Engineering Applications of Artificial Intelligence 76 (2018) 130–157

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Survey paper

How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark

Nick F. Ryman-Tubb [a,*], Paul Krause [b], Wolfgang Garn [c]

[a] Room 28MS02, The Rik Medlik Building, University of Surrey, Stag Hill, Guildford GU2 7XH, UK
[b] Department of Computer Science, University of Surrey, Guildford, UK
[c] Business Analytics Group, Department of Business Transformation, University of Surrey, Guildford, UK

ARTICLE INFO

Keywords:
Fraud detection
Financial crime
AI
Machine learning
Payments card
Cyber-crime
Translational research

ABSTRACT

The core goal of this paper is to identify guidance on how the research community can better transition their research into payment card fraud detection towards a transformation away from the current unacceptable levels of payment card fraud. Payment card fraud is a serious and long-term threat to society (Ryman-Tubb and d'Avila Garcez, 2010) with an economic impact forecast to be $416bn in 2017 (see Appendix A).[1] The proceeds of this fraud are known to finance terrorism, arms and drug crime. Until recently the patterns of fraud (*fraud vectors*) have slowly evolved and the criminals *modus operandi* (MO) has remained unsophisticated. Disruptive technologies such as smartphones, mobile payments, cloud computing and contactless payments have emerged almost simultaneously with large-scale data breaches. This has led to a growth in new fraud vectors, so that the existing methods for detection are becoming less effective. This in turn makes further research in this domain important. In this context, a timely survey of published methods for payment card fraud detection is presented with the focus on methods that use AI and machine learning. The purpose of the survey is to consistently benchmark payment card fraud detection methods for industry using transactional volumes in 2017. This benchmark will show that only eight methods have a practical performance to be deployed in industry despite the body of research. The key challenges in the application of artificial intelligence and machine learning to fraud detection are discerned. Future directions are discussed and it is suggested that a cognitive computing approach is a promising research direction while encouraging industry data philanthropy.

1. Introduction

For the first time, fraud detection works are all consistently benchmarked and ranked contemporaneously using industry volumes from 2017. This industry benchmark and survey indicates that despite the academic validity of the research surveyed, its impact on the payment card industry has been minimal. Additional evaluation metrics to explicate the business impact of each fraud detection approach are identified. These show that whilst a fraud detection algorithm may perform well in terms of standard academic measures of accuracy, they can fail to address the broader business context. It is argued that it is important to broaden the evaluation criteria in this way in order to transition this programme of research into a level of technical readiness that is required for impact and to attract the interest of industry (Campolo et al., 2017). This need to meet the challenges of industry is increasingly being recognised globally. For example, the UK Government Industrial Strategy White Paper, specifically highlights the need for funding to "*help service industries to identify how the application of these technologies can transform their operations*" (UK-Government, 2017).

Cashless payments can be made to purchase services/goods using a payment card without the need for physical banknotes. Payment card fraud is the criminal act of deception using a physical (plastic) card or Card-Holder Data (CHD) without the knowledge of the genuine cardholder (Ryman-Tubb and Krause, 2011). CHD is vulnerable to being compromised by criminals who use it to undertake fraud so as to be monetised. A fraud vector consists of a specific sequence of operations to undertake payment card fraud that have been subsequently recognised or detected by law enforcement or fraud experts and reported. There are a wide range of fraud vectors discussed in detail in Shen et al. (2007).

* Corresponding author.
E-mail address: n.ryman-tubb@surrey.ac.uk (N.F. Ryman-Tubb).

[1] A prefix of $ indicates the USA Dollar (USD) value for that variable. 1m = One million (1x10^6), 1bn = One billion (1x10^9) and 1tn = One trillion (1x10^{12}). Appendix A details terms, abbreviations, sources and computation of industry data used. Plotted points and values may contain errors due to the uncertainties in industry figures; error-bars are omitted. Where tables are sorted this is indicated.

https://doi.org/10.1016/j.engappai.2018.07.008
Received 17 November 2017; Received in revised form 28 June 2018; Accepted 24 July 2018

Since the launch of general payments cards in 1950s, fraud vectors have become established over time and became well-known to the industry. Until recently, criminal methods have changed only slowly (Mann, 2006b) which may partly explain the lack of research impetus. Until the 1970s every transaction was processed using paper documents that were physically posted (Evans and Schmalensee, 2005). With the development of the magnetic stripe to store CHD that could be automatically read by terminals, the process could be automated (Svigals, 2012). It was at this point that early research started to focus on the simple automation of detecting fraud and to devise new methods using rules (Parker, 1976). It was not until 1994 that the earliest significant work (Ghosh and Reilly, 1994) was published in this domain.

It will be demonstrated in Sections 2 and 3 that from the earliest work, only a small improvement has been made by the research community, bringing limited impact on the reduction of payment card fraud detection. It is discussed in Section 4 that some of this earliest work is ranked in the top quartile of all works. It is then identified in Section 5 that there is a gap in research into improved systemic methods to manage fraud and future directions are suggested. The following sections outline the contact of payment card fraud, the research challenges. Industry metrics are proposed so that the effectiveness of each method is determined and can then be usefully ranked in a benchmark. Thus, the "state of the art" in fraud detection methods is established.

1.1. The growth of payments and payment card fraud

It is important to review the background of payment card fraud so that the motivation to devise methods to tackle the problem can be understood in context. It is argued here that the economic health, day-to-day government social and cultural existence of citizen's is threatened by the continued growth in payment card fraud and yet research has made slow progress in terms of impact. Society is now a cyber-society dependent on the continued availability, accuracy and confidentiality of information stored, processed and communicated by computers. Businesses and citizens all benefit from this infrastructure and the rapid advancement of cyber-technology including the ability to make rapid secure payments. If fraud reaches a point where security or an economy is sufficiently threatened, trust in these systems will be damaged and their use endangered.

Unfortunately, general society perceive payment card fraud as a minor crime where its effects are mitigated by their issuer refunding any personal fraud; the individual impact to the victim of fraud is softened. There is a common belief that (1) payment fraud only affects banks, big business and government and (2) that the fraud is undertaken by individuals and typically by "bedroom hackers" (Castle, 2008). However, it has been identified that criminal enterprises and Organised Crime Groups (OCGs) use payment card fraud to fund their activities including arms, drugs and terrorism (Financial-Fraud-Action-UK, 2014). The activities of these criminals include violence and murder (Everett, 2003; Jacobson, 2010)—individual acts of fraud have a human cost. In 2017, it is forecast that there will be 349 bn payment card transactions with Card Expenditure Volume ($\$CEV$) at \$26.3 tn with direct fraud losses ($\$fraud$) at \$24 bn; it is here calculated that the economic impact is a minimum of \$416 bn (Appendix A). Fig. 1 shows the exponential growth of $\$CEV$ and $\$fraud$. In 2017, it is forecast that for the first time $\$fraud$ will grow more rapidly than $\$CEV$. As argued in Ryman-Tubb (2011), the same technology that has enabled cashless payments is fuelling exponential growth in payment card fraud.

1.2. Payment card transaction process

There are multiple participants that are involved when a cashless transaction takes place (see Fig. 2). When a merchant wishes to take payment from a cardholder's payment card, then the details of that transaction are passed to the merchant's acquirer. The acquirer then requests authorisation from the cardholder's card issuer and the transaction is approved or declined. This decision is then passed back to the merchant to complete the transaction. If the transaction is *authorised* then the sale is completed and the goods are taken or dispatched.

1.3. Fraud Management System (FMS)

To determine if a payment card transaction is authorised, a number of processes are undertaken, one of which includes the FMS. The FMS receives a transaction, makes a decision using some form of classifier and returns this as part of the authorisation process. If the transaction is determined to be suspicious it is typically blocked or declined and a *fraud ticket* is created. This fraud ticket contains sufficient information for a human *reviewer* to understand the transaction and then make a decision. In most organisations, a team of reviewers check fraud tickets and an investigation is undertaken that might include contacting the cardholder or merchant.

1.4. Major challenges in real-world fraud detection

The timely understanding and detection of fraud vectors is fundamental to reducing the growing payment card fraud problem. The complex scientific and industry challenges of detecting payment card fraud through the use of AI and machine learning have been identified in this survey and each is discussed in the following sections. Specific applications in the near future and research directions are discussed in Section 5.

1.4.1. Transparent decisions

It is argued that an important factor limiting the impact of research is that the majority of published methods are *black-boxes* where their workings are mysterious; the inputs and its decision on fraud can be observed but how one becomes the other is opaque. They cannot easily explain their decisions or reasoning so that humans cannot understand the new emerging fraud vectors. However, industry considers that it is only the timely understanding of new fraud vectors that will allow improved prevention methods to be put in place. For fraud practitioners, it is argued that comprehensible classifiers are essential to guide them towards a particular type of investigation and towards creating prevention that is more effective.

> "*Gaps in knowledge, putative and real, have powerful implications as do the uses that are made of them. Alan Greenspan, once the most powerful central banker in the world, claimed that today's markets are driven by an 'unredeemably opaque' version of Adam Smith's 'invisible hand' and that no one (including regulators) can ever get more than a glimpse at the internal workings of the simplest of modern financial systems*". (Pasquale, 2015).

1.4.2. Cost of fraud detection to the payments industry

If academic research is to have a greater industry impact then it is argued that researchers need to understand that costs are a key motivation within the payments industry. For example, in practice most FMS produce a large volume of $AlertD$ that must be matched against available and costly human review resource and so the issue of prioritisation requires attention. It is argued that only if the various costs are taken into account that a more effective FMS can be created (Hand et al., 2008). The output of a fraud detection system requires human reviewers to investigate alerts generated. There is an operational cost for such a process — with the number of reviewers, experts and the required IT being a significant proportion (typically 30% of the value of fraud write-offs in 2017). An illustration of the size of a review team is given in Appendix A.

The accuracy of a fraud detection model can be set so as to detect all fraud but this will have a resultant uneconomical increase in the operational cost to detect the fraud, as $AlertD$ becomes unrealistic. Therefore, a commercial decision must be made between these costs and the impact and savings by detecting fraud (Bose, 2006). This is further complicated as "disturbing good customers" by contacting them about an alerted transaction that is not fraud does not inspire customer confidence; implying to the innocent customer that there is the suspicion of fraud is likely detrimental to good relations (Leonard, 1993). Few methods take this into account.