

A Fraud Identification Model to Detect Deception for Online Shopping Networks

Qixia Sun

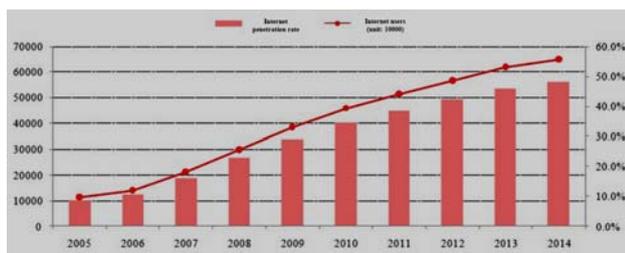
1. School of Economics and Management, China University of Geosciences, Wuhan 430074, China.
2. School of Economics and Management, East China University of Technology, Nanchang 330013, China.
E-mail: huntercet@qq.com

Abstract - To promote the healthy development of e-commerce and maintain the stability of the economy and society, it is particularly important to study online shopping fraud, which has a strong practical significance. Firstly, we use the framework of law case base to calculate the similarity of the case using the nearest neighbor method. This facilitates the determination of the weights by proposing a design method based on the analytic hierarchy process to make adjustments, which leads to the identification of the case with high confidence as fraud in online shopping.

Keywords - Online shopping, fraud identification, deception mechanism

I. INTRODUCTION

CNNIC, Chinese iNternet Network Information Centre <https://cnnic.com.cn/>, shows that the number of Internet users has exceeded 640 million in January 2015, Internet penetration rate has reached 47.9%, and compared with 2013, the popularization rate of the increase was 2.1%. The number of mobile phone users has reached 557 million in the last year, compared with 2013, the growth exceeded the number of 56 million, and the number rose to 85.8%. The rural users accounted for 27.5%, which reached 178 million, increased to 1.8 million. The popularity of the family WiFi reached a higher level, the proportion was 81.1%, and the family WiFi has a strong leading role for the elderly family members in the Internet, which can enhance the promotion of city Internet penetration rate [1,2].



Notes: The data comes from CNNIC

Figure 1. Internet penetration rate in China

Fraud refers to deliver parties intended to conceal or provide false information associated with the transaction to obtain economic benefits, is a kind of perceived risk. The reasons for online auction fraud can be explained by "asymmetric information". The seller has more information about the transaction object than the buyer. In the case of asymmetric information (Akerl, 2012) [3], which causes the online seller to easily provide inappropriate product information, thereby increasing the risk of fraud. The problem of asymmetric information in online auction is

particularly prominent in the online trading of most goods than traditional brand products or second-hand products, and cannot see the kind, only consumers are evaluated according to the specific information provided by the seller (De Figurer, 2013) [4], this feature has produced a typical "lemon" problem. The anonymous existence, the information asymmetry in online auction is not only in the product area, but also in the seller's location, scale, identity and so on, which also makes it difficult to investigate and deal with it.

II. STATE OF THE ART

The types of online auction fraud include payment, non-delivery, and intentional delivery of false information, concealment of fees, illegal or inferior products, long auctions, and the hiring of false bids (Gregg, 2014) [5]. The seller also uses the reputation system for fraud, such as accumulating high credit values through small transactions, and then swindling in large transactions. Factors affecting online auction fraud include three aspects of products, traders, and trading processes (Macinnes, 2015) [6]. Domestic research on online reputation system involves system construction, mechanism design and related factors put forward five major factors related to a product reputation system, namely online product utility evaluation, physical characteristics, online product online product content, online products and the producers of the prior network diversity (Song Guangxing, 2014) [7]. The relevant literature (Zhang Wei 2015) [8] on the effectiveness of the existing trust model in online reputation system is analyzed, five factors of trust proposed the influence of online auction trust, including user feedback score, recent trust score, people trust, transaction value and evaluation time weight, and constructs C2C in online auctions the computational model of trust. Researcher use the theory of incentive constraint mechanism in game theory, a reputation rating alliance mechanism is designed to promote the role of online reputation system.

III. METHODOLOGY

The deception mechanism of network fraud mainly involves the following methods that semantic network method, frame representation method, XML method and object oriented method.1) Semantic network uses graph to express knowledge, and its logical structure is directed graph of two element relation. It can express directly the structure relation between concept and even entity, it includes arc and node. Whether arc or node, they have some weight, which is a representation of the degree of correlation. For complex knowledge structures, the method can also express the representation, which is widely used in organizations with relatively distributed knowledge resources. This method is used in the information resource organization of Internet, and then a brand-new knowledge network is produced.2) Frame representation method has very clear structure, can be very good description of the case can be retrieved, so this method is the use of a very wide range, for a large and complex problem when this method is particularly suitable for reasoning. 3) XML is created by the Internet and organizations, has a number of syntax rules that can be used to describe data. For XML, it is a text that uses attributes and flags to describe the nature of the data. In addition to describing the relationship of the data, the method also stores the corresponding data.4) Object oriented representation method is realized and the case of instantiation, which is a reality of the abstract. It is a specific object set, which includes the service and the same properties, and the framework is to encapsulate the behavior and properties. The method used is object-oriented in the case, then the storage space is only the case, and the corresponding solution and scheme only occupy the common storage space.

IV. THE CONSTRUCTION OF FRAUD IDENTIFICATION MODEL

The method is to describe the characteristics of fraud cases in online shopping, and to integrate the case into feature library. In case data organization, its attribute is the representation of its basic situation, mainly including the time, cause, experience and so on. First of all, we should collect and build a case library of online shopping fraud. Suppose we can collect n online shopping fraud cases, then the original case number of this case library is written as n,

and the case library can be represented as represents the I case in the library, represents the set of features that has, the n is represented by fin, and its solution set uses that is the n [36] in the corresponding to the solution.

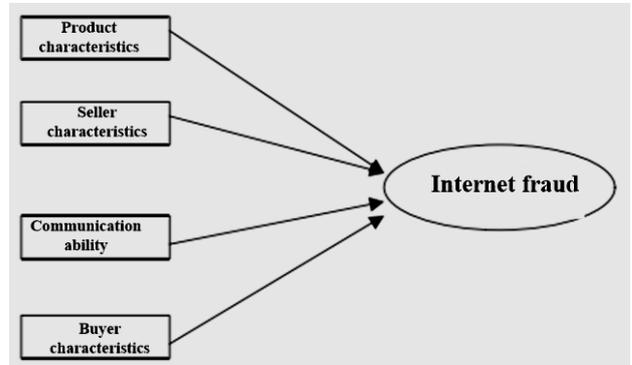


Figure 2. Network deception model.

A. Case Representation of Fraud

The case can be generally described using that problem, solution description, description, which is three tuple representation, similarly, fraud case reasoning decision model in the online shopping, the case also includes three aspects: the description of the characteristics of fraud case has the problem is described; solve the fraud case solution is the solution the description, description of solutions of fraud cases corresponding to the result is the effect. In case of fraud in online shopping, a fraud case is composed of many kinds of complex information level, in the similarity of case are defined, which is based on the similarity of attributes in case of. The frame representation structure is relatively clear, which can well describe the case, easy to search, and can be very good to deal with large complex problems. Therefore, this paper uses frame representation method, the method in hierarchical relationship based application, appear in the online shopping fraud, using this method not only said it is feasible, and is very suitable.

TABLE I. FRAUD CASE FRAMEWORK

Frame name	fraud case name in online shopping	Frame name	fraud case name in online shopping
Level 1	fraud information description	Level 2.2	Fraud characteristic 2 (index 1, value 1, weight 1; index 2, value 2, weight 2...)
Level 1.1	types of fraud	Level 2.n	Fraud characteristic n (index 1, value 1, weight 1; index 2, value 2, weight 2...)
Level 1.2	fraud description	Level 3	result set
Level 1.3	reasons for fraud discovery	Level 3.1	How long has fraud been discovered?
Level 2	characterization of fraud behavior	Level 3.2	loss caused by fraud
Level 2.1	Fraud characteristic 1 (index 1, value 1, weight 1; index 2, value 2, weight 2...)	Level 4	relevant knowledge

B. Network Fraud Case Matching Calculation

The cumulative feature is similar for each case and target case in case base, the overall similarity was determined, according to the similarity of the corresponding set of threshold. The case similarity exceeds the threshold, the case will be returned to the user. The similarity of attributes is expressed according to the distance of attribute values, when the global similarity is expressed, the representation method is weighted by the similarity of attributes, and the formula is as follows:

$$Sim(x_i)^n = \sum_{j=0}^n \binom{n}{j} x \tag{1}$$

The i^{th} case targets the case similarity on the whole, which is the attribute index used for feature matching retrieval. The importance of the j attribute to its weigh W_j , the total of weight is 1, in order to sure, which can use hierarchical analysis method, probability, and expert evaluation method. The x case is the case database, the similarity is given by the attribute. The guide strategy will use heuristic information with information gain in which, which is using the nearest neighbor algorithm. The maximum of its shortcomings is the search using the fixed attributes, and classified guidance strategy undertook improvement to some extent in practice, the ID3 algorithm is the most frequently used. The theoretical method has the least expected attribute but for the heuristic function with the information gain. The classification attribute is unable to find the minimum value, which is to measure the similarity of case, analyze its characteristics and practical situation. The most important thing is to analyze the case, compare the feature attributes and weights to see whether it is consistent. The method of measurement used is based on attribute characteristics. In this paper, we use similarity to measure the case retrieval, and the process is as follows: (1) the query is suitable for the case, in order to get the most suitable collection of the two cases. In terms of the same attributes, there may be many cases that are appropriate for the observed values, and these appropriate cases form the set of cases. (2) Case matching uses a case matching algorithm in retrieving, and sets the appropriate case. Firstly, there are attributes similarity calculation, also is the local similarity, similarity calculation global, finally obtains the most similar case, the specific process is as follows that the problem of case attributes → weights are given → the calculation problem of case and case source attribute similarity → global similarity calculation → global similarity ranking, screening → case.

C. Assessment Result

The selection of attributes based on extensive literature study and combined with the experience of the research

conclusion, integrating different researchers, summarized the five factors are that the degree of standardization, product business credibility, business communication, consumer characteristics and the degree of fraud. The extensive literature research, we according to the degree of fraud from low to high order, divided into the receipt of the goods but to return service are not satisfied, but to receive the goods arrival status not satisfied but to receive the goods, the consistent degree of dissatisfaction, but not with the receipt of the goods, the delivery time of the seller after the money is not delivered 5 degree of fraud. According to the above analysis, the above 5 attributes can be quantified low. The product is a standard product, which the seller is divided into 20 levels.

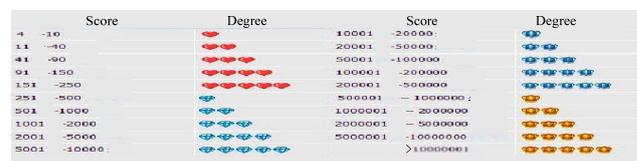


Figure 3. Credit Rating Map

All in all, we is according to the case based reasoning of online shopping fraud identification model test, this case is judged to be fraud case. This cell phone is to the Samsung aftermarket identified, and concluded that this cell phone is indeed a fake and shoddy product. the similar case detection value is not equal to 1, that the case is not exactly the same case, the similarity is between the threshold and alpha 1, the system will enter the case of the learning process, which is according to the successful case of the combination of expert experience. The seller is not too difficult for the indefensible, logistics information display the seller receive the goods after the entry, automatic refund procedures, time after the system will automatically return money to the author. This paper constructs the recognition model of case, which based on the case of online shopping fraud, and preliminarily verified by examples, the result is true and reliable, proved that this model can be applied to the analysis and judgment on the case of online shopping fraud.

V. CONCLUSION

This paper introduced the method of distance education network extending to all disciplines, many university teachers are trying to integrate network teaching mode in the information age into the classroom, which is high efficiency and convenience brought by the development of society, is worthy of promotion. Compared with the simple teaching in the classroom, the new modern teaching mode which combines the network with the classroom has unique advantages. This paper focuses on the research in the relevant technical support, according to the requirements of the parties, the creation of a Tibetan folk dance teaching management information system, this paper has obtained

some achievements, first analyzed at home and abroad about the Tibetan folk dance teaching management information system research present situation and the direction of development, on the basis of the teaching mode as a tool to network feasibility study, at the same time from the user demand analysis of the role and significance of the Tibetan folk dance teaching management information system. At the same time, the network technology involved in dance teaching management information system is discussed. The.NET platform enables the system to run across platforms because of its good compatibility, which has positive implications for future extensions. At the same time in order to speed up the progress of complete system, C#, session state management, validation and some other advanced technologies have been introduced to the system construction, greatly uninstalled the Tibetan folk dance teaching management information system.

The many domestic and foreign scholars research reasoning theory and practice, but the research fraud in online shopping is nearly blank. This paper analyzes the research background, and through the depth of the literature research, detailed analysis of the related theory of online shopping fraud and case based reasoning, which build online shopping fraud recognition model based on case reasoning, fully describes the principle. The operation process and application characteristics, and select the larger of fraud case for a simple test on it, this model of suspected fraud the behavior encountered in the process of online shopping were judged more scientific and accurate, and

provides some solutions. Therefore, this paper based on the protection mechanism of the United States, the European Union and Japan and other countries advanced on, according to China's national conditions in China are put forward in the network of shopping consumer protection measures mainly from the aspects of legislation, construction, system and credit system are expounded.

REFERENCES

- [1] Trung Dong Huynh, Nicholas R, "An integrated trust and reputation model for open multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 2, pp. 109-152, 2016.
- [2] WEI Min, "A Research on the E-commerce Applied to the Construction of Marketing Model," *Journal of Sichuan University*, vol. 3, no. 4, pp. 80-88, 2017.
- [3] Zhang Yu. A resource allocation mechanism of trust incentive compatibility in grid environment. *Software journal*. vol. 11, pp. 42-45, 2015.
- [4] Gilboa, D. Schmeidler, "Case-based Decision Theory," *Quarterly Journal of Economics*, vol. 7, no. 2, pp. 600-636, 2015.
- [5] Deng P S, "Using Case-based Reasoning Approach to the Support of LII-structured Decision," *European Journal of Operational Research*, vol. 93, pp. 509-519, 2016.
- [6] Akerlof, G, "The market for "lemons": Quality under uncertainty and the market mechanism," *Quarterly Journal of Economics*, vol. 84, no. 3, pp. 488-500, 2015.
- [7] Mishra, D.P.; Heide, J.B.; and Cort, S.G, "Information asymmetry and levels of agency relationships," *Journal of Marketing Research*, [J] vol. 35, no. 3, pp. 277-295, 2014.
- [8] De Figueiredo, J.M, "Finding sustainable profitability in electronic commerce," *Sloan Management Review*, vol. 14, no. 4, pp. 41-52, 2016.