Abnormal Vehicle Load Identification Method Based on Genetic Algorithm and Wireless Sensor Network

Professor Dr. Sorush Niknamian

BSC in Applied Mathematics (IAUCTB), BSC in Mechatronics Engineering (SUT)

and PhD in Cell and Molecular Biology and Military Medicine, Liberty University

(LU), USA

Email: so.ninamian@gmail.com

Abstract: The current abnormal wireless sensor network vehicle load data recognition method is more complex, which leads to low recognition rate, false alarm rate and slow recognition speed. Based on the genetic algorithm, the accurate method for abnormal wireless sensor network vehicle load data recognition is proposed. The effective feature set of abnormal vehicle load data in the wireless sensor network is constructed, to remove irrelevant features and redundant features from existing abnormal wireless sensor network vehicle load data. The abnormal wireless sensor network vehicle load data in the effective feature set are coded, to reduce the recognition time of abnormal wireless sensor network vehicle load data. The adaptive fitness function, crossover operator and mutation operator are applied to genetic algorithm, which can improve the recognition rate, reduce the false alarm rate, and realize the recognition of abnormal vehicle load data wireless sensor network. The experimental results show that the recognition rate of this method is high, the false alarm rate is low, and the time of recognition is less.

1 Introduction

With the extension of wireless sensor network applications in the field of vehicles, in-vehicle sensor networks, as a new generation of network technology that has attracted much attention, have broad application prospects in urban road condition monitoring and traffic anomaly detection. In view of the fact that traffic anomalies can easily lead to secondary accidents and road congestion problems, how to use vehicle-mounted sensor networks to timely and accurately detect traffic anomaly information, and provide practical information release and subscription services to mitigate the adverse effects of anomalous events on urban traffic. Become the core of promoting smart city construction [1]. Urban road condition monitoring is one of the typical applications of vehicle-mounted sensor networks. It is also the basis of traffic anomaly information detection. During the driving process, the sensor nodes can sense and collect various road condition data in real time, send data to mobile nodes, or aggregate. Go to the roadside unit to assist the traffic control center in monitoring urban road conditions [2, 3]. However, because in-vehicle systems and vehicle users have the willingness to minimize energy consumption and protect personal privacy, the number of nodes participating in routing and forwarding is reduced, and the unfairness of energy consumption makes the problem more severe, leading to existing data collection methods. The work efficiency is low and the transmission reliability cannot be

guaranteed. Therefore, designing an efficient and reliable road data acquisition method is not only conducive to improving and improving the monitoring level of urban road conditions, but also promoting and promoting the in-depth application of in-vehicle sensor network technology in the field of intelligent transportation [4].

At present, there were many defects in the wireless vehicle network abnormal vehicle load data identification method, the recognition rate was low, the false alarm rate was high, and the identification takes a long time. Reference [5] proposed an adaptive wireless sensor network anomaly vehicle load data identification method. This method constructed a heterogeneous complex information wireless sensor network system model, and used an inherent modal decomposition to analyze the abnormal vehicle load data signal. Decomposed into multiple narrow-band signals, design a second-order lattice trap structure, used multiple fixed notch cascades to suppress interference components, use matching projection method to find optimized feature solutions, find all matching feature points, and complete wireless transmission. Identification of network abnormal vehicle load data. Reference [6] proposed a vehicle load data stream based wireless sensor network anomaly vehicle load data identification method, combining the misuse identification mode with the anomaly recognition mode, and establishing the normal mode and the abnormal mode by initial clustering. Knowledge base; using the dissimilarity between the vehicle load data points and the vehicle load data cluster to measure the similarity between the wireless sensor network access vehicle load data and the normal mode and the abnormal mode, thereby determining the wireless sensor network accessing the vehicle load data. Legitimacy;

update the knowledge base by re-clustering to reflect the recent state of wireless sensor network access, and identify abnormal vehicle load data of the wireless sensor network. Reference [7] proposes a method for identifying abnormal vehicle load data in complex wireless sensor networks. This method is based on modularity clustering and graph calculation ideas, applying new node and edge update methods, and using balanced binary tree to increase modularity. Indexing is performed to identify vehicle load data in a wireless sensor network. In order to solve the above problems, this paper proposes an accurate method for identifying abnormal vehicle load data in wireless sensor networks based on genetic algorithm.

2 Wireless sensor network vehicle load data transmission model

In order to realize the abnormal vehicle load detection of the wireless sensor network, the wireless sensor network data transmission model is first constructed. The wireless sensor network carries different protocols and network vehicle loads, and realizes resource scheduling through cloud storage. The wireless sensor network vehicle load model has three types: the rotation protocol, the random access protocol and the vehicle load division protocol. The vehicle load division protocol is mainly FDMA, CDMA, TDMA, etc. [8] It is a non-competitive protocol. The network model results are divided into resource classes, resource instances, attribute classes, attribute instances, etc., in the design of vehicle load data protocol for wireless sensor networks. The SeerSim analysis algorithm is used to construct the network data transmission. The limited queuing model of the NAV task in each vehicle load can use the state space as:

$$S = \left\{ (k,n), 0 \le k \le K, 0 \le n \le N \right\} \tag{1}$$

In the process of network communication, resource scheduling is realized through cloud storage, and abnormal vehicle loads are generated during the continuous transmission and transmission of a large amount of data. At this time, the local cross-term information chain is expressed as:

$$CUB(i) = \frac{CBT(i)}{T} \times CB$$

(2)

Among them, CB represents the number of packets in the system, and i represents the number of packets processing stages. For a general receiving node, when the receiving power is $n = N$, the CPU performs packet processing. When $n = 1,...N-1$, after calculating the state probability of the information chain, it is assumed that there are n neighbor nodes in the communication range of the k node, and the system average throughput of the network. The quantity γ can be expressed as:

$$\gamma = \mu \sum_{k-1}^{K} p_{k,N}$$

(3)

The ratio of all vehicle loads to CUBP, where the path loss factor is:

$$\Delta \tau_{ij}^{k}(t) = Q / z_k$$

(4)

Where z_k is used to describe the length of the route required for the kth data in the current topology. The above analysis process is used to construct the vehicle load data transmission model for the wireless sensor network, as shown in Figure 1, in this model. Data transmission and scheduling, analysis of abnormal vehicle load detection and characterization methods.
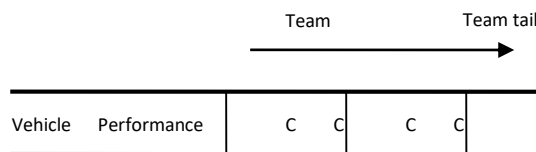

Figure 1 Vehicle load switching queue data structure

In the above-mentioned wireless sensor network vehicle load data transmission model, the design of abnormal vehicle load detection algorithm is carried out [9]. It is necessary to construct the abnormal vehicle load signal model first, and introduce modern signal detection technology to realize signal detection in multiple known interferences. In order to resolve the discrete data of abnormal vehicle load signals, an abnormal vehicle load signal analysis model for wireless sensor networks is constructed:

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)} \qquad (5)$$

In the above formula: $z(t)$ represents the abnormal vehicle load signal, $x(t)$ represents the real part of the abnormal vehicle load signal analysis model, $y(t)$ represents the intrinsic mode function of the abnormal vehicle load signal, and uses the natural mode decomposition to decompose the abnormal vehicle load signal analysis model into Multiple narrowband signals, the envelope characteristics of the obtained data are:

$$a(t) = \sqrt{x^2(t) + y^2(t)}, \theta(t) = \arctan\frac{y(t)}{x(t)} \qquad (6)$$

Among them, $\theta(t)$ represents the high-frequency component, $a(t)$ and $\theta(t)$ are the interference characteristic amplitude and phase information in the wireless sensor network, respectively, and the abnormal vehicle load interference frequency characteristic expression of the heterogeneous network is:

$$f(t) = \frac{1}{2\pi} \times \frac{d\theta(t)}{dt} \qquad (7)$$

The above equation describes the abnormal vehicle load generation check digits of the heterogeneous network in the multi-source process node, which is used to

characterize the data interference frequency. The probability weights of the abnormal vehicle load appear as:

$$w_{ij} = \beta \times w\left(e_p k_q\right) \quad (\beta > 1)$$

(8)

In summary, the signal-to-interference ratio SIR of each vehicle load can be calculated, and the information component of the data information acquisition target is obtained as:

$$r_1 = x(t) - c_1$$

(9)

Through the above processing, a data flow model of abnormal vehicle load is constructed, which provides a signal basis for effective detection of abnormal vehicle loads.

3 Accurate vehicle load accurate identification method

3.1 Construction of effective feature set

Feature selection is to select a valid attribute subset for describing an effective model containing a relatively large, redundant and unrelated attribute vehicle load data set in the abnormal wireless sensor network vehicle load data. Feature selection requires a minimal and most descriptive subset of vehicle load data sets at a tolerable cost within an allowed time, and each feature of the selection of the feature sets is related and non redundant for the feature set [10, 11].

Assuming that the amount of abnormal vehicle load data contained in the feature x may be v_i, and the probability of occurrence of each value is $p(v_i)$, then $H(x) = \sum_{i=1}^{n} p(v_i) \ln p(v_i)$ is called the information entropy of feature x, which means that X contains the number of abnormal vehicle load data.

Setting threshold ε_1, the feature x is randomly selected from the abnormal recognition vehicle load database and its information entropy is calculated. If $H(x) \geq \varepsilon_1$, the feature is selected as the first effective feature of the effective feature library; otherwise, until the first effective feature is selected.

Setting x is as the first effective feature in the effective feature library, and ε_2 is as a predetermined threshold, y is selected from the abnormal recognition vehicle load database randomly, if $H(x \mid y) \geq \varepsilon_2$, then y is as the second features of an effective feature library, otherwise, until second features are selected.

Setting F is a combination of all features, and F_i is one of the features, and C is a given class. Then:

$$S_i = F - \{F_i\} \tag{10}$$

If the feature F_i satisfies $P(C \mid F_i, S_i) \neq P(C \mid S_i)$, the feature F_i is strongly correlated with the class C. If the feature F_i satisfies $P(C \mid F_i, S_i) = P(C \mid S_i)$, and there is a $S_i' \subset S_i$, and $P(C \mid F_i, S_i) \neq P(C \mid S_i')$. Then the feature F_i is weakly correlated with the class C. If feature F_i satisfies any $S_i' \subseteq S_i$ and $P(C \mid F_i, S_i') \neq P(C \mid S_i')$, then the feature F_i is not related to the category C.

After selecting features x and y, a set of x and y sets up a class C. The features in the abnormal recognition vehicle load database are selected one by one, to remove the irrelevant features and redundant features, and form an effective feature set L.

Whether a new feature variable is necessary to enter the model or whether a characteristic variable can be excluded from the model can be judged by the partial F test method based on its significance. It is assumed that there are m independent

variables $x_1, x_2, \cdots, x_m$, and the full model of the m independent variables is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon \qquad (11)$$

The independent variable x_i is eliminated from the model, and the subtraction model is made up of $m-1$ independent variables.

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{i-1} x_{i-1} + \beta_i x_i + \cdots \beta_m x_m + \varepsilon \qquad (12)$$

Supposing that the sum of squares of the complex correlation coefficient of the formula (2) is R^2, the sum of squares of the complex correlation coefficient of the formula (3) is R_i^2, then $\Delta R_i^2 = R^2 - R_i^2$ is defined. If ΔR_i^2 is close to zero, it shows that the independent variable x_i has no significant influence on the dependent variable y. If the ΔR_i^2 is bigger, it shows that the influence of the independent variable x_i on the dependent variable y is greater.

This process is equivalent to the test hypothesis $H_0 : \Delta R_i^2 = 0$, $H_1 : \Delta R_i^2 \neq 0$, which is called the partial F test, and it is the basis of variable selection. In this paper, the feature selection operation of abnormal wireless sensor network vehicle load data is as follows: the characteristics of the effective feature library are checked one by one. If the partial F test of the characteristic variable is significant, the feature is retained; otherwise, the feature is removed from the feature library.

3.2 Accurate recognition of abnormal wireless sensor network vehicle load data

(1) Coding

In this paper, decimal method is used to code the abnormal vehicle load data in the wireless sensor network. In order to accurately recognize the abnormal vehicle load data in the wireless sensor network, all parts of a specific wireless sensor network

connection vehicle load data set are recognized. For a wireless sensor network connection, the protocol type is taken as an example to consider some of the attributes in the Table 1.

Table 1 The characteristic attribute of the wireless sensor network connection

Attribute	Range of value	Example	Description
IP source address	0.0.0.0-255.255.255.255	b2.0f.0.0	B2.0f.f.f-b2.0f.ff.ff
IP destination address	0.0.0.0-255.255.255.255	58.12.d0.00	58.12.d0.00-58.12.d0.ff
Source port	0-65535	35296	Connection source port
Target port	0-65535	00080	Connection target port
Connection time	0-99999999	0000286	The connection time is 286 seconds
Protocol type	1-9	2	The protocol used is TCP
Bytes sent	0-99999999	00002532	2532 bytes of vehicle load data are sent
Response bytes	0-99999999	00008538	Responding to 8538 bytes of vehicle load data

The following rules can be obtained by encoding the example in Table 1:

If the wireless sensor network connection has IP source address, IP destination address, source port number, target port number, connection time, protocol type, sending byte number, response byte number, terminate the connection, the above rules are converted into chromosomal forms, as shown in Figure 2.

(b,2,0,f,0,0,5,8,1,2,d,0,0,0,3,5,2,9,6,0,0,0,0,8,0,
0,0,0,0,2,8,6,2,0,0,0,0,2,5,3,2,0,0,0,0,8,5,3,8)

Figure 2 Diagram of chromosome structure

The newly generated rules match the normal connection. If the match is successful, the rule is misreported, and the chromosome will be punished. If the match is not successful, the rule can get the abnormal accurately, and the chromosome will get a reward. To evaluate the accuracy of the new rule, we need to recognize the vehicle load data in the actual wireless sensor network environment [12-19].

(2) Selection of fitness function

The selection of fitness function is an important aspect of abnormal vehicle load data in the search and optimization process. The core work of genetic algorithm is based on the fitness function. Genetic algorithm is based on the idea of survival of the fittest. Search is based on fitness function value, it represents the quality of individuals, so fitness function is one of the most important parameters in genetic algorithm.

In the process of abnormal vehicle load data recognition, wireless sensor network vehicle load data is classified and encoded into individual form in all areas of the recognized wireless sensor network, and the colony is formed. Then, search rules are applied to group space, and individuals are optimized for search and recognition.

$$Q = \sum_{i=1}^{50} Match \times Wi \tag{13}$$

If $Match$ is equal to 1, the recognized vehicle load data in wireless sensor network matches the rule; if $Match$ is equal to 0, the recognizeed vehicle load data in wireless sensor network does not match the rule, where Wi is the weights of different abnormal wireless sensor network vehicle load data; if the Q value is equal to the number of recognized abnormal vehicle load data, the maximum Q value is as the recognized value. According to the characteristic attributes of wireless sensor network connection in Table 1, the selection order of Wi value is in turn from high to low.

Target IP address

Source IP address

Destination port number

Connection duration

The number of bytes sent by the initiator

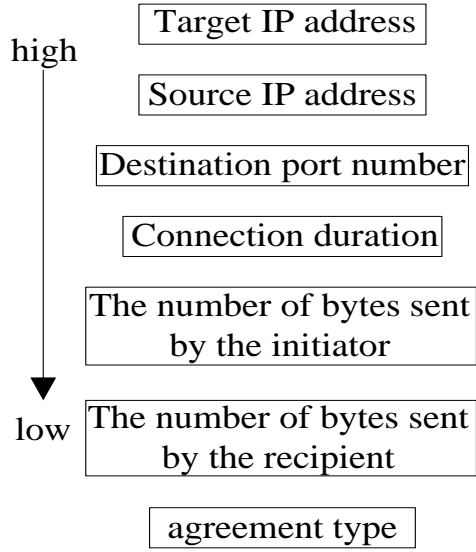low The number of bytes sent by the recipient

agreement type

Figure 3 Weight order

According to the importance of each attribute characteristics, different weights are assigned. In order to get the most important and most direct information of the abnormal vehicle load data, the weight of the target IP address is the first consideration, because the target IP address is the target. Through the IP address of the target, it can identify the abnormal vehicle load data address, and accurately respond to the abnormal vehicle load data in the protected wireless sensor network or host. The next is the source address, because the source IP address is the address that marks the generated exception vehicle load data, and it can protect the normal vehicle load data in the wireless sensor network by blocking the IP connection. Port corresponds to a certain service or wireless sensor network program of host, so the target port has higher weight. It means to open service on the target system, and through the port, it can judge the wireless sensor network program running by the port, and display the running program of the abnormal vehicle load data.

The higher the Q value of the wireless sensor network vehicle load data is, the

more the vehicle load data's suspicious is, and therefore, by judging the size of the Q value, it is analyzed whether the wireless sensor network vehicle load data is abnormal. The specific algorithm is as follows:

$$\Delta = \sum_{i=1}^{50} 1 \times Wi - Q \qquad (14)$$

$$Penalty = \Delta \times Rank \qquad (15)$$

Where, $Rank$ is the error level, and the greater the error level is, the greater the $Rank$ value is. When a very clear normal connection vehicle load data is misclassified as abnormal vehicle load data, the error level of $Rank$ increases, the corresponding punishment received is more; if the judgment of wireless sensor network vehicle load data is not clear on an attribute value, because the vehicle load data attribute is not clear, there is a big error rate, when the error occurs, the corresponding value of $Rank$ is small, and it can be punished less. The adjustment range of fitness can be controlled by the degree of punishment.

$$Fit = Q - Penaly \qquad (16)$$

Through continuous selection, crossover and mutation, the corresponding rules of individual fitness will be selected into the feature library, and the corresponding rules of low fitness individuals will be eliminated.

The fitness function indicates that the individual is close to the optimal solution in the genetic algorithm, and the fitness function is designed to ensure that the fitness is non negative. In the initial stage of evolution, there may be a high degree of individual, but most individual fitness is very low, so the individual selected probability is much higher than other individuals, which may fill the next generation, and the evolution is

difficult to continue and the premature phenomenon occurs; While in the late stage of evolution, the fitness degree of individual in the group is similar, the choice probability difference of the better individual is similar to that of the poor individual, which makes the evolution slow and reduces the convergence speed. This paper uses the fitness difference value to measure the difference size of fitness, which is defined as:

$$\alpha = \Delta z / \Delta Z \tag{17}$$

$$\Delta z = z_{max} - z_{min} \tag{18}$$

$$\Delta Z = Z_{max} - Z_{min} \tag{19}$$

Where, ΔZ is the maximum difference of fitness theory, Δz is the population fitness difference of this generation, z_{max} is the maximum fitness for the population of this generation, z_{min} is the minimum fitness for the population of this generation, Z_{min} is the minimum value of fitness theory, and Z_{max} is the maximum of the fitness theory. It is known that the range of α values is [0,1]. The adaptation is transformed according to α.

$$z'(x) = z(x) + k(\alpha - 0.5) \tag{20}$$

In formula (11), k is constant, $z(x)$ is fitness before transformation, and $z'(x)$ is adaptation after transformation. According to specific circumstances, the greater the adaptation is, the greater the intensity of transformation is. It is tried to ensure that individual fitness is not negative

When the individual fitness difference is large ($\alpha > 0.5$), individual fitness increases $k(\alpha - 0.5)$, however, the increased proportion of individuals with small fitness is larger and the probability of selection increases, which helps to maintain group

diversity and search for wider solution space and avoid falling into local optimum. When the individual fitness difference is small ($\alpha < 0.5$), the fitness of individuals are reduced by $k(0.5-\alpha)$, but the reduced proportion of the individuals with larger fitness is smaller, and the probability selection increases, which helps to increase the search direction, and avoid blind search.

(3) Crossover operator

In this article, the single point crossover in binary crossover is used, to determine the intersection after the target IP address, namely that following thirty-second chromosome, because the target IP address has abnormal wireless sensor network vehicle load data, source IP address is the position to generate an abnormal vehicle load data, both of which are not fixed, A can attack B, and can also attack C, however, M can also attack on B and C. Their attack methods are similar, that is, different attackers can use the same attack means to attack different targets. In this way, there is a greater probability of producing a better fitness than a new chromosome. Assuming that N is the size of the population and Pc is a cross probability, there are $N \times Pc$ crossover operations at each time in the population. After the Wi and Wj are made crossover, a new generation of chromosome X and Y are produced as follows:

Let the vector as:

$$
\begin{aligned}
&\quad\;\; \overset{16\ \text{bit}}{} \qquad \overset{34\ \text{bit}}{} \\
&P_1 = (1,1,\ldots1 | 0,0,\ldots0) \\
&\quad\;\; \overset{16\ \text{bit}}{} \qquad \overset{34\ \text{bit}}{} \\
&P_2 = (|0,0,\ldots0 | 1,1,\ldots1)
\end{aligned}
\tag{21}
$$

then

$$
X = P_1 Wi + P_2 Wj
\tag{22}
$$

$$Y = P_j Wi + P_2 Wi \qquad (23)$$

(4) Mutation operator

Mutation operator refers to the change of a specific gene on a chromosome. Its role is to solve the problem that relying on the implementation of selection and crossover can not generate the optimal solution. In the group generated after crossover operation, there may be evolution to the later stage, and search efficiency decreases. The result may not be global optimal solution or premature convergence. The mutation operation is to select a number of individuals according to a certain strategy and probability, and to change the operation of the mutation, which is to change the value of one of the coded values. The group diversity is increased to avoid premature falling into the local optimal solution.

When the chromosomal differences are small, the mutation probability increases, and the group diversity is maintained, which makes it easier to search for wider solution space and avoid falling into local optimum. When the individual difference is large, the mutation probability is reduced, and the amount of calculation is reduced and the diversity of the group is guaranteed. So the definition of mutation probability as:

$$P_m = 0.2(1-\alpha) \qquad (24)$$

Because the mutation operation can also destroy the good individual, the variation probability is generally smaller, and the range of P_m is [0, 0.2].

(5) The flow of genetic algorithm in wireless sensor network abnormal vehicle load data recognition
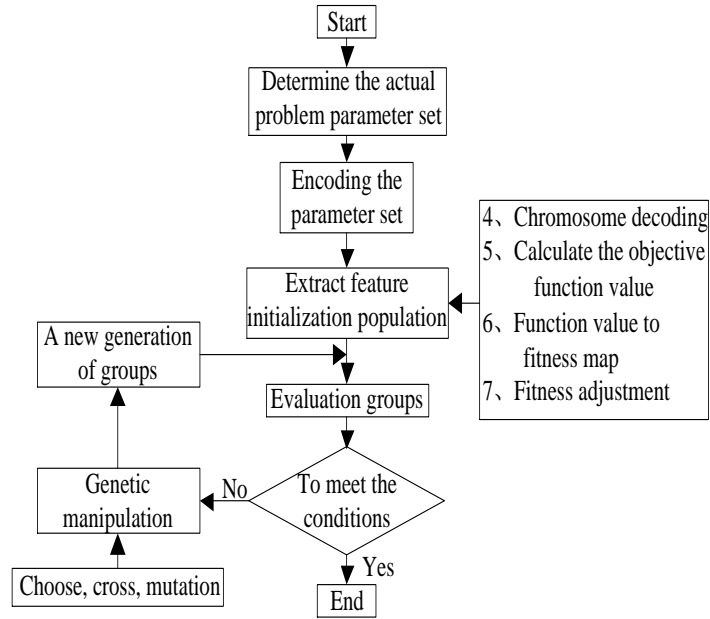
Figure 4 Flow chart of wireless sensor network abnormal vehicle load data recognition based on genetic algorithm

Figure 4 shows that the attribute features are extracted from the current recognized vehicle load data traffic to determine whether it contains an exception feature field. The vehicle load data package with abnormal features is encoded into the form of Figure 1 according to the corresponding encoding rules, and then it is selected, crossed and processed to generate new chromosomes. Then, the fitness of new chromosomes is evaluated, the chromosomes with high fitness are picked out and the above tests are repeated. After that, they can be added to the feature library to become the new rules.

4 Experimental results and analysis

In order to verify the application of the proposed method for abnormal vehicle load data recognition in wireless sensor network, KDD Cup experimental vehicle load data are selected. The vehicle load data are set up by the Lincoln Laboratory of Massachusetts Institute of Technology, and the TCP/IP vehicle load data are extracted

from the typical American air force wireless sensor network. 41 characteristics are extracted from the vehicle load data. The vehicle load dataset includes about 4.9 million training sets and 2.9 million testing sets. These vehicle load data are divided into 4 categories: (1) remote unauthorized access, (2) denial of service, (3) vulnerability recognition of system, and (4) super users' unauthorized access.

After reprocessing the original KDD competition vehicle load data, a total of 62 cases of flow vehicle load data of super user unauthorized access, 14951 cases of traffic vehicle load data of system's vulnerability recognition, 1002 cases of traffic vehicle load data of distal unauthorized access, 247371 cases of traffic vehicle load data of denial service, and 812815 cases of normal traffic vehicle load data are received

The experimental software platform is Windows XP and Matlab6.5, and the hardware platform is IBM Z61t, 1 G memory. During the experiment, there are 14 chromosomes in the initial population. Each chromosome is encoded with length of 50, the population size is 1700, the number of generations is 50000, the crossing rate is 0.5%, the mutation rate is 0.05, the maximum migration number is 4, and the migration interval is 5. According to the test by the following formula, Table 2 is the experimental vehicle load data of the proposed method, and Table 3 is the experimental vehicle load data of the current method.

$$\text{Detection rate} = \frac{\text{The number of abnormal data detected correctly}}{\text{The number of all abnormal data}} \times 100\%$$

$$(25)$$

$$\text{Error rate} = \frac{\text{The number of misinformation events}}{\text{The number of all events}} \times 100\%$$

$$(26)$$

Table 2 Experiment vehicle load data of the proposed method

Protocol type	Test number	Number of abnormal vehicle load data	Number of recognition	Recognition rate (%)	Number of false recognition	False recognition rate (%)
TCP	4383	1936	1857	95.92	64	3.44
IP	3759	1494	1407	94.18	43	3.06
ICMP	3466	1259	1174	93.25	32	2.73
UDP	2919	1013	945	93.29	23	2.43

Table 3 Experiment vehicle load data of the current method

Protocol type	Test number	Number of abnormal vehicle load data	Number of recognition	Recognition rate (%)	Number of false recognition	False recognition rate (%)
TCP	4383	1936	1645	84.97	89	5.41
IP	3759	1494	1258	84.20	63	5.00
ICMP	3466	1259	963	76.49	57	5.92
UDP	2919	1013	785	77.49	42	5.35

Figures 4 and 5 can be obtained from vehicle load data in Table 2 and table 3.
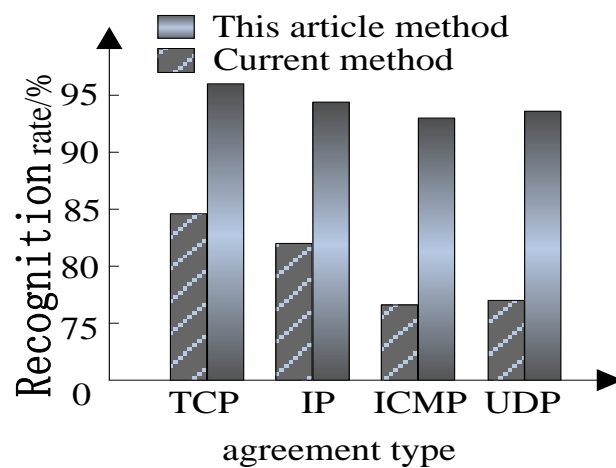
Figure 5 The recognition rate of two different methods
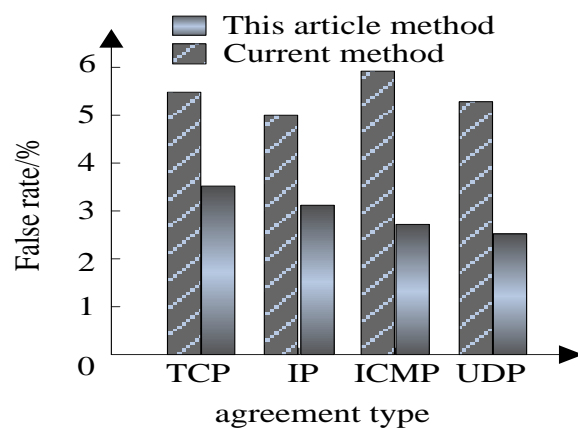
Figure 6 False recognition rate of two different methods

The recognition rate and error rate of this method and the current method are shown in Table 2 and Table 3. The recognition rate is the ratio of the number of abnormal vehicle load data in the wireless sensor network to the number of all abnormal vehicle load data. The false recognition rate refers to the ratio of the number of the normal vehicle load data in the wireless sensor network but recognized as abnormal vehicle load data to the number of all normal vehicle load data. Analysis of Table 2, Table 3, Figure 5 and Figure 6, we can see that the minimum recognition rate of the current method is 76.49% and the maximum false recognition rate is 5.92%; while the minimum recognition rate of the proposed method is 93.25% and the maximum false recognition rate is 3.44%. It shows that the accurate recognition method of abnormal wireless sensor network vehicle load data based on genetic algorithm has good effectiveness and feasibility.

The optimal solution is approximated by the present method and the current genetic algorithm respectively. The maximum fitness values of the two methods are analyzed, and the results are shown in Figure 7.
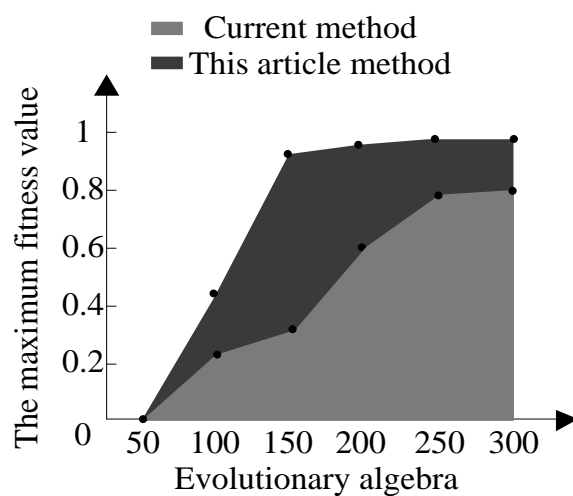

Figure 7 Maximum fitness of two different methods

From Figure 7, we can see that with the increase of evolutionary generations, the maximum fitness values of the proposed method and the current genetic algorithm increase gradually, and then tend to be stable. However, the proposed method tends to be stable with fewer evolutionary generations, which indicates that the proposed method tends to be more stable.

In order to verify the recognition efficiency of the proposed method, the proposed method and the current method are used to test it. When the vehicle load data volume is the same, the time spent by two different methods for recognizing wireless sensor network abnormal vehicle load data is compared. The test results are shown in Figure 8.
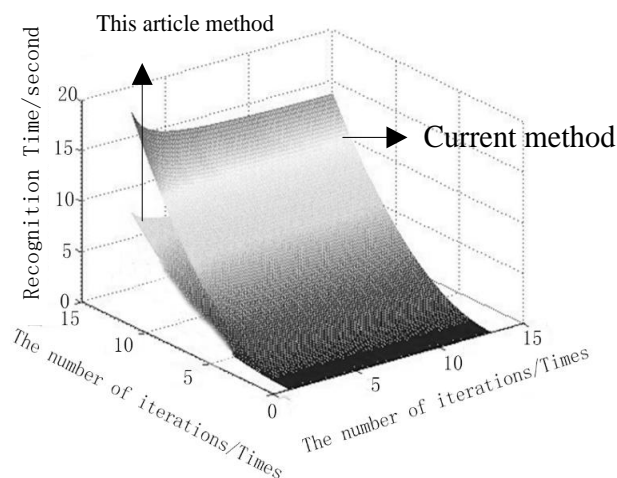

Figure 8 Test results of two different methods

Analysis of Figure 8Figure shows that when the conditions are all the same, it is much less time to recognize abnormal vehicle load data in the wireless sensor network based on genetic algorithm than that of the current method. Because before recognizing, a wireless sensor network outliers recognition algorithm based on genetic algorithm is used to construct the effective feature set, which removes the redundant and

uncorrelated vehicle load data in the wireless sensor network abnormal vehicle load data, and encodes the wireless sensor network outliers, reducing the recognition time.

5 Conclusions

When the current method is used to recognize abnormal vehicle load data in the wireless sensor network, the recognition algorithm is too complicated, which leads to the problem of high error recognition rate and long recognition time. On the basis of genetic algorithm, this paper proposes an accurate recognition method for wireless sensor network abnormal vehicle load data, and applies genetic algorithm to wireless sensor network abnormal vehicle load data recognition. The experimental results show that this method has high recognition rate and less time to recognize abnormal vehicle load data in the wireless sensor network, which provides a guarantee for the safe operation of the wireless sensor network.

Reference

[1] Xu Y, Xu N, Zhuang Z, et al. (2017). An Algorithm of Abnormal Data Detection for Internet of Vehicles Based on Crowdsensing. Hunan Daxue Xuebao/journal of Hunan University Natural Sciences, 44(8): 145-151.

[2] Kailkhura, B., Han, Y.S., Brahma, S., et al. (2015). Distributed Bayesian Detection in the Presence of Byzantine Data. IEEE Transactions on Signal Processing, 63(19), 5250-5263. DOI: 10.1109/TSP.2015.2450191.

[3] Gao, Y.-H., Zhang, L.-T., Liang, J., et al. (2016). Abnormal noise analysis for commercial vehicle cab based on psychoacoustics. Journal of Jilin University, 46(01): 43-49.

[4] Kumarage, H., Khalil, I., and Tari, Z. (2015). Granular Evaluation of Anomalies in Wireless Sensor Networks Using Dynamic Data Partitioning with an Entropy Criteria. IEEE Transactions on Computers, 64(9), 2573-2585. DOI: 10.1109/TC.2014.2366755.

[5] Mehrjoo S, Khunjush F. (2018). Accurate compressive data gathering in wireless sensor networks using weighted spatio-temporal compressive sensing. Telecommunication Systems, 68(12): 1-10.

[6] Mu, L.W., Peng, X.B., and Huang, L. (2015). Abnormal Data Detection Algorithm in Heterogeneous Complex Information Network. Computer Science, 42(11), 34-137. DOI: 10.11896/j.issn.1002-137X.2015.11.028.

[7] Xin H, Liu X. (2017). Energy-Balanced Transmission with Accurate Distances for Strip-based Wireless Sensor Networks. IEEE Access, PP. (99): 1-1.

[8] Rathore, M.M., Ahmad, A., and Paul, A. (2016). Real time intrusion detection system for ultra-high-speed big data environments. Journal of Supercomputing, 72(9), 3489-3510. DOI: https://doi.org/10.1007/s11227-015-1615-5.

[9] Bui T O, Xu P, Phan N Q, et al. (2016). An Accurate and Energy-Efficient Localization Algorithm for Wireless Sensor Networks// IEEE, Vehicular Technology Conference. IEEE, 1-5.

[10] Zhang, J., Li, H., Gao, Q., et al. (2015). Detecting anomalies from big network traffic data using an adaptive detection approach. Information Sciences, 318(C), 91-110. DOI: org/10.1016/j.ins.2014.07.044.

[11] Zhao, Q.Q. (2017). Under the Environment of Dynamic Data Network Duplicate

Data Simulation Test Method. Computer Simulation, 34(6), 445-448. DOI: 10.3969/j.issn.1006-9348.2017.06.097.

[12] Liu D, Cao Z, Zhang Y, et al. (2017). Achieving Accurate and Real-Time Link Estimation for Low Power Wireless Sensor Networks. IEEE/ACM Transactions on Networking, PP. (99): 1-14.

[13] Dayal, A. (2018). Improving adaptive frameless rendering. International Journal of Computers and Applications, 40(2): 110-120.

[14] Qiu, K. and Wang, J.R. (2018). A Fractional Integral Identity and its Application to Fractional Hermite-Hadamard Type Inequalities. Journal of Interdisciplinary Mathematics, 21(1): 1-16.

[15] Chen, D.Z. (2018). Performance Evaluation Model for Government-Supportive Urbanization. Journal of Discrete Mathematical Sciences and Cryptography, 21(4): 859-868.

[16] Rintala, J. and Kolari, M. (2017). Better Hygiene in Food Packaging Board at Reduced Risk of Rejected Tonnage and Machine Corrosion. Paper Asia, 33(5): 20-23.

[17] Khorramnejad, K.; Ferdouse, L.; Guan, L., and Anpalagan, A. (2018). Performance of integrated workload scheduling and pre-fetching in multimedia mobile cloud computing. Journal of Cloud Computing, 7(1)

[18] R, A. K., Dr. K, B. (2017). Pi and sliding mode speed control of permanent magnet synchronous motor fed from three phase four switch vsi. Journal of Mechanical Engineering Research and Developments, 40(4): 716-725.

[19] Nie, Z.Q. (2018). Discounted cash flow (dcf) model detection based on goodwill impairment test. Journal of Discrete Mathematical Sciences and Cryptography, 21(4): 959-968.