Expert Systems with Applications 37 (2010) 6070–6076

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A hybrid model for plastic card fraud detection systems

M. Krivko

Department of Mathematics, University of Leicester, Leicester, LE1 7RH, UK

ARTICLE INFO

Keywords:
Fraud detection
Hybrid model
Plastic card fraud
One-class classification

ABSTRACT

In this paper we present the framework for a hybrid model for plastic card fraud detection systems. The proposed data-customised approach combines elements of supervised and unsupervised methodologies aiming to compensate for the individual deficiencies of the methods. We demonstrate the ability of the hybrid model to identify fraudulent activity on the real debit card transaction data. We also explore the model's efficiency against that of the existing monitoring system of the collaborating bank, using appropriate performance assessment criteria.

1. Introduction

Plastic cards have successfully become an essential part of the modern payment system, providing a broad range of services to the users of the system. Despite being one of the most advanced forms of payment, modern plastic cards still suffer from the same fraud related problems that cash does, namely being counterfeited and stolen. In the present context, we consider plastic card fraud as an unauthorized account activity committed by means of the debit/credit facilities of a legitimate account. In this paper we consider the problem of detecting potentially fraudulent activity on a debit card account and describe the results of our collaborated work with a bank in tackling this issue.

Plastic card fraud is growing along with an increasing volume of payment traffic, advancement and expansion of modern technology, and sophistication of fraudulent tactics. This causes significant losses and great inconvenience to issuing companies, merchants and customers world-wide. In 2007, total card fraud losses on UK issued cards increased by 25% from the previous year and amounted to £535 million (APACS, 2008). The range of fraud tactics observed in the industry can be broadly described within the following categories: lost and stolen card fraud, counterfeit card fraud, card not present fraud, mail non-receipt card fraud, account takeover fraud and application fraud. This list evolves over time as fraudsters adapt new strategies in response to practices of issuing companies and merchants to protect against identified tactics in the future. Currently the largest type of plastic card fraud in the UK is Card-not-Present (CNP) fraud, where the physical card is not present at the point-of-sale (POS). This includes fraud conducted over the Internet, by telephone, fax and mail order and amounts to 54% of all fraud on UK cards. It is expected that the volume of CNP fraud will continue to grow as face-to-face fraudulent transactions become increasingly difficult.

The nature of transaction data and some particular operational issues present a number of challenges for designing a fraud detection system:

- The volume of transactions processed by plastic card issuers daily is high, furthermore each transaction includes more than 70 fields of coded information. Transaction data is heterogeneous and time-varying within and between accounts. Patterns and trends vary significantly for different groups of merchants, holiday seasons and geographical regions.
- The generally accepted fraud rate within the plastic card industry is 0.1–0.2%, i.e. the occurrence of fraud is relatively rare. Frequently this leads to the problem that the majority of cases flagged by the fraud detection system as being potentially fraudulent are in fact legitimate. This type of error is referred to as false positive (FP). As the number of FPs increase so do the associated costs and customer inconvenience.
- Alerts arising from the fraud detection system are usually passed on to the fraud department for further investigation. The suspected cases are followed up with a call to a cardholder for verification of the transactions, where it is required by the bank policy. As a result of this, the number of alerts should be kept at a level such that it can be handled by the available number of investigators and fraud analysts.
- Fraudulent cases missed by the fraud detection system are reported to the issuing company when the cardholder identifies that their account has been compromised. This can take up to several months, resulting in a delay in correctly labelling each case. Some fraudulent cases remain unidentified and therefore mislabelled. Thus, a fraud detection model is almost certainly trained on noisy data.

In order to discourage fraud and to decrease the losses suffered due to fraud, the industry and their member banks employ various technologies to detect and prevent plastic card fraud. Some of the

E-mail address: mk211@le.ac.uk

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.02.119

preventive measures on the cards are consistency checks based on chip and pin, 3-D Secure for online transactions, card reader security and security questions for internet banking, etc. Fraud detection comes into play once prevention has failed and aims to stop the abuse in progress as quickly as possible after it's first occurrence.

Fraud detection systems can be based on various approaches (Bolton & Hand, 2002; Fawcett & Provost, 2002; Phua, Lee, Smith, & Gayler, 2005). The emphasis on fraud detection methodology is usually put upon supervised classification at transaction level that constructs an assignment procedure for new cases from the given training samples of fraudulent and non-fraudulent transactions (Maes et al., 2002; Brause, Langsdorf, & Hepp, 1999). An example of such a system in the banking industry is a rule-based system that consist of rules of the form: If {*assertions*}, Then {*consequence*}. Typically, the in-use set of rules combines the results of a non-statistical expert analysis by a fraud team, findings of investigators, and rules derived from a tree-based algorithm. The strategy is to monitor individual transactions and combinations of the short-term history of transactions. This approach is proven to reliably detect patterns of fraudulent activity which have previously been observed. To extract a rule with confidence there should be an adequate number of cases perpetrated in the same fashion. Time is required to collect the cases, extract an appropriate rule and put it into operation. By the time this circle is complete fraudsters may have changed their tactics. Fraud is an organized criminal enterprise which evolves over time; furthermore there are over 20 main plastic card issuers in the UK and there is no centralized system that collects all identified fraudulent cases. This dissemination of information may prevent a clear and accurate understanding of the incidence of plastic card fraud.

In contrast to the supervised approach, fraud detection systems based on an unsupervised methodology monitor account activity and flag transactions inconsistent with an account's usual behaviour observed over a period of time (Bolton & Hand, 2001; Juszczak, Adams, Hand, Whitrow, & Weston, 2008). Some banks deploy the unsupervised methodology in the form of so-called "behavioural models" which build an individual profile for each account. This includes characteristics of account typical transaction activity, such as merchant types, time of day, monetary values, geographic locations, etc. With a vast number of variations of behaviour and even larger number of opportunities of adapting new patterns it is difficult to cover all possible scenario of legitimate transaction activity. Very often an unusual transaction is in fact legitimate. For instance, purchase of airline tickets or transfer of a lump sum to credit card can be an event which has not previously occurred or has been observed with different parameters. As a result, alerts created by the system in many situations are for incorrectly implicated legitimate cases.

In this paper, we approach the fraud detection problem with a hybrid model that incorporates one-class classification and rule-based approaches at account level. This model has arisen gradually over the implementation stage of our collaborated work with the bank and tackles the issues which supervised and unsupervised approaches may lack individually, by their combination. Given that the use of "behavioural models" might be accompanied by a high number of incorrectly alerted cases whereas the use of rule-based system might result in a poor performance, for instance, when fraud tactics had changed, the logical extension is to apply a combined methodology.

In the data pre-processing step we adapt methodology proposed in (Whitrow, Hand, Juszczak, Weston, & Adams, 2008) for transaction aggregation over a period of time. Since the transaction aggregation yields a smoothed data representation, it is expected to result in a more consistent and stable model than a system built at transaction level. This framework along with a data-customised methodology is deployed in order to build a model of aggregated spending behaviour of an account in a time window. The data-customised methodology is based on separating accounts into several behavioural groups. A model is fitted to each group of accounts, rather than handling each account individually. This results in a reduction of the number of parameters while not adversely affecting the matching of account behaviour to models.

The proposed fraud detection system operates on two levels. At the first level the system monitors any deviation from the account model of aggregated spending behaviour in the time window and assigns a score according to the level of suspicion of fraud. The aggregated sequence of transactions scored above a prescribed threshold is passed on for further refinement to the second level of hybrid model – rule-based filters. A case that contravenes any of the rules is flagged as suspected to be fraudulent. The rules extracted from the transaction records are aimed to enhance the output of the system.

The paper is organized as follows. Section 2 presents the framework for aggregating transaction information at the data pre-processing step. The hybrid model is described in Section 3. Criteria for performance assessment of fraud detection systems are discussed in Section 4. Section 5 contains results of experiments with the real debit card transaction data and performance comparison of the implemented hybrid system with the rule-based fraud detection system of the collaborating bank. Concluding remarks are given in Section 6.

2. Debit card transaction data pre-processing

Let us assume that each transaction $x_i(t)$ of an account i at time t is an object described by a d-dimensional vector of features containing a set of real-valued measurements and categorical indicators such as account number (integer), transaction amount $m_i(t)$ (real non-negative number), transaction type (categorical indicator), etc. Transaction type is an important indicator that defines whether the transaction was conducted at an automatic teller machine (ATM) or at a point-of-sale (POS) terminal. The latter category is subdivided into the point-of-sale type when card is present (POS (CP)) and the point-of-sale type when card is not present (POS (CNP)), i.e. when the transaction is made through Internet, mail or telephone orders.

Consider a time period $[t_1, T]$. For an account i, suppose that there have been n_i number of transactions over the time period $[t_1, T]$ and denote the time-ordered series of transactions as $x_i(t_1), x_i(t_2), \ldots, x_i(t_{n_i})$ such that jth transaction occurred at time t_j. Then we introduce the time-ordered series of transactions over a time window, Δt, of k-day length as

$$X_i(t) = \{x_i(t_j),\ \textit{where } t_j : t - \Delta t \leqslant t_j \leqslant t\} \textit{ for any } t \in [t_1 + \Delta t, T].$$

The transformation from the transaction level to the account level requires an account level summary of the transaction data, i.e. $Y_i(t) = \Phi(X_i(t))$, where Φ is a pre-processing transformation.

To this end, we introduce the notation. Consider the set of transaction types $\mathcal{S} = \{POS(CNP), ATM, POS(CP)\}$. For a transaction $x_i(t_j)$, introduce a three-dimensional column vector $z_{ij} = \left(z_{ij}^1, z_{ij}^2, z_{ij}^3\right)^T$ such that

$$z_{ij}^1 = \begin{cases} 1, & \text{if the transaction } x_i(t) \text{ is of type POS(CNP)}, \\ 0, & \text{otherwise}, \end{cases}$$

and z_{ij}^2 and z_{ij}^3 are defined analogously for the types *ATM* and *POS*(*CP*), respectively.

We choose the pre-processing transformation to be the total value and count of particular type of transactions in the time window Δt. Then the account summary of transaction data over the time window Δt of k-day width is,