Introduction

Here you have to give some known facts about the field you will work on.

Try to focus on the problems that are most common and then state the goals of the project.

* You can try to answer to the following question:
* Which questions do we want to answer ?
* What is known about the problem?
* How do we define the outcome(s)?
* What is known to influence the outcome?
* Does we have any possible new knowledge that has not been in use before?

This part must be between half to one and half page.

# Introduction

Ecommerce fraud, also known as payment fraud, is a criminal deception conducted during a commercial transaction over the Internet with the goal of financial or personal gain of the fraudster while negatively affecting the bottom line of the merchant.

Chargeback is the term used to define the reversal of a credit card payment that comes directly from the bank or card issuer. When customers report a transaction is fraudulent, merchants are “charged back” the amount requested to be charged by them for delivering goods or services, so they are left with no payment on the goods or service provided.

As more business moves from brick and mortar stores to online retailing, ecommerce fraud is on the rise, with a staggering annual growth rate of 30%. In 2017, ecommerce fraud lead to chargeback losses of $57.8 billion across 8 big industries. [[1]](#footnote-1)

In this document, we offer a method to identify such fraudulent transactions when they occur and before they affect the merchant or end user.

The document addresses the following questions:

1. What is a fraudulent transaction (definition)
2. How to identify a fraudulent transaction (Boolean predictor)
3. How to mitigate false positives and false negatives in identifying a fraudulent transaction
4. Methods to improve fraudulent transaction detection accuracy
5. Suggested method uniqueness vs other methods used

What is a fraudulent transaction?

An ecommerce transaction occurs online, where the buyer is not present at a physical store, and does not present a credit card and identity measures to the seller.

On one hand, this provides an opportunity for fraudsters to execute fraudulent transactions, as their real identity is not disclosed, and they do not need to physically be in the seller location. In fact, fraudulent transactions can (and often do) be executed very far from the seller location, including other countries and even other continents.

On the other hand, since these transactions occur online, there are digital footprints associated with the transaction that can help trail it back or correlate it with other transactions or actions with the same characteristics. This include the origin transaction country, time of day of the buyer and seller, ip address used, the type of device used, currency, physical addresses of the alleged buyer, and other footprints that can indicate a fraudulent transaction.

The purpose of this model is to attempt and identify these correlations and anomaly in data and thus calculate the probability of a fraudulent transaction. If the probability crosses a certain threshold, than the transaction will be deemed fraudulent.

## Labeling logic

The logic of our labeling is to define a reported chargeback on the card as a fraudelent transaction (isFraud=1) and transactions posterior to it with either user account, email address or billing address directly linked to these attributes as fraud too. If none of the above is reported and found beyond 120 days, then we define it as a legit transaction (isFraud=0).

## Handling fraudulent transactions mislabeling as legit ones (reduce false negative findings)

In the real world fraudulent activity might not be reported at all, e.g. the cardholder was unaware, or neglected to report it during the acceptable the claim period. In such cases, an actual fraudulent transaction might be labeled as a legit one, and thus missed by the model.

However, the assumption is that these types of fraudulent transactions are not common, and would not have a significant impact on the model, due to the following reasons:

1. Unreported real fraudulent transactions are bound to have minimal impact on the affected buyer and merchant, otherwise these would have been picked up by any one of them and reported as fraudulent. Therefore these are either transactions having very low sums, and/or very rare. In either case these would represent a very small percentage portion of the overall fraudulent transactions number and total amount, and thus would not have an impact on the model, which is aimed at preventing impactful fraudulent transactions.
2. As can be seen from the model, there is a reoccurring pattern of fraudulent transactions that is derived from same or similar entities or behaviors that represent the transaction: sourced from the same countries, using same email or IP addressed, or having similar amounts, currencies, transaction times, etc. This finding, along with the model capability to back propagate and mark transactions from the same sources as fraudulent, much further reduces false negative findings and eventually marks also most non-impactful transactions as fraudulent based on their source.

## Handling legit transactions mislabeling as fraudulent ones (handle false positives)

The model assumes a strict approach that labels all transactions coming from a fraudulent source as fraudulent transactions.

While this approach may label legit transactions as fraudulent ones, it favors such markings, which can then be further examined by other models or human inspectors, rather than allow fraudulent transactions go undetected.

The model marking of a fraudulent source as a high weighted input for marking a fraudulent transaction also further mitigates this risk, as only transactions coming from such sources are exposed to being marked fraudulent.

## KPIs for the suggested method

The suggested method accuracy should be

1. <https://review42.com/ecommerce-fraud-statistics/> [↑](#footnote-ref-1)