Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 4, April 2015, pg.92 – 95

RESEARCH ARTICLE

Credit Card Fraud Detection Using Decision Tree Induction Algorithm

[1]Snehal Patil, [2]Harshada Somavanshi, [3]Jyoti Gaikwad, [4]Amruta Deshmane, [#]Rinku Badgujar

Department Of Computer Engineering, JSPM's Bhivrabai Sawant Institute of Technology and Research, Wagholi Pune

[1] svpatil008@gmail.com, [2] harshada2910@gmail.com, [3] jrgaikwad121@gmail.com,

[4] deshmaneamruta55@gmail.com, [#] rinku.badgujar@gmail.com

Abstract-- Today Fraud is increasing all over the world, resulting in the vast financial losses. Chip and pin are developed for credit card systems, through fraud prevention mechanisms and these mechanisms do not prevent the most common fraud types such as fraudulent credit card usages over virtual POS (Point Of Sale) terminals or mail orders so known as an online credit card fraud. So as a result, fraud detection becomes the important tool and possibly the best way to stop such types of fraud. A new cost-sensitive decision tree approach which reduces the sum of misclassification costs while selecting the splitting attribute at each non-terminal node is advanced and the act of this approach is compared with the well-known traditional classification models on a real world credit card data set. The data mining layers prevent fraudsters to attack and improve a safe transaction. This research is totally concerned with credit card application fraud detection by performing the process of asking security queries to the persons intricate with the transactions and as well as by eliminating real time data faults.

Keywords— E-Commerce Security, Credit Card Fraud Detection, Data Mining, ID3 Decision Tree, Visual Cryptography

I. INTRODUCTION

The use of credit cards has increased, because of a rapid advancement in the electronic commerce technology. The credit card becomes the most popular type of payment for both online as well as regular purchase, cases of credit card fraud also increasing. In Modern day the fraud is one of the major effects of great commercial losses, not only for merchants, the individual clients are also affected.

Credit Card Fraud: Credit card fraud has been distributed into two types: Offline fraud and On-line fraud.

- The Offline fraud is dedicated by using a stolen physical card at call center or any other place.
- The On-line fraud is dedicated via phone, shopping, web, internet or in absence of card holder.

In the commercial practice a large-scale data-mining techniques can improve on the state of the art. The scalable techniques to analyze massive amounts of transaction data that powerfully compute fraud detectors in a timely manner is an important problem, especially for e-commerce. Moreover scalability and efficiency, the fraud-detection job exhibits technical difficulties that include slanted distributions of training data and non-uniform cost per error, both of which have not been usually studied in the knowledge-discovery and data mining community.

The use of machine learning in fraud detection has been an exciting topic. Due to the confidentiality of financial information and non-availability of public databases, few researches have the chance to work on developing methods exact to credit card fraud detection. However, the literature on credit card fraud detection is increasing and it has been shown that machine learning can be used effectively for this problem, in particular: neural

networks, artificial immune systems, association rules, Bayesian learning, support vector machines, and peer group analysis.

II. IMPLEMENTATION

A. Decision Tree Induction algorithm

A Decision tree algorithms are a method for approaching discrete-valued target functions, in which the learned function is denoted by a decision tree. These types of algorithms are famous in inductive learning and have been successfully applied to abroad range of tasks. We examine the decision tree learning algorithm – ID3.
The decision tree is a structure that contains root node, branch and leaf node. Every internal node indicates a test on attribute, every branch indicates the outcome of test and each leaf node holds the class tag. The uppermost node in the tree is the root node .A Decision trees organize circumstances by sorting them down the tree from the root to some leaf node, which delivers the classification of the instance. Each node in the tree specifies a test of some attribute of the instance and each branch descending from that node links to one of the possible values for this attribute. For example figure below explains a decision tree based on attribute name outlook.

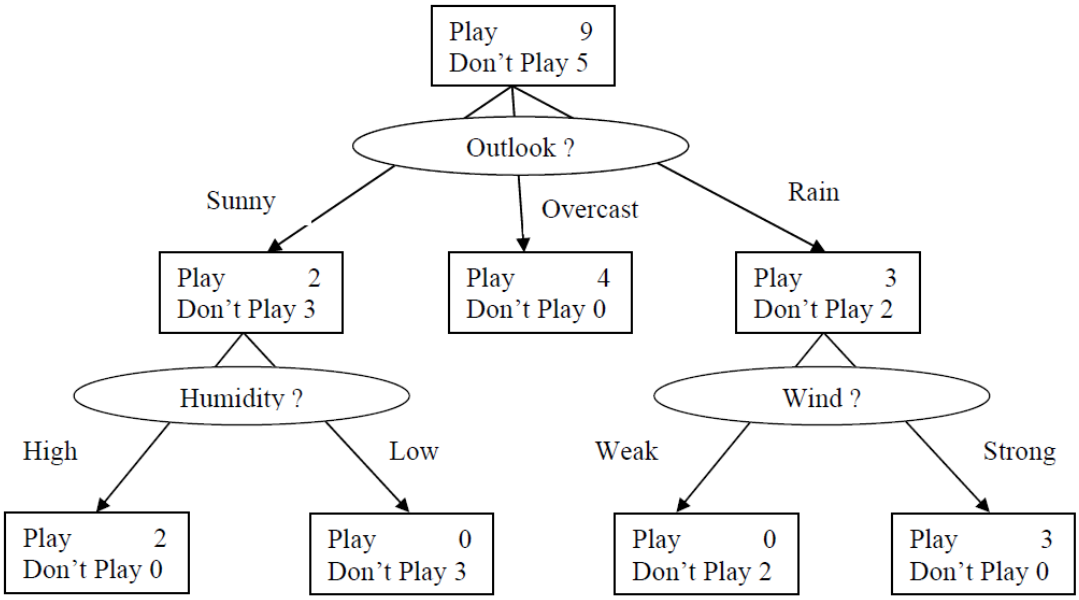


Fig. Decision Tree

1.1 Entropy

The entropy is a measure in the information theory, which illustrates the impurity of an arbitrary collection of examples. For e.g. if training data has 14 instances with 6 positive and 8 negative instances, the entropy is calculated as

Entropy ([6+, 8-]) = −(6 /14) log2 (6 /14) − (8/14) log2 (8/14) = 0.985

A key point to reminder here is that the more uniform is the probability distribution, the greater is its entropy.

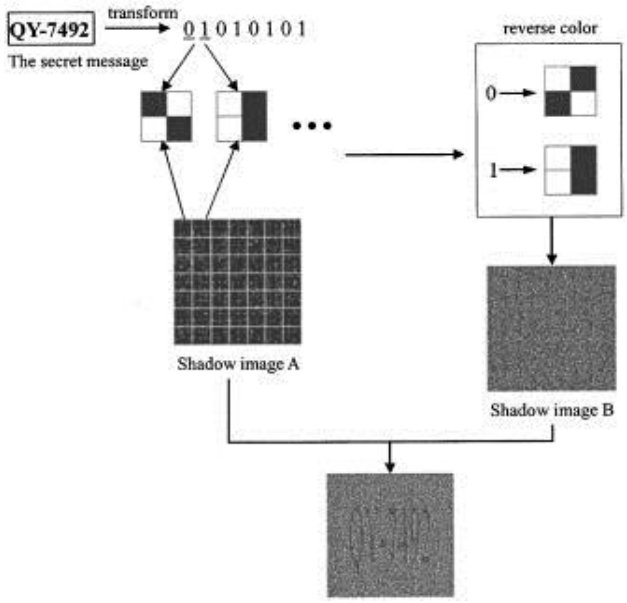
1.2 Information gain

The information gain measures the likely reduction in entropy by partitioning the examples according to the attribute.

B. Visual Cryptography

A visual cryptography is a kind of cryptography that can be decoded openly by the human visual system without any special calculation for decryption. The visual information text, allows through a cryptographic technique which is known as Visual cryptography to be encrypted in such a way that decryption becomes a mechanical operation that does not require a computer.
Visual cryptography is a popular solution for image encryption. The encryption procedure encrypts a secret image into the shares which are noise-like secure images which can be communicated or distributed over an untrusted communication channel by using secret sharing concepts. The secret image is decrypted without additional computations and any knowledge of cryptography by using the properties of the HVS to force the

recognition of a secret message from overlapping shares. The Visual cryptography is proposed in 1994 by Naor and Shamir who introduced a simple but perfectly secure way, which allows secret sharing without any cryptographic computation, which they known as Visual Cryptography Scheme (VCS). The simplest Visual Cryptography Scheme is given by the idea of a secret image contains of a collection of black and white pixels where each pixel is treated individually. There are many algorithms to encrypt the image in a different image, but a exceptional of them have been in visual cryptography for colour image. So we are using visual cryptography method.



Visual cryptography uses the characteristics of human vision which is an emerging cryptography technology, for decrypting the encrypted images. It wants neither complex computation nor cryptography knowledge. For security concerns it also ensures that hackers cannot perceive any clues about a secret image from individual cover images.

III.RESULTS

The procedure of decision tree for classification problems includes two steps: using a training data set to construct a decision tree; for each of the elements, applying decision tree to determine the elemental groups. Table shows an example includes the ten articles about fraud and related information, and records as data set S, now ID3 algorithm is used to create fraud classification decision tree about the credit card:

Number (Not as decision attribute)	Sex(Sex)	Marital Status(Marraige)	Level of Education(Education)	Place of crime(Place)	Is It Fraud
1	Man	Unmarried	High school and below	Bank	Yes
2	Women	Married	Junior	Supermarket	No
3	Women	Unmarried	Undergraduate	Bank	No
4	Man	Married	Undergraduate	Other	No
5	Man	Other	Graduate and above	other	No
6	Women	Unmarried	High school and below	Bank	Yes
7	Man	Married	Undergraduate	Supermarket	No
8	Man	Other	Junior	Bank	Yes
9	Man	Unmarried	High school and below	Other	No
10	Women	Married	Undergraduate	Supermarket	Yes

Table. Data Set Of Fraud Information

Every data has five properties in the table, which has four attributes need to calculate their information gain, and takes this as the source to build a decision tree, the four properties are:
Sex, Marital status, Level of education, Place of crime.

IV. CONCLUSION

As Credit card fraud has become more and more widespread in recent years. To increase merchants risk management level in an automatic and active way, structure an perfect and easy handling credit card risk monitoring system is one of the key tasks for the merchant banks. So we propose credit card fraud detection problem for the resolution of reducing the bank's risk. With the historical profile patterns, make use of credit card fraud detection models to equal the transaction information to predict the probability of being fraudulent for a new transaction. It offers a scientific basis for the authorization mechanisms.

V. ACKNOWLEDGEMENT

We acknowledge the effort and hard work by the experts who have contributed towards the development of Credit Card Fraud Detection System. We also acknowledge the reviewers of the journal for the suggestions and modifications to improve the quality of the paper.

REFERENCES

[1] Alejandro Correa Bahnsen, AleksandarStojanovic, DjamilaAouada and Bj¨ornOttersten" Cost Sensitive Credit Card Fraud Detection using Bayes Minimum Risk".
[2] Tatsuya Minegishi, AyahikoNiimi "Detection of Fraud Use of Credit Card by Extended VFDT".
[3] V.Dheepa,Dr. R.Dhanapal"Analysis of Credit Card Fraud Detection Methods". International Journal of RecentTrends in Engineering, Vol 2, No. 3, November 2009.
[4] Krishna Kumar Tripathi, Mahesh A. Pavaskar" Survey on Credit Card Fraud Detection Methods". International Journal of Emerging Technology and Advanced Engineering,Volume 2, Issue 11, November 2012.
[5] Y. Sahin and E. Duman" Detecting Credit Card Fraud by Decision Treesand Support Vector Machines".
[6] AlkaHerenj,Susmita Mishra" Secure Mechanism for Credit Card TransactionFraud Detection System". International Journal of Advanced Research in Computer and Communication EngineeringVol. 2, Issue 2, February 2013
[7]AnandBahety,Department of Computer Science" Extension and Evaluation of ID3 – Decision Tree Algorithm".
[8] KaiqiZou, Wenming Sun" ID3 decision tree in fraud detection application".
[9] SozanAbdulla"New Visual Cryptography Algorithm ForColored Image"Volume 2, Issue 4, April 2010.
[10]SnehalPatil,HarshadaSomavanshi,JyotiGaikwad,AmrutaDeshmane,Dept of Computer Engg JSPM'S BSIOTR Pune, India "Credit Card Fraud Detection Using Induction Algorithm".

