An Assessment Method of Internet of Vehicles User Behavior Based on Hidden Markov Model

Peng-Shou Xie, Yi-Fan Wang, Zong-Liang Wang, Nan-Nan Li, Tao Feng, and Yan Yan
(Corresponding author: Yi-Fan Wang)

School of Computer and Communications, Lanzhou University of Technology
No. 36 Peng-jia-ping Road, Lanzhou, Gansu 730050, China
Email: 844782234@qq.com

(Received July 21, 2021; Revised and Accepted Apr. 13, 2022; First Online Apr. 23, 2022)

Abstract

In order to solve the problem that some attackers steal, tamper or destroy the stored sensitive data after passing the identity authentication on the Internet of Vehicles service cloud platform, an assessment method of Internet of Vehicles user behavior based on the Hidden Markov Model is proposed in this paper. Firstly, the critical behavior features of Internet of Vehicles users with higher weight values are extracted using Term Frequency-Inverse Document Frequency. Then, according to the behavior features, the Internet of Vehicles user behavior model is established using the Hidden Markov Model. Finally, the weight difference of the user's behavior sequence before and after is calculated by improving the Forward algorithm of the Hidden Markov Model to assess whether the user's current behavior is legal or not. The experimental results show that the method improves the comprehensive assessment index F1-Score on Internet of Vehicles user behavior assessment results.

Keywords: Behavior Assessment; Data Security; Hidden Markov Model; Internet of Vehicles Service Cloud Platform; Term Frequency-Inverse Document Frequency

1 Introduction

As an important communication server, the Internet of Vehicles service cloud platform has been widely used in the scene of the Internet of Vehicles. According to the risk analysis, it constructs a multi-dimensional integrated protection system for site security, host security, data security and business security, and so on [20]. In recent years, the state has promulgated the strategy of traffic power Internet of Vehicles (IoV) is taken as the key application scenario of 5G communication infrastructure construction. At the same time, many enterprises and individual car owners have begun to store data related to vehicle nodes in the service cloud platform. When IoV users use the service cloud platform in the process of stealing by others or unintentionally illegal operations leading to the leakage of the user's information is a very common security problem [5]. At the same time, it also increases the risk of data storage in the IoV service cloud platform [6]. Attackers forge and impersonate the identity of legitimate IoV users, illegally access the service cloud platform, tamper, steal or damage the stored sensitive data which poses a great threat to the property and information security of the IoV users.

User behavior analysis provides a new solution for the research of an access control model in an open network environment. People's daily behavior pattern is regular and it is the same when visiting the IoV service cloud platform [3]. Therefore, through the statistical analysis of the daily behavior data of users, the habitual behavior patterns of users can be obtained, to distinguish abnormal behaviors and judge whether the user is legitimate. Researchers at home and abroad use a variety of techniques to assess user behavior more accurately [10].For example, building analytical models [16], data mining [11], machine learning [19], artificial intelligence [15] and log auditing [7]. Although these methods improve the accuracy of behavior assessment, they often ignore the relevance and individual differences between the operation behaviors, failing to fully describe the user's operation behavior, which cannot meet the accuracy requirements of the access control of the IoV service cloud platform.

In order to better solve the above problems, this paper studies the behavior pattern of the IoV users in transaction processing on the service cloud platform and finds the differences between the behavior sequences in the behavior pattern. An assessment method of the IoV user behavior based on the Hidden Markov Model (HMM) is proposed. After the key features of command samples are effectively extracted by using Term Frequency-Inverse Document Frequency (TF-IDF), the user behavior model of IoV is established based on HMM, and an improved Forward algorithm to calculate the weight difference of the before and after behavior sequence. Finally, the effectiveness of the proposed assessment method is verified based on simulation experiments, which adds a security

guarantee of behavior authentication for access control of the IoV service cloud platform.

2 An Assessment Method of IoV User Behavior Based on HMM

By analyzing the behavior of IoV users, we can gain insight into the information hidden behind each transaction processing behavior and discover the corresponding network security problems, thus enhancing the network situational awareness and enhancing the defense ability of network security [17]. This method not only optimizes the access control system but also gives a timely warning of identity theft attacks, preventing attackers from modifying permissions, tampering with information stealing data in the network by stealing the legitimate certificates of IoV users, thus improving the security of the IoV service cloud platform.

When the user is a newly registered user or an old user has some new behaviors, the historical behavior data do not exist in the service cloud platform, so it cannot be judged only based on the historical behavior of a single user. In order to comprehensively assess the user behavior of the IoV, this paper studies the various behaviors of IoV user groups on the transaction processing of the service cloud platform, so as to assess the user's abnormal behavior, and dynamically adjust the access strategy [1].

1) First of all, collect and sort out the behavior information of the IoV user group in the transaction processing of the service cloud platform. The IoV user group is divided into known users and unknown users to learn the normal and abnormal behavior data of several known users, so as to assess the behavior of unknown users more accurately;

2) Secondly, TF-IDF [24] is used to extract the behavior features of known users, remove the redundant information in the original data set, and finally get the key behavior features with higher weight value;

3) Then, the IoV user behavior model is established based on HMM. The elements of HMM are redefined to make it more suitable for the IoV service cloud platform. According to the Baum-Welch algorithm [21], the model is trained and optimized;

4) Finally, the Forward algorithm of HMM is improved to calculate the difference of the weight value ΔW of the behavior sequence before and after the IoV users, so as to more accurately assess whether the IoV users' behavior is legal at a certain time. In case of abnormal behavior, adjust the access right immediately and conduct secondary identity authentication; If it is normal behavior, accept and continue to assess the follow-up behavior until it exits the IoV service cloud platform.

The flow chart of the IoV user behavior assessment method based on HMM is shown in Figure 1.

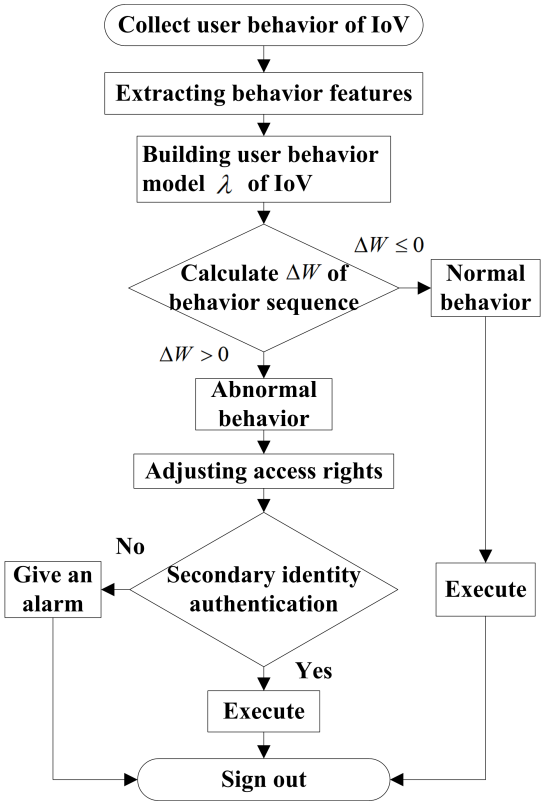


Figure 1: Flow chart of the IoV user behavior assessment

2.1 IoV User Behavior Feature Extraction

The operational command like send mail is a very common behavior in the user operation behavior, so the frequency of this command is higher. But after data desensitization, the command will not contain any important information and become an operation command with no difference. These operational commands would not be significant if they were the primary features of the IoV user behavior assessment. So this paper should look for those behaviors which are more frequent in a single user's operation but less frequent in all users.

In order to better solve the above problems. The TF-IDF method is used to extract features from the data set, select more meaningful behavior features and accurately assess the IoV user behavior. TF-IDF is a method of text data vectorization that is used in text analysis, Chinese and English keywords extraction, information retrieval, document classification, and so on.

Let $t_i(i = 0, 1, \cdots, n-1)$ be an operation command of an IoV user, and $d_j(j = 0, 1, \cdots, n-1)$ be a behavior data set of an IoV user including t_i.

Step 1. Calculate TF. That is, the probability of an operational command of the IoV user appearing in his behavior data set. Where $n_{i,j}$ is the number of times t_i appears in d_j, and $\sum_k n_{k,j}$ is the total number of operational commands contained in behavior data set

d_j. The calculation is shown in Equation (1):

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

Step 2. Calculate IDF. That is the probability that an operational command of the IoV user appears in all IoV user behavior data sets. Where $|D|$ is the total number of IoV users. $|\{j : t_i \in d_j\}|$ is the total number of IoV users including.If the operation command is not in any behavior data sets of IoV users, the denominator will be zero, so the denominator uses $|\{j : t_i \in d_j\}|+1$. The calculation is shown in Equation (2):

$$idf_i = lg\frac{|D|}{|\{j : t_i \in d_j\}|+1} \quad (2)$$

Step 3. Calculate TF-IDF. That is the weight value W_{a1} of an operational command of the IoV users. The calculation is shown in Equation (3):

$$W_{a1} = tf_{i,j} \times idf_i \quad (3)$$

If the weight value calculated by the TF-IDF is higher, it indicates that is more important in all IoV user behaviors. Taking this behavior as a key behavior feature is very helpful for a behavior assessment.

2.2 User Behavior Model of IoV is Established Based on HMM

In this paper, the observation variables, state variables, state transition probability matrix, observation variable probability matrix, and initial state probability vector of HMM are redefined. The user behavior model of IoV is established by using HMM, which is represented by $\lambda = (X, Y, A, B, \pi)$, the following will describe five parameters.

1) Observation variable X of the IoV user behavior model
The observation variables in this paper represent the behavior sequence when users process transactions on the IoV service cloud platform. Behavior sequence can be directly observed and contains key behavior features. Set X as $\{x_1, x_2, \cdots, x_m\}$,Where $x_i \in X$ represents the operation behavior sequence of the IoV users at i time. The number of different behavior features in each state is denoted by m. The observation set is O_i, that is $O_i \in X = \{x_1, x_2, \cdots, x_m\}$.

2) State variable Y of the IoV user behavior model
Set the state variable Y of the IoV user behavior model to $\{y_1, y_2, \cdots, y_n\}$, the state at i time is q_i, that is $q_i \in Y = \{y_1, y_2, \cdots, y_n\}$. The number of states in the model is denoted by n. Generally, Y cannot be observed. According to the behavior features of the IoV users, the state of sensitive data stored in the IoV service cloud platform is taken as the state variable of the model in this paper. According to real-life, the state of sensitive data is divided into two kinds: dangerous and safe.

3) State transition probability matrix A of the IoV user behavior model
A is used to represent the state transition probability matrix of the IoV user behavior model, denoted as $A = (a_{ij})_{n\times n}$. A is the probability distribution of mutual transfer between sensitive data states stored in the service cloud platform. $a_{ij} = P(q_{t+1} = y_j|q_t = y_i), 1 \le i \le n, 1 \le j \le n$, a_{ij} means the state of sensitive data stored by the IoV users at t time is q_t. The probability of its transition to the state of $t+1$ time is $q_t + 1$.

4) Observation variable probability matrix B of the IoV user behavior model
B is used to represent the probability matrix of the observed variables of the IoV user behavior model, denoted as $B = (b_{jk})_{n\times m}$. Each state variable corresponds to an observation variable and its relative probability distribution is calculated as $b_{jk} = P(O_t = x_k|q_t = y_j), 1 \le j \le n, 1 \le k \le m$. b_{jk} is the probability of generating the sequence of operation behavior x_k at t time sensitive data is in the state q_t.

5) nitial state probability vectorof the IoV user behavior model
π is used to represent the initial state probability vector of the IoV user behavior model, denoted as $\pi = (\pi_i)_{1*n} = \{\pi_1, \pi_2, \cdots, \pi_n\}$. π is the initial probability distribution matrix of order $1 \times N$. $\pi_i = P(q_1 = y_i), 1 \le i \le n$ is the probability that the stored sensitive data is in the state q_i when $t = 1$.

In order to establish the model more quickly and conveniently, based on the initial values of n, m, A, B and, combined with the Baum-Welch algorithm of HMM, the parameters of the IoV user behavior model is determined and the model is optimized.

Baum-Welch algorithm [22] starts from the initial estimation of parameters, and achieves the local optimal value through parameter learning training, and considers that the probability distribution of the initial state is uniform. Therefore, this method assumes that the user behavior model of the Internet of vehicles has T states and K observation symbols. Then the probability of taking any states as the initial state is $1/T$. The transition probability of each step between each state in the model is also $1/T$. In each of these states, the probability of behavior feature is $1/K$.

2.3 The Weight Calculation of User Behavior Sequence of IoV

There are three basic algorithms for HMM: Forward algorithm for model assessment, Viterbi algorithm for decoding, Baum-Welch algorithm for parameter learning. The

Forward algorithm can effectively calculate the probability $P(X|\lambda)$ of an observed variable X in the model λ.

In order to better reflect the relevance of user transaction processing behavior. In this paper, the weight of the behavior sequence of the IoV users is calculated by improving the Forward algorithm, and the difference between the weight values of the front and rear behavior sequences is calculated in order to assess whether the current behavior of the IoV users is legal. The weight of the behavior sequence refers to the number of times that the sequence of behaviors appears in all operating commands. If a large number of IoV users often perform the same transaction operation, the weight of the behavior sequence will be increased.

At t, λ outputs the observation variable $X_1 = x_1, x_2, \cdots, x_R$ with the length of R. This variable is the operational command of the IoV users from time 0 to time t, forming an initial observation variable. The weight valueof W_1 is X_1 calculated as shown in Equation (4):

$$W_1 = P(X_1|\lambda) = P(x_1, x_2, \cdots, x_R|\lambda) \tag{4}$$

At $t+1$, λ outputs the behavior x_{R+1}, discards x_1 in X_1 and obtains a new observation variable $X_2 = x_2, x_3, \cdots, x_R, x_{R+1}$ with R length. The weight value W_2 of X_2 is calculated as shown in Equation (5):

$$W_2 = P(X_2|\lambda) = P(x_2, x_3, \cdots, x_R, x_{R+1}|\lambda) \tag{5}$$

Calculate the difference ΔW between the weight values of the two observation variables, as shown in Equation (6):

$$\Delta W = W_1 - W_2 \tag{6}$$

If $\Delta W > 0$, indicates that the weight value of X_2 is smaller than that of X_1. Therefore, the probability of newly observed variability being accepted by the model is low, X_{R+1} may be an abnormal behavior. Discard X_{R+1}. Immediately adjust the user's access rights and conduct secondary identity authentication. If it is the user himself, accept and continue to assess the follow-up behavior; If it is an attacker, it will immediately report to the police and exit the IoV service cloud platform; But if $\Delta W \leq 0$,it shows that for the trained model, the weight of new observation variables is increasing. X_{R+1} is normal behavior. X_{R+1} is added to the observation variable as a new sequence which is used as the basic sequence to determine the validity of the next behavior.

As time goes on, the behavior patterns of the IoV users may be changed, adding new behavior sequences to the observed variability all the time. Its significance is to continuously learn the behavior patterns of the IoV users, reduce the false alarm rate and prevent missing some new malicious behaviors, so as to better adapt to the changes of the IoV users' behavior.

3 Simulation Experiment and Result Analysis

3.1 Experimental Environment and Assessment Index

The system environment of this paper is a win10 64-bit operating system. Python 3.6 is selected as the programming environment. Pycharm community 2019.2.1 is selected as the programming platform. The Confusion matrix is selected and a comprehensive assessment is made from four indicators: Accuracy, Precision, Recall and F1-Score. Because the transaction data of the IoV users on the service cloud platform cannot be obtained directly. SEA data set [14] was selected as the simulation experimental data in this paper.

The SEA data set is composed of operation command files of 50 users in the UNIX system. Each user's data set contains 15000 operation commands,such as $\{cpp, sh, cpp, mvdir \cdots\}$. The first 5000 commands in the dataset are the normal operation commands of users, in the remaining commands, the operation commands of 20 masquerades who are not among the 50 users are inserted with a certain probability as the abnormal behavior data of the masquerades. And every 100 commands as a command block, 0 means that there is no abnormal operation behavior in the command block, 1 means that there is an abnormal operation behavior in the command block.

In this paper, a confusion matrix is used to assess the assessment results. The assessment results of the IoV user behavior can be divided into four situations: TP, FN, FP, TN [18]. TP refers to the number of normal behaviors of IoV users assessed as normal behaviors; FN refers to the number of normal behaviors of IoV users assessed as abnormal behaviors; FP refers to the number of abnormal behaviors of IoV users assessed as normal behaviors; TN refers to the number of abnormal behaviors of IoV users assessed as abnormal behaviors. The confusion matrix is shown in Table 1:

Table 1: Confusion matrix

Actual	**Assessment**	
	Normal	Abnormal
Normal	TP	FN
Abnormal	FP	TN

According to the elements in the confusion matrix [9], four standard assessment indexes can be obtained, as shown in Equation (7), Equation (8), Equation (9) and Equation (10).

1) Precision

$$\Pr e = \frac{TP}{TP + FP} \tag{7}$$

2) Recall

$$\mathrm{Re} = \frac{TP}{TP + FN} \tag{8}$$

3) Accuracy

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \tag{9}$$

4) F1-Score

$$F1 = \frac{2 \times \mathrm{Pr}\, e \times \mathrm{Re}}{\mathrm{Pr}\, e + \mathrm{Re}} \tag{10}$$

Accuracy is the only assessment index used by some researchers, but Accuracy cannot fully represent the assessment method. For example, the behavior in the user behavior data set is normal and some kinds of assessment methods evaluate all behaviors as normal behaviors. In fact, this is wrong, the Accuracy of the method will be close to 100%. However, this method will bring huge losses to users when they encounter identity theft attacks. Therefore, four assessment indicators were selected by this paper to assess the method more comprehensively and scientifically [2].

3.2 Experimental Simulation and Result Analysis

Because the operation behavior habits and environment of each user are different, the behavior between users and users is different. Therefore, it will take a lot of time to model the operation behavior of each user separately. Moreover, the number of abnormal behaviors in the SEA data set is less [8]. Although this is more in line with the actual situation, if the training set and test set are randomly assigned, it is very easy to appear the training set without abnormal behaviors, which makes the discrimination of data not enough and affects the assessment results.

In order to better simulate the real situation in the simulation experiment in this paper, the normal behavior data and abnormal behavior data from user 1 to user 9 in the data set are selected as the known user behavior data of IOV, namely the training set; The behavior data of User 10 is regarded as the unknown user behavior data of IoV, namely the test set. The sampling distribution of the data set is shown in Table 2:

Table 2: Sample distribution of data set

Num	Type	Number of Samples	
		Training	Testing
0	Normal	131300	13700
1	Abnormal	3700	1300

TF-IDF is used to calculate the weight value W_{a1} of the nine users' behaviors, and the key behavior features of the IoV users' behavior assessment are selected according to the calculation results. The feature?s weight value W_{a1} of IoV user behavior is shown in Table 3:

Table 3: Features weight value W_{a1}

Num	Features	W_{a1}
0	rm	0.067231
1	whodo	0.040209
2	su	0.025281
3	write	0.022363
4	cp	0.021354
5	ls	0.016226
6	cat	0.015828
7	gcc	0.012199
8	egrep	0.012052

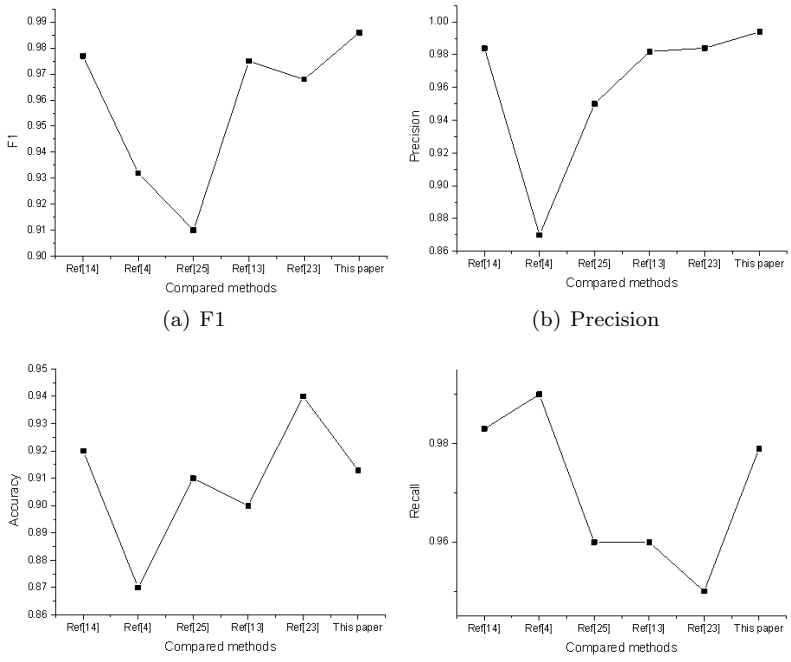
As can be seen from this Table: among all the operating commands of the nine IoV users, the most discriminative behavior features are rm and whodo. Because of the sixth behavior, the weight value W_{a1} is less than 0.02. If these behaviors are also considered as key behavioral characteristics of IoV user behavior assessment, the assessment results will be affected and time-consuming. Therefore, delete these redundant features or less important features. Only the first five operational behaviors were selected as key behavioral features. The behavior features of the IoV users are shown in Table 4:

Table 4: Behavior features of the IoV users

Num	Features	Description
0	rm	Delete sensitive data files
1	whodo	Query a user's access history
2	su	Change user account password
3	write	Talk to another user
4	cp	Copy sensitive file data

TF-IDF extracts five kinds of key behavior features of nine IoV users, which are rm, whodo, su, write, cp. That is, the number of corresponding behavior features in each state is $m = 5$, and the model parameter is; There are two states of sensitive data: dangerous and safe. That is, the number of model states is $n = 2$, and the model parameter. Using the Baum-Welch algorithm to train the model and determine the remaining parameters [12], a complete IoV user behavior model based on HMM is obtained to assess the behavior of unknown users of IoV. The confusion matrix of this paper is shown in Table 5:

In order to verify the performance of this method in the assessment of IoV user behavior, according to the confusion matrix, four indicators, Precision, Recall, Accuracy and F1-Score are compared with the existing research. In this paper, the methods proposed in reference [14], reference [4], reference [25], reference [13] and reference [23]



(a) F1 (b) Precision

(c) Accuracy (d) Recall

Figure 2: Comparison of experimental results

Table 5: Confusion matrix of this paper

Actual	**Assessment**	
	Normal	Abnormal
Normal	13413	289
Abnormal	86	1214

are selected as comparative experiments.

In reference [14], the bag of words model is used to process the data in the SEA dataset, and the LSTM algorithm is combined with a bimodal threshold mechanism to assess user security behavior. In reference [4], the bag of words model, N-Gram model and XG-Boost algorithm are combined to assess user behavior. N-Gram model can assess the difference between two strings and effectively compare the correlation of operational behavior before and after. In reference [25], the Random Forest algorithm is used to extract behavior features based on the user input frequency and effectively detect user's malicious operations. In reference [13], BiLSTM and Attention mechanisms are used to process serialized data and detect abnormal behavior. In reference [23], TF-IDF is used for feature extraction and LSTM training behavior model, so as to assess abnormal behaviors. The comparison of experimental results is shown in Figure 2.

According to this figure, Figure (a) is the experimental comparison diagram of F1-Score. Figure (b) is the experimental comparison diagram of Precision. Figure (c) is the experimental comparison diagram of Accuracy. Figure (d) is the experimental comparison diagram of Recall. Although the Recall index of the proposed method is lower than that of the reference [14] and reference [4]. Accuracy is lower than reference [14] and reference [23]. However, the comprehensive assessment indexes F1-Score and Precision are higher than other assessment methods. F1-Score is the harmonic average of Precision and Recall, and it is also the weighted average of Accuracy and Recall. When F1-Score is higher, it means that the assessment method is more effective in comprehensive behavior assessment. The experimental results show that the method based on HMM has obvious advantages over other methods in the field of IoV user behavior assessment.

In order to better verify the generalization ability of the proposed method, four users from User 11 to User 50 are randomly selected as test objects to detect whether the abnormal behaviors of these four users can be accurately assessed. The experimental results are shown in Figure 3.

From the analysis of the experimental results in figure, it can be seen that when the unknown users of IoV have many abnormal behaviors, such as (User24, User43). The method in this paper can accurately assess abnormal behavior; However, when there are few abnormal behaviors, such as (User25, User30), the model does not learn enough about abnormal behaviors, which leads to some errors in the assessment results. In addition, the improved weight calculation method for the user behavior sequence of IoV fully considers the correlation of user behavior, but it leads to the decrease of the weight value of some subsequent behaviors, which is regarded as abnormal behavior so that the assessed value of the experimental results is greater than the real value. This will be the need for

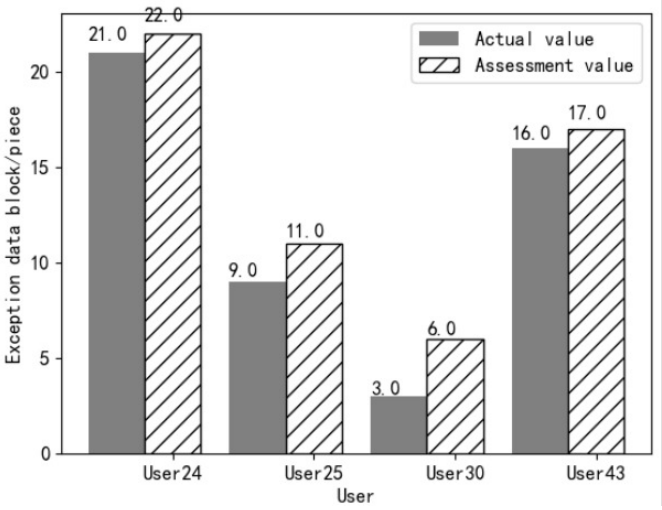


Figure 3: Assessment of abnormal behavior

improvement in future research.

4 Conclusion

This paper describes how to assess IoV user behavior based on HMM and TF-IDF, which adds a security guarantee of behavior authentication for the access control model of the IoV service cloud platform. Experimental results show that the method has the ability to assess the abnormal behavior of unknown users of IoV and it can also effectively protect the integrity, confidentiality and availability of sensitive data, and promote the popularization and promotion of the IoV service cloud platform.

In this paper, the assessment method of IoV user behavior based on HMM does not consider any special identity of IoV users, such as the administrator of the IoV service cloud platform and when there are abnormal behaviors in the behavior sequence, some subsequent behaviors will be regarded as abnormal behaviors. In future research, the method still needs to be improved to classify different identities of IoV users, so as to make the assessment method more comprehensive.

Acknowledgments

This research is supported by the National Natural Science Foundations of China under Grants No.61862040, No.61762059 and No.6176 2060. The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] S. Ashry, W. Gomaa, and M. Abdu-Aguye, "Improved IMU based Human Activity Recognition Using Hierarchical HMM Dissimilarity," in *17th International Conference on Informatics in Control*, pp. 702–709. France, 2020.

[2] C. M. Chen, G. H. Syu, and Z.X. Cai, "Analyzing System Log based on Machine Learning Model," *International Journal of Network Security*, vol. 22, no. 6, pp. 925–933, 2020.

[3] G. R. Chen, K. Wang, and J. Tan, "A Risk Assessment Method based on Software Behavior," in *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 47–52. Shenzhen, 2019.

[4] M. S. Chen and K. H. Wu, "Internal Attack Detection based on Shell Commands," *Computerand Modernization (in chinese)*, no. 1, pp. 56–60, 2021.

[5] Y. Chule, A. Renzaglia, and A. Paigwa, "Driving Behavior Assessment and Anomaly Detection for Intelligent Vehicles," in *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS)*, pp. 524–529. Bangkok, 2019.

[6] M. Fouad and A. H. Amr Talaat, "On Detecting IOT Power Signature Anomalies Using Hidden Markov Model HMM)," in *2019 31st International Conference on Microelectronics (ICM)*, pp. 108–112. Cairo, 2019.

[7] T. L. Gao, T. Li, and R. Jiang, "Research on Cloud Service Security Measurement based on Information Entropy," *International Journal of Network Security*, vol. 21, no. 6, pp. 1003–1013, 2019.

[8] J. B. He, J. Yang, and K. J. Ren, "Network Security Threat Detection under Big Data by Using Machine Learning," *International Journal of Network Security*, vol. 21, no. 5, pp. 768–773, 2019.

[9] L. C. Huang, C. H. Chang, and M. S. Huang, "Research on Malware Detection and Classification based on Artificial Intelligence," *International Journal of Network Security*, vol. 22, no. 5, pp. 717–727, 2020.

[10] J. Kim, J. Kim, and H. Kim, "Cnn-Based Network Intrusion Detection Against Denial of Service Attacks," *Electronics*, vol. 9, no. 6, pp. 916–937, 2020.

[11] M. Li, G. S. Xing, and Z. H. Wang, "Research on Real-time Online Intelligent Detection Technology of SQL Injection Behavior," *Journal of Hunan University (Natural Science Edition)(in chinese)*, vol. 47, no. 8, pp. 31–41, 2020.

[12] G. S. Mahmood, D. J. Huang, and B. A. Jaleel, "Achieving an Effective, Confidentiality and Integrity of Data in Cloud Computing," *International Journal of Network Security*, vol. 21, no. 2, pp. 326–332, 2019.

[13] C. Z. Mu, Z. Xue, and Y. Shi, "Command Sequence Detection Method based on BILSTM and Attention," *Communication Technology(in chinese)*, vol. 52, no. 12, pp. 3016–3020, 2019.

[14] X. L. Tao, K. C. Kong, and F. Zhao, "Internal User Security Behavior Evaluation Method based on LSTM," *Journal of University of Electronic Science and Technology of China(in chinese)*, vol. 48, no. 5, pp. 775–789, 2019.

[15] X. N. Wang and X. Y. Li, "A Research of Gcforest Methods for Network Abnormal Behavior Detection," in *2020 International Conference on Computer Engineering and Application (ICCEA)*, pp. 218–221. Guangzhou, 2020.

[16] Z. N. Wu, L. Q. Tian, and Z. G. Wang, "Network User Behavior Authentication based on Hidden Markov Model," in *2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE)*, pp. 76–82. Chengdu, 2021.

[17] L. M. Xia and Z. M. Xia, "A New Method of Abnormal Behavior Detection Using LSTM Network with Temporal Attention Mechanism," *The Journal of Supercomputing*, vol. 4, no. 77, pp. 3223–3241, 2020.

[18] P. S. Xie, C. Fu, T. Feng, Y. Yan, and L. L. Li, "Malicious Attack Detection Algorithm of Internet of Vehicles based on CW-KNN," *International Journal of Network Security*, vol. 22, no. 6, pp. 1004–1014, 2020.

[19] P. S. Xie, C. Fu, X. Wang, T. Fen, and Y. Yan, "Malicious Atack Pevention Model of Internet of Vehicles based on IOV-SIRS," *International Journal of Network Security*, vol. 23, no. 5, pp. 835–844, 2021.

[20] H. W. Xue, Y. Liu, and W. C. Zhuang, "Vehicle Abnormal Behavior Detection Method based on Stacking Integrated Learning in Internet of Vehicles," *Automotive Engineering (in chinese)*, vol. 43, no. 4, pp. 501–508+536, 2021.

[21] P. Yang. *Research on Reverse Analysis Method for Network Protocol Behavior Model(in chinese)*. PhD thesis, Harbin Institute of Technology.

[22] X. Y. Ye, S. S. Hong, and M. Han, "Feature Engineering Method Using Double Layer Hidden Markov Model for Insider Threat Detection," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 20, no. 1, pp. 17–25, 2020.

[23] Y. H. Zhang, Z. Xue, and Y. Shi, "Rebound Shell Detection Method based on LSTM and TF-IDF," *Communication Technology (in chinese)*, vol. 53, no. 12, pp. 3046–3050, 2020.

[24] X. Zhou. *Research on Improved TF-IDF Feature Selection and Short Text Classification Algorithm(in chinese)*. PhD thesis, Anhui University.

[25] Z. Q. Zuo, S. Yong, and X. Zhi, "Malicious Command Detection Method Based on Machine Learning," *Communications Technology(in chinese)*, vol. 53, no. 11, pp. 2775–2779, 2020.

Biography

Peng-shou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Security on Internet of Things.E-mail: xiepsh_lut@163.com

Yi-Fan Wang was born in Aug. 1996. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 844782234@ qq.com

Zong-liang Wang was born in Mar. 1997. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: 1292094887@ qq.com

Nan-nan Li was born in Feb. 1997. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 2500466296@ qq.com

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn

Yan Yan was born in Oct.1980.She is a associate professor and a supervisor of master student at Lanzhou University of Technology. Hermajor research field is privacy protection, multimedia information security.E-mail:yanyan@lut.cn