Research on Dynamic Reputation Evaluation Model of Machine Learning Based on Multi-features

Jingya Huang

School of Software and Applied Technology, Zhengzhou University, 450000

Keywords: Feature extraction; Machine learning; Reputation; Dynamic reputation

Abstract: Reputation evaluation is the public's credit evaluation of a series of historical behaviors of individuals, organizations or enterprises. It is a comprehensive evaluation formed by obtaining relevant feedback information from the public in social networks. Its key is to comprehensively quantify the accumulation of historical behaviors of entities. Feature extraction and multi-objective machine learning algorithm is proposed based on multi-objective cooperative FTEA. The algorithm finds out its kernel attribute by feature extraction of multi-attribute of learning samples, and the kernel attribute and other non-kernel attributes form an attribute group, thus improving the classification accuracy. This paper presents a fuzzy reputation level evaluation algorithm (FTEA) for generating evaluation rules and reputation levels. The algorithm adopts the rule-based machine learning method, which has the ability of self-learning from a large number of input data to obtain evaluation rules. In reputation management, the model adopts the cluster head management mechanism, which alleviates the problem of slow reputation convergence caused by complex computing, and is suitable for the network expansion needs. In the process of reputation update, the reputation information update of related nodes drives the reputation update of each role node.

1. Introduction

In an open network environment, how to maintain the security of the network and how to infiltrate security measures into it is a very complicated project. Traditionally, encryption, authentication and other security technologies have been used to achieve this goal [1]. Compared with traditional business activities, due to the anonymity and virtuality of the network, it is more difficult to establish trust in practical applications of e-commerce. Nodes in the network exchange information by using their own wireless transceiver equipment. When the days of each other are not within the communication range of each other, they can realize multi-hop communication by means of relay of other nodes, thus realizing the sharing of information and services [2]. Crowd sourcing has been repackaged on the basis of the open source concept and is also a new strategic model for providing problem-solving solutions that appeal to the public by leveraging individual hobbies and motivations. In fact, the security of the network can not only be considered from the perspective of ensuring user access security, but also the establishment of trust between users and resources [3]. In order to ensure resource security, the credibility of users needs to be evaluated; the selection and learning algorithms are closely related, and the results of feature selection are ultimately evaluated by the performance of the learning algorithm, so the combination of feature selection and learning algorithms is also very important. Therefore, how to effectively regulate the behavior of participants, eliminate the phenomenon of fraud for gaining benefits, and ensure a good platform environment has become an urgent problem to be solved by crowdsourcing virtual platforms.

2. Dynamic Reputation Evaluation Based on Multi-feature Machine Learning

Generally, the evaluation of reputation level needs to consider both subjective and objective factors. For example, the evaluation of users' credibility by resources is determined not only by the direct reputation obtained locally, but also by the indirect reputation recommendation (reputation) obtained by the computing agency. This model divides the credibility relationship according to the roles and functions of nodes, and evaluates the credibility of nodes according to their behaviors. At the same time, it takes full consideration of the cooperative relationship between the nodes of each role, and improves the viability and profitability of nodes by means of the cooperation between nodes, and serves as the basis for dynamically determining the credibility of nodes [4]. This process includes a series of closely linked links, and each link also needs to obtain corresponding resources such as target permissions. In the evaluation model of reputation evaluation, after each service interaction, a result is generated by the mutual services of both parties, and a corresponding evaluation of each other is given based on the result, and a reputation value of both parties is recalculated based on the evaluation [5]. But these factors are difficult to measure quantitatively, and the fuzzy reasoning mechanism in the fuzzy mathematics theory can solve this uncertainty well. Therefore, we establish a fuzzy dynamic reputation evaluation system FTEs based on machine learning.

Generally, mobile application behavior needs to acquire resources through a series of operations to complete specific functions. From the perspective of system permission, operation execution, that is, calling system functions, needs to obtain specific permission. In other words, there is a corresponding relationship between operation and permission [6]. Whether a node can safely and smoothly perform its responsibilities in the routing process according to its role requires the close cooperation of related nodes. Any relationship rupture may bring unstable factors to the routing, and even have a negative impact on the performance of the whole network. Therefore, it is considered to improve the calculation process of the reputation value of the reputation service module, reduce the workload of the reputation service module as much as possible, improve the overall performance of the system, at the same time, correctly reflect the change trend of the reputation of the service, and punish the illegal behaviors in the service process [7]. In the structure of FTEs, the evaluation system obtains the evidence about the historical status of user resources in the evidence database, combines the existing rules in the knowledge base, deduces to obtain the direct reputation, and then carries out fuzzy synthesis on the direct reputation and the indirect reputation from the evidence base to obtain the user resource reputation list (URTL). In this process, if the currently input evidence cannot match any rules in the existing knowledge base, a new rule is generated and added to the knowledge base.

The main dimensions of the multi-dimensional reputation evaluation model include subjective evaluation and punishment, which mainly reflect the subjective evaluation users receive when publishing and completing tasks and whether they are punished for illegal and malicious operations. The rest of the dimension information is auxiliary information, in order to give a comprehensive and intuitive understanding of the platform users' comprehensive reputation value. The cumulative subjective evaluation value is calculated as follows:

$$S_{\mathrm{j}} = \frac{1}{\sum_{i=1}^{n}(S_r)}, (0 < (S_{\mathrm{j}}) \leq 1) \tag{1}$$

S indicates the rating value of the evaluator, and Sr is the reputation parameter factor. In order to prevent the extreme brushing operation of the malicious user, the high reputation may be poor, and the low reputation may not be poor; the comprehensive multi-dimensional reputation evaluation model is as follows:

$$S(r_k) = \sum_{r_k \neq r_i} w(r_i) D_r(r_k, r_i) \tag{2}$$

First, the input and output of FTEs should be established, and the experience database should be established as the training data of the system. These training data come from the data source: monitoring center. It is responsible for collecting common parameters in the process of user and computing resource interaction. After the behavior sequence is updated, each role node calculates the behavior reputation according to the quality within the routing nature cluster or between the clusters and the quality of its forwarding. The cluster members have a cooperative relationship due to the routing behavior of the cluster head. The change of the cluster head behavior will inevitably bring about the update of the member reputation [8]. For example, a certain software needs to call the contact information to the server. If the contact is first invoked and then the network operation is performed, the acquired resource is an unupdated contact, and the updated contact information is obtained, thereby obtaining different contacts. information. Obtaining parameters such as the success rate of resource execution of user jobs, execution time, online time of resources, self-defense capability of resource sites (security level), communication security level, reputation recommendation of resources to users, etc. [9]; Assuming that crowdsourcing platform uses a single reputation evaluation mechanism, each participant exists in the platform for the purpose of maximizing its benefits, and some participants may blindly give publishers extremely high or very low evaluations for the purpose of benefits [10]. Evolutionary algorithm realizes global search by maintaining a population of potential solutions between generations. This population-to-population method is very useful for searching the optimal solution set of multi-objective optimization problems.

After obtaining the operation tasks of open source mobile applications, the designed algorithm is used to simulate the resource acquisition in the applications. Both node reputation algorithm and pruning algorithm are implemented in C++language. During the experiment, different model test parameters were designed to verify the model and algorithm, which ensured that the test parameters of the model were different during each operation. The experimental design parameters are shown in Table 1.

Table 1. Parameter design of simulation experiment

Experimental sequence	Number of atomic operations	Number of nodes	Reputation value
First time	20	15	9
Second time	30	20	10
Third time	40	25	12

The reliability of atomic operations is quantified by reputation values, reflecting the possibility of the user selecting which atomic operations to acquire resources. For m in the model σ, the reputation value σ of the atomic operation is:

$$\sigma = \sqrt{\ln(1+\frac{\mathrm{v}_r}{m^2})} \tag{3}$$

Frequent use reflects the trust in atomic operations to a certain extent, and the more you use it, the more you experience the mobile application, which means the more proficient the use. Therefore, the historical reputation value is affected by three factors, and the influence of the three factors on the historical reputation value shows a linear relationship, as follows:

$$\mu = \ln(\frac{m^2}{\sqrt{v_r + m^2}}) \tag{4}$$

Therefore, the new evaluation value is affected by three factors, and the influence of the three factors on the new reputation value shows a linear relationship, as follows:

$$M_{\mathrm{k}} = c_k^d M_{k-1} \tag{5}$$

The value of node credibility increases gradually from the leaf node to the root node. The reason for this is that as the number of resource nodes acquired by users increases, the level of permissions acquired by users will also increase, thus the mobile application behavior also has higher approval. According to the design parameters of the first experiment, the dynamic graph of the node credibility of the sub-credit model is generated to dynamically evaluate the credibility of the resource, as shown in Figure 1.

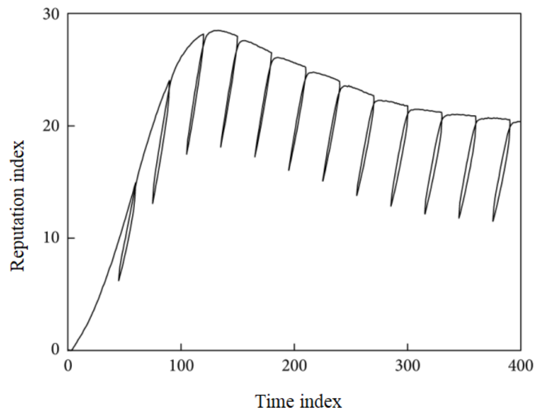

Figure 1. Node credibility dynamic change graph

The fuzzy reputation level evaluation in FTEs adopts the algorithm FTEA, which includes two parts: evaluation rule generation algorithm (ERGA) and direct reputation level inference algorithm (DTRA). It corresponds to direct trust. There is no direct interaction experience between two entities, but based on the recommendation of other entities establishes a trust relationship. The recommended trust value is based on the evaluation of other entities and belongs to indirect trust. Therefore, feature selection has a non-negligible effect on learning algorithms in different situations. Choosing good features not only reduces computational complexity, improves classification accuracy, but also helps to find more streamlined and understandable algorithm models. By the same token, publishers will use profit means to drive malicious participants to obtain extremely high reputation values for them. All these phenomena will lead to imbalance in the credibility system of crowdsourcing platforms, thus undermining the stability of the platforms. Similarly, when different operations with the same function replace each other, the information to be obtained is also different. The node reputation layer is used to calculate the node reputation and the impact factor of the node reputation on the cluster reputation. The cluster reputation layer is used to find the cluster members, calculate the cluster reputation and calculate the gateway reputation. The gateway reputation layer is used to find the relevant clusters and calculate the gateway reputation. The reputation service module receives the application of service provider participating in service evaluation received by the service registrar, obtains objective data and subjective data of service evaluation through detection and feedback, and calculates the data to obtain the reputation of the required service.

3. Result Analysis and Discussion

Node reputation is used to evaluate the cooperative ability of nodes to process messages in the routing process. When a node first joins the network, it has not yet received any information from other nodes. At this time, it needs to assign an initial reputation value to the node. The system finds a set of fuzzy rules through some self-learning method to reproduce the input and output behaviors of the system, and evaluates the future input data directly according to these rules. The reputation value is based on the performance of the mobile application over a period of time and the user's evaluation of it. Develop an effective incentive mechanism to attract a certain group of users to participate in an active manner. Excellent quality control strategies can effectively eliminate the garbage results and prevent or detect malicious participants to a certain extent, thus ensuring the stability of the crowdsourcing platform. And availability. Trust is not simply believed or unbelievable. It can be measured by trust degree. The range of general trust is set in [0,1]. The closer to 1 is, the more trust, the closer to 0, the less trust. If the evidence to be tested cannot be To effectively match any rule in the knowledge base, the system gives an inference result that approximates the existing rules in the knowledge base, and then combines the input tag combination and the output inference result into a new rule and adds it to the knowledge base. Principal component analysis is the most famous algorithm in this kind of methods. It can reduce dimensions for many learning tasks, but the understanding of features is very poor, because even simple linear combination can make the constructed features difficult to understand, and in many cases, the reasonable solution of features is very important.

If node k has i pre-resource nodes, corresponding nodes use i ($0 \le 1 \le k$), and the corresponding nodes of the pre-nodes are represented by SN, and the relationship between them is AND, the credibility of node k is:

$$\mathrm{k} = \frac{S_N(k)}{\sum S_N(k_i)} \tag{6}$$

If node k has x front-end resource nodes, the corresponding node is represented by I ($0 \le 1 \le k$), the corresponding edge of the front-end node is represented by I, and the relationship between them is or, then the credibility of node k is:

$$\beta_{xy} = \frac{I_{xy}}{\sum_{a-1}^{g-1} I_{xa}} \tag{7}$$

If node K has only one pre resource node, the reputation of node R is:

$$R_i = M_i \sum_{i=1}^{N_r} \hat{R}_{i,r} \tag{8}$$

If node K does not have a pre-existing resource node, that is, node I is a leaf node, the credibility of node K is specified as 2. Among them are:

$$I = I_0 + \sum_{i=1}^{n} A_i \exp\left(\frac{-t}{\tau_i}\right) \tag{9}$$

This process of assigning initial reputation value is actually a subjective process. Nodes assign different initial values to different nodes by virtue of their own experience. Under normal circumstances, a small number of the first networking nodes will trust each other and believe that they are all trustworthy. That is, when the difference between the existing rules and the evidence to be tested is greater than 1, we think that there are significant differences between the two parties. Therefore, the logical conclusion of the existing rules cannot be directly used as the final direct credibility evaluation result, but should be modified to some extent. Analysis shows that the historical reputation value of atomic operation is reflected by the usage time. The longer the usage time, the higher the reliability of atomic operation. According to the user's daily such as login, interaction, completion of tasks to obtain the corresponding points, high score users will have unique identification, and therefore will obtain more operating rights, thus to encourage users to actively participate in the website. For a fixed sample set, the number of inconsistencies in a pattern is defined as the difference between the number of samples in this mode and the number of samples in the category with the largest proportion. The resource provider constructs the income function according to the consumer's bid price, trust degree and completion time constraints, provides differentiated services, and encourages the nodes to share resources and obtain higher trust. Therefore, it should allow a small amount of initial evidence to be collected as training. The sample, in the process of using the knowledge base for fuzzy reasoning, constantly expands the inference rules in the knowledge base.

The following takes Best Buy,Circuit City model as a reference, and verifies and analyzes the validity and feasibility of its reputation service module model by comparing the influence of malicious behaviors on reputation calculation under malicious services. In the experiment, the fraudulent behavior of malicious services was analyzed. The average error of each model is shown in Table 2. The experimental results are shown in Figure 2.

Table 2. Average error of each model

Average error	Best Buy	Circuit City	Reputation service module
No conspiracy (maximum)	0.341	0.225	0.334
Conspiracy (minimum)	0.145	0.471	0429

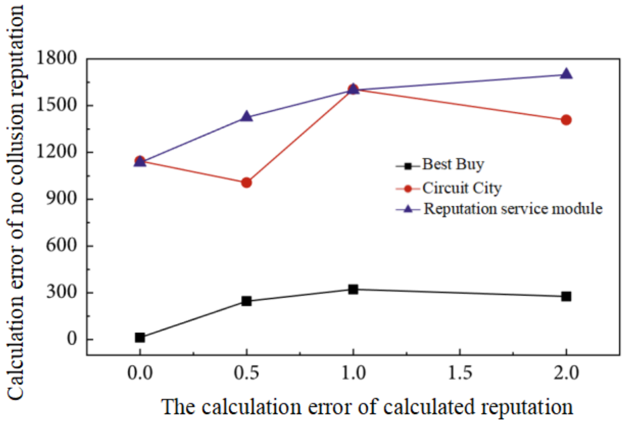

Figure 2. There is no calculation error of conspiracy reputation and calculation error of conspiracy reputation.

In real life, the credibility of an entity to another entity is established through previous contacts between the two. At the beginning, there is a basic initial value of credibility between entities that have not been contacted, and the level of trust will change through the assessment of each contact. The time factor is reflected in the time of obtaining resources. The shorter the time, the higher the new evaluation value of atomic operation. The user feedback factor is reflected in the user's expectation of the use result. The higher the expectation, the higher the new evaluation value brought by the user feedback. If the content involves violation of the law or social ethics, the administrator will delete it and punish it accordingly. After all, the administrator's review power is limited and can only be controlled in the general direction. Moreover, since the feature selection algorithm using the inconsistency evaluation needs to consider the combination of the values of various features under the combination of various feature subsets, the calculation cost is high for high-dimensional data. The research of reputation based grid resource management and scheduling is still in the initial stage. In the design process, the actual performance, reliability, availability and other characteristics of resources need to be considered as a whole, and resources need to be effectively managed and reasonably scheduled according to different needs. The local FTEs can be used to normalize the reputation level evaluation results of all recommenders to obtain the weight value as the credit weight of each indirect reputation recommendation.

4. Summary

In the network computing environment, in order to ensure the security of interaction between users' resources and improve users' satisfaction with the services provided by resources, an automatic rule generation algorithm is designed based on multi-sample data and using machine learning method. Trust is a complex social relationship and subjective psychological cognition. A clear definition of trust is the basis of related properties, expressions and measures. Existing researches propose models for their respective definitions. Trust involves many aspects such as security and resource performance in grid environment. In this paper, we design and implement a multi-dimensional reputation evaluation model based on multi-feature machine learning. By collecting and analyzing the behavior of system users, we transform the scores into corresponding dimensions, improve and optimize the single problem of traditional reputation evaluation, and analyze the user's credit value from different aspects. The model is based on resource and atomic operation, and credit value and credit degree are introduced to measure the reliability of atomic operation and resource respectively. The main contribution of this model is to establish a reputation relationship network through the node behavior and the node cooperation in the routing process, so that the security evaluation is more accurate and reasonable to reward the good cooperation behavior of cluster members to cluster heads, and enhance the probability of cluster members participating in network activities.

References

[1] Shao Zhipeng, Lareina C, Ma Yuanyuan, et al. Research on credibility evaluation model of mobile applications based on operation behavior [J]. Journal of Taiyuan University of Technology, 2018, 49(06):46-52.

[2] Zhang Chaoning, Zhang Ying, Xu Yingtao. Online Reputation Evaluation Model Based on Discrete Time Optimal Control [J]. Journal of Zhejiang Normal University (Natural Science Edition), 2018, 41(04):29-35.

[3] Xing Xuliang, Tian Xiuxia. A novel reputation-based dynamic access control model [J]. Journal of Shanghai Electric Power University, 2017(06):80-85.

[4] Shang Yuwei, Ma Zhao, Peng Chenyang, et al. Exploration of machine learning method with embedded professional knowledge and experience (2): application and practice of guided learning [J]. Journal of China Electrical Engineering, 2017(20):26-35.

[5] Jiao Jiafeng, Liu Yun. Overview of Typical Machine Learning Platforms under Big Data [J]. Computer Applications, 2017(11):7-15+20.

[6] Shang Yuwei, Ma Zhao, Peng Chenyang, et al. Exploration of machine learning method with embedded professional knowledge and experience (1): the introduction of guided learning [6] and theoretical basis [J]. Journal of China Electrical Engineering, 2017(19):44-55+317.

[7] Dong Qingling. Machine Learning and Conflict Prediction-An Interdisciplinary Perspective in International Relations Research [J]. World Economy and Politics, 2017(07):102-119+160-161.

[8] Dai Tingting, Han Yan, Hu Xiaofei. Personal Credit Evaluation Based on PLS-SVM Algorithm [J]. Journal of Dehong Teachers College, 2017(26):107.

[9] Hu Xiangdong, Xing Quanquan, He Wenxiang. WSN Energy Efficient and Secure Clustering Algorithm Combining Reputation Evaluation and Patrol Mechanism [J]. Telecommunications Technology, 2019, 59(02):5-12.

[10] Wang Qing, Yao Kang. Research on SME Credit Evaluation Method Based on Stochastic Subspace Ensemble Learning [J]. Shanghai Management Science, 2018, v.40; No.228(03):98-101.