Expert Systems With Applications 91 (2018) 235–251

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Sequential fraud detection for prepaid cards using hidden Markov model divergence

William N. Robinson*, Andrea Aria

Computer Information Systems, Georgia State Unviserisity, Atlanta, GA, USA

ARTICLE INFO

Article history:
Received 14 March 2017
Revised 23 August 2017
Accepted 24 August 2017
Available online 12 September 2017

Keywords:
Stored value cards
Transaction processing
Fraud detection
Hidden Markov model
KL divergence
Security

ABSTRACT

Stored-value cards, or prepaid cards, are increasingly popular. Like credit cards, their use is vulnerable to fraud, costing merchants and card processors millions of dollars. Prior techniques to automate fraud detection rely on a priori rules or specialized learned models associated with the customer. Mostly, these techniques do not consider fraud sequences or changing behavior, which can lead to false alarms. This study demonstrates how a transaction model can be dynamically created and updated, and fraud can be automatically detected for prepaid cards. A card processing company creates models of the store terminals rather than the customers, in part, because of the anonymous nature of prepaid cards. The technique automatically creates, updates, and compares hidden Markov models (HMM) of merchant terminals. We present fraud detection and experiments on real transactional data, showing the efficiency and effectiveness of the approach. In the fraud test cases, derived from known fraud cases, the technique has a good F-score. The technique can detect fraud in real-time for merchants, as card transactions are processed by a modern transaction processing system.

1. Introduction

This work presents a store-centric approach to fraud detection. Sequential anomalies are detected using hidden Markov model analysis over a merchant's stream of financial card transactions. This approach detects fraud that would not be found using the more common single-card transaction analysis.

1.1. Cash-card transactions

This research began with a real-world problem. A prepaid-card transaction processor was increasingly experiencing fraud that was not detected by their transaction monitoring system (TMS). A prepaid card is also known as a stored-value card or cash-card. It is a branded product, like Starbucks, AT&T, or Visa, that has the equivalent of cash stored on the card.

It important to observe that such prepaid cards are not associated with a person, commonly. Moreover, cards are rarely reloaded with money. Thus, the lifetime of a card is relatively short—from months to a year, for example. Therefore, there is little information from which to create a card model.

The company experiencing fraud is the international card processor, Card Communications International (CardCom[1]). CardCom has a distribution network of more than 75,000 retail locations. It was the first national point-of-service-activation (POSA) and distribution partner for all major wireless telephone carriers in the USA. CardCom provides many services and products.

CardCom's TMS was typical system in that transaction attributes were checked against threshold values—values outside of the specified ranges are considered potential fraud. Consider this illustrative rule: *if a terminal sold more than $3000 of a specific product* (e.g., *a $25 AT&T card*) *then fraud may have occurred* (and thus, any subsequent transactions from the terminal shall be voided). Initially, this rule-based approach worked well. However, as the ruleset grew, the effort of maintaining the TMS grew too, and consequently the accuracy to detect fraud fell.

The company's rule-based TMS was based on detailed analysis of the transaction histories of each merchant, store, terminal, and product, as well as season (e.g., holiday) and sales specials. The result is hundreds of rules for the different combinations. Moreover, the thresholds used in the rules are derived from transaction histories, and thus require constant update to be current. Specifying, updating, and monitoring the rules themselves is a complex and

* Corresponding author.
E-mail addresses: wrobinson@gsu.edu (W.N. Robinson), aaria@gsu.edu (A. Aria).

[1] CardCom is a pseudonym for the real, international card processor based in Georgia, USA.

http://dx.doi.org/10.1016/j.eswa.2017.08.043
0957-4174/© 2017 Elsevier Ltd. All rights reserved.

burdensome chore. Thus, while the rule-based attribute threshold technique is simple in theory, its use in practice is overly complex.

CardCom increasingly faces a new kind of fraud, where individual transaction attributes are within normal ranges, but a sequence of many small transactions leads to big fraud. Consider a fraudster who takes a bundle of 50 cards and activates them in sequence. Such real cases have occurred, for example when an employee is engaged in fraud. Each activation fails to trigger a fraud detection rule. Nevertheless, the sequence of 50 cards is an anomaly, and should be considered as potential fraud. Sequence analysis is needed to find such fraud. CardCom needs to improve fraud detection, using an approach that requires little maintenance and address fraud sequences.

1.2. Approaches to transaction fraud

In the prepaid card context, the fraud detection system (FDsing a known level of detection, albeit faulty, the new TMS component should ideally work alongside the existing TMS.

A review of the literature reveals two general kinds of fraud detection techniques for transactions: supervised and unsupervised methods. In supervised methods, samples of both fraudulent and non-fraudulent records are used to construct models, which classify observations as either fraudulent or non-fraudulent (Bahnsen, Aouada, Stojanovic, & Ottersten, 2016). Unsupervised methods require little or no prior classifications to identify anomalies, which are subject to subsequent review for the determination of fraud. Rule-based models illustrate the supervised approach. The rules can be automatically derived from training data or directly specified by experts, often to counteract a recent unidentified fraud case (Sánchez, Vila, Cerda, & Serrano, 2009). Other supervised approaches include neural networks, Bayesian models, genetic algorithms, etc., which use machine learning to derive a model of common transaction entities. Unsupervised methods include monitoring unusual transaction characteristics, such as unusually expensive purchases or unusual locations or merchants. A single credit card typically has a near linear slope of cumulative credit spending. Thus, a significant rise above mean slope triggers the threshold for fraud consideration (Hand & Blunt, 2001).

These prior approaches do not meet the four prepaid card criteria. As noted previously, the rule-based method is overly complex for the many prepaid card contexts. It may be improved through automated rule generation; however, this requires a transaction history. Transaction history trained models can be adapted to the prepaid card context, but must address two issues. First, note that card history is central to most of these techniques. Unfortunately, prepaid cards have little history, as they are often discarded after their value is used. Additionally, fraud often occurs near the time of purchase, when the card has no consumer transactions. Alternatively, this approach may be adapted to model the terminal, store, and merchant histories, although we are not aware of any examples. We call this the *store-centric approach*. A second issue is concept drift, which occurs when the real behavior moves away from the modeled behavior. For example, past transactions can be classified as high-, medium-, or low-cost based on the user's transaction history (Srivastava, Kundu, Sural, & Majumdar, 2008). As the user's wealth or inflation increases over time, fewer transactions will be labeled low and medium, while more will be labeled high, and thus the labels are less informative. This drift can be addressed by generating a new set of classified labels whenever the distribution of current labels fits the data poorly. Some approaches consider such concept drift.

Sequential fraud is yet another concern. Many sequence-based analyses consider a valid prefix sequence and then identify a single new transaction as fraudulent. In contrast, a sequence of new transactions may be anomalous, while any single transaction within is valid. Rules can specify sequence characteristics, such as *if a store sells more than 50 of a product in sequence, then fraud may have occurred*. The variety of sequential frauds makes such rules numerous and therefore complex to maintain. Moreover, such rules depend on data windows that are preprocessed to present the necessary metrics. Fraud models may be trained, but again they require sequence analysis, such as provided by Markov models.

1.3. Finding fraud with little history

The store-centric approach introduced in this article uses a little transaction history to detect fraud. It satisfies the four prepaid card context issues: (1) automated model maintenance, (2) little transaction history, (3) automated customization to context, and (4) high detection accuracy for sequentially fraudulent transactions. It fills this prepaid card fraud-detection gap. The approach relies on sequence analysis provided by hidden Markov models (HMM), which are automatically created and compared around the most recent store transactions. HMM software is commonplace; for example, there are open-source versions for Java and R. The idea is to use an HMM modeler for the data. Then, the HMM divergences will raise alerts whenever a threshold is reached. The HMM divergence method provides a baseline of sequence analysis that can easily be added to existing fraud detection systems. The approach is demonstrated using actual prepaid card data and associated real fraud cases. Experiments using the dataset explore method parameters, including number of fraud incidents, fraud threshold, and window size. The findings indicate that the method works well under a variety of conditions. However, it is not intended to replace all fraud techniques. Instead, it provides an effective baseline for the construction of a more comprehensive TMS.

1.4. Article overview

This article continues with a summary of related fraud detection systems. Next, we introduce a well-known approach, HMM's for card histories, for comparison with our store-centric, windowed-history approach. The following sections present our hypotheses, data collection, experiments, and findings. The article ends with a discussion of the approach and conclusions.

2. Background on finding transaction fraud

2.1. Related work on fraud detection systems

A fraud detection system (FDS) reviews behaviors involving a financial instrument, including its transactions, to identify unusual behavior and classify it as fraudulent (Phua, Smith-Miles, Lee, & Gayler, 2007). There are a number of literature surveys on these systems (Bolton & Hand, 2002; Kou, Lu, Sirwongwattana, & Huang, 2004). Here, we highlight a few approaches to a FDS.

Dempster–Shafer theory is use to combined information sources to classify transactions (Singh, Shukla, Rakesh, & Tyagi, 2011). For example, a cardholder activity profile characterizes transaction and shopping behavior. A rule-based component measures the degree of fraud in a transaction. Then, the Dempster–Shafer theory is applied to combine several such information sources to derive an overall belief (Panigrahi, Kundu, Sural, & Majumdar, 2009). Finally, a Bayesian learner can be applied to weaken or strengthen this belief using labeled genuine and fraudulent transactions. Similarly, Krivko (2010) combines expert rules with unsupervised individual profile models to build a hybrid detection model. The general idea is to improve detection by combining models, with the preceding combining unsupervised with supervised models, while Louzada and Ara (2012) combines multiple supervised classifiers.