doi: 10.1007/978-3-642-53778-3_22.
[12] Z. Qin, J. Ruan, D. Huang, and Y. Huang, “Mechanical fault diagnosis of high voltage outdoor circuit breaker based on motor current signal analysis,” in Proc. Int. Conf. Power Syst. Technol., Oct. 2014, pp. 1193–1198, doi: 10.1109/POWERCON.2014.6993501.
[13] I. Aydin et al., “Particle swarm based arc detection on time series in pantograph-catenary system,” in Proc. IEEE Int. Symp. Innov. Intell. Syst. Appl., 2014, pp. 23–25, doi: 10.1109/INISTA.2014.6873642.
[14] F. Jiang, H. Yang, and S. Wang, “Asymptotic stability of stochastic pantograph differential equations with Markovian switching,” J. Intell. Contr. Electron. Meas. Instrum., to be published. [Online]. Available: http://www.math.sci.ac.cn/ijse/index.php/jnao/about
[15] M. Karakose, O. Yaman, I. Aydin, and E. Karakose, “Real-time condition monitoring approach of pantograph-catenary system using FPGA,” in Proc. IEEE 14th Int. Conf. Ind. Informat. (INDIN), Jul. 2016, pp. 481–486, doi: 10.1109/SMIC.2013.337.
[16] F. A. Ghasemi, S. Daneshpayeh, and J. Ghasemi, “An investigation on the Young’s modulus and impact strength of nanocomposites based on polypropylene/linear low-density polyethylene/titan dioxide (PP/LLDPE/TiO2) using response surface methodology,” Polym. Bull., vol. 73, no. 6, pp. 1741–1760, Jun. 2016, doi: 10.1007/s00289-015-1574-2.
[17] C. Chrysouhou, A. Angelis, G. Tsisinakis, S. Spetsioti, I. Nasis, D. Tsiairis, P. Papakoulia, C. Pitsavos, N. G. Kouliouris, I. Vogiatzis, and T. Dimitriou, “Cardiovascular effects of high-intensity interval aerobic training: combined leg strength exercise in patients with chronic heart failure. A randomized phase III clinical trial,” Int. J. Cardiol., vol. 179, pp. 269–274, Jan. 2015.
[18] X. Zhang, J. Dong, Z. Huang, X. Nie, H. Hong, J. Wang, W. Sun, C. Leo, and J. Wang, “Study of soil structures strength and stiffness loss based on thermodynamics and continuum mechanics,” Environ. Earth Sci., vol. 73, no. 8, pp. 4143–4149, Apr. 2015, doi: 10.1007/s12665-014-3699-2.
[19] G. S. Cantrell, B. K. Schilling, M. R. Paquette, and Z. Murlin, “Maximal strength, power, and aerobic endurance adaptations to concurrent strength and sprint interval training,” Eur. J. Appl. Physiol., vol. 114, no. 4, pp. 763–771, Apr. 2014, doi: 10.1007/s00421-013-2811-8.
[20] Z. Zhi-Jing, C. Lin-Liang, and S. Hua, “Simulation of air traffic flow optimization prediction,” Comput. Simul., vol. 33, no. 8, pp. 54–57, Jan. 2016, doi: 10.3969/j.issn.1006-9348.2016.08.012.
[21] P. Frank, E. Andersson, M. Pontén, B. Eklom, M. Eklom, and K. Sahlin, “Strength training improves muscle aerobic capacity and glucose tolerance in elderly,” Scandin. J. Med. Sci. Sports, vol. 26, no. 7, pp. 764–773, Jul. 2016, doi: 10.1111/smss.12537.
[22] K. Harada and Y. Nakamura, “Awareness of role of strength training in care prevention, negative perception, and stages of change for strength training behavior among Japanese older adults,” Bull. Chem. Soc. Jpn., vol. 62, no. 12, pp. 3869–3876, Jan. 2015, doi: 10.1039/bj100640c.
[23] J. J. Rödersch, U. Füssel, C. Gresch, C. Großmann, and B. Hoffmann, “Data mining in resistance spot welding: A non-destructive method to predict the welding spot diameter by monitoring process parameters,” Int. J. Adv. Manuf. Technol., vol. 99, nos. 5–8, pp. 1085–1099, Nov. 2018.
[24] K. Rogulski, D. Wilk-Nowicka, J. Kaczprzyk, G. Gumieny, G. Rojek, and B. Mrzygłód, “Approximation of ausferrite content in the compacted graphite iron with the use of combined techniques of data mining,” Arch. Foundry Eng., vol. 17, no. 3, pp. 117–122, Sep. 2017.
[25] W. Liu, J. Rostami, and E. Keller, “Application of new void detection algorithm for analysis of feed pressure and transition pressure of roof bolters,” Int. J. Mining Sci. Technol., vol. 27, no. 1, pp. 77–81, Jan. 2017.
[26] K. Mathan, P. M. Kumar, G. Manogaran, R. Varadarajan, and P. Panchatcharam, “A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease,” Des. Autom. Embedded Syst., vol. 22, no. 9, pp. 225–242, 2018.
L. Guo: Research on Anomaly Detection in Massive Multimedia Data Transmission Network Based on Improved PSO Algorithm

[27] J. Rojas, C. E. M. Marin, P. A. Gaona García, J. Forero, and R. G. Crespo, “Analysis of physico-chemical variables and their influence on water quality of the Bogota river using data mining,” Int. J. High Perform. Syst. Archit., vol. 8, nos. 1–2, p. 3, 2018.

[28] J. Górecki, M. Hofert, and M. Holeňa, “An approach to structure determination and estimation of hierarchical Archimedean copulas and its application to Bayesian classification,” J. Intell. Inf. Syst., vol. 46, no. 1, pp. 21–59, Feb. 2016.

[29] Z. S. Pourtaghi, H. R. Pourghasemi, R. Aretano, and T. Semeraro, “Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques,” Ecol. Indicators, vol. 64, pp. 72–84, May 2016.

[30] P. Pinto et al., “Data mining and social Web semantics: A case study on the use of Hashtags and memes in online social networks,” IEEE Latin Amer. Trans., vol. 15, no. 12, pp. 2276–2281, 2017.

[31] Q. A. Kester, “Using formal concepts analysis techniques in mining data from criminal databases and profiling events based on factors to understand criminal environments,” in Computational Science and Its Applications—ICCSA (Lecture Notes in Computer Science), vol. 9790. 2016, pp. 480–496.

[32] D. Tien Bui, T. A. Tuan, H. Klempe, B. Pradhan, and I. Revhaug, “Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree,” Landslides, vol. 13, no. 2, pp. 361–378, Apr. 2016.

[33] Y. Han and F. Moutarde, “Analysis of large-scale traffic dynamics in an urban transportation network using non-negative tensor factorization,” Int. J. Intell. Transp. Syst. Res., vol. 14, no. 1, pp. 36–49, Jan. 2016.

LIQIANG GUO was born in May 1981. He received the master’s degree in computer application technology from Henan University. He is currently working as an Associate Professor with the School of Educational Science, Luoyang Normal University. His research interests focus on information technology and computer algorithms research.