Csaba Elekes

Wannabe **Senior Data Analyst Verification Team**

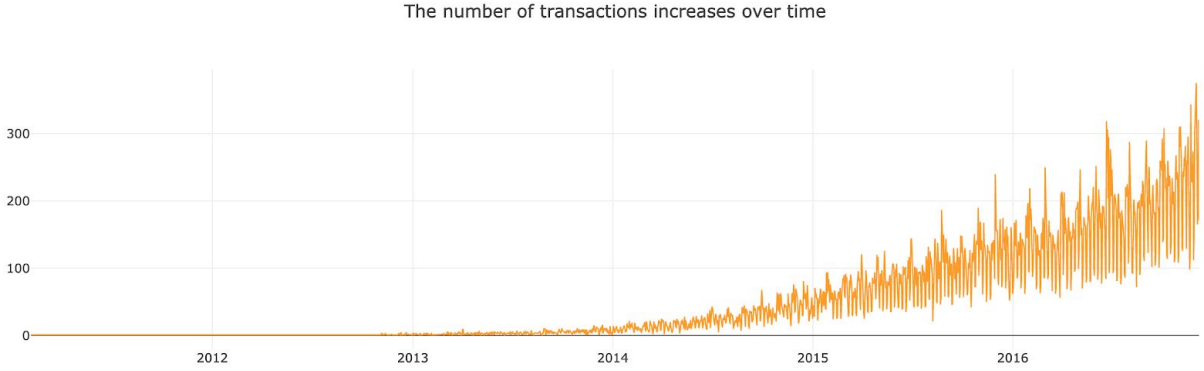
Summary

For the take-home assignment t100 000 records of previous **transactions** were handed out. The goal was to **find** risky customers who might be trying to **launder money** and give suggestions on **how treat them**. I provided several methods to identify suspicious individuals and activities. I divided the risky transactions to two groups - one-off and regular transfers. **One-off transfers** are risky if the customer's identity cannot be verified, sender or receiver is on a blacklist, transferred amount is too high or does not align with the user's regular behaviour. **Regular transfers** might show patterns that indicate fraud, e.g. transfers submitted too frequently, cancelled transactions, too many transfers to different recipients, money moves in circles or chains ending up in one account, etc. Business accounts follow different patterns than personal ones. The attached notebook contains SQL and Python code to find risky customers of each kind.

Introduction

The mission of the Verification Team is to make sure TransferWise only provides services to legitiate individuals and businesses, and flags suspicious individuals and transactions involved in money laundering. The take-home test contained 100 000 records about transactions that happenned between 2011 and 2016. The goal was to find high-risk customers and give recommendations on how to treat them, what kind of info to acquire for the decision.
The dataset contained:

- Details about senders and recipients: hashed ID, country, date of first attempt and success, personal or business, device category
- Details about transfers: dates (submitted, received, completed, cancelled), currencies, reference classification, transfer sequence, days since previous transfer



Limitations

Data is sampled, only a proportion of the records are handed out, information is incomplete

- Impossible to have a complete view on the suspicious users' activity
- Can't tell what was the point when the service was denied
- Can't find loops, chains, accounts where the money ends up

Data is not labelled, hard to tell which users were deemed too risky by TransferWise.

- The best indicators for blocked accounts and transactions are whether they were successful or cancelled. Transfers can be cancelled for many reasons by the user and by TW, and this information is missing

Other hidden / missing info -- columns are less useful when info is sparse

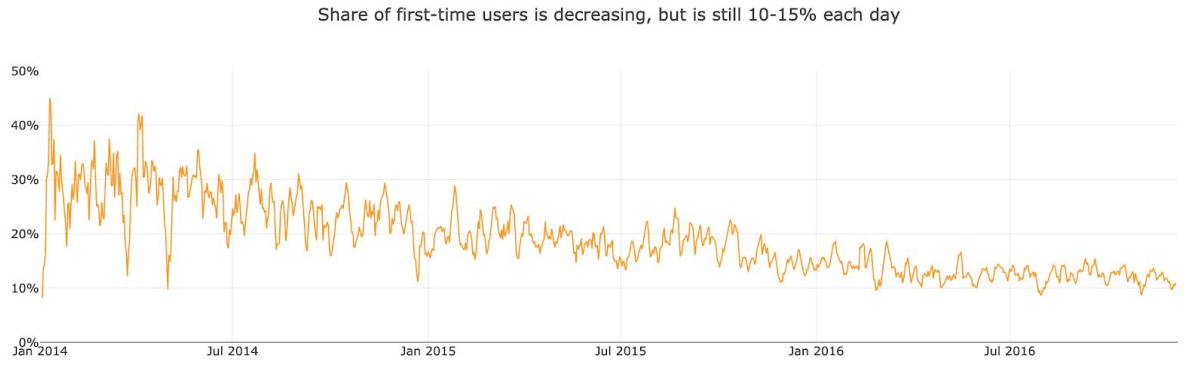
- Reference
- Bank country, bank name

One-off transfers

Incomplete registration

Just like with traditional banks, TransferWise customers have to prove that they are eligible to access the services of the company. During the registration process customers have to provide basic information about their identities like pictures of ID cards, passports or electronic signatures where it is avaliable.
Many users sign up when they actually want to transfer money abroad, and they going through the identification process during their first transfer.

- 10-15% of the daiy transaction comes from new users (`transfer_sequence = 1`)
- 5600 out of 15000 first-time transfers were unsuccessful (`payment_status != 'Transferred'`)
- About half of them solve the issue later (`first_success_date is not null`)
- E.g. `8b70ed5969d08d28ba67807ce2af2284`



Treatment / solution:

- Simple and straightforward registration and identification process
- Detailed and easy-to-find FAQ
- Reminder emails

Blacklists

There are national and international lists containing people and organisations that are involved in criminal activities and money laundering.

- TransferWise should check whether the sender or the receiver of a transfer is on any of these lists.
- Newly registered users should also be verified this way
- Certain countries are problematic (e.g. Iran, Iraq, Afghanistan, Cuba)
    - Data set did not contain any transfers to and from these countries
    - I assume TransferWise's service is not available there

I tried to find suspicious transfers that might have been blocked because of blacklists. As there is no information in the table about the reason of cancellation, I could only make guesses which info might point to declined customers.

- First-time users (`transfer_sequence = 1`)
- Who have alreay sent the amount to TW's account (`date_request_received is not null`)
- But they have never had any successful transaction (`first_success_date is null`)
- There are 234 such transactions in the dataset
- E.g. `eb7d8621b0e718b9d2d410afc738e9a2`

Amount limits

Below certain amount (and frequency) the transfer doesn't pose a high risk of money laundering → more relaxed rules

Above certain amount (e.g. $1000) users can be asked for extra documents

- The UI + email can inform users that the transfer will be blocked until the documents are provided and checked
- Customer support may try to contact users who haven't taken an action
- Contacting CS before a large transfer might speed up the approval process
- Never succeeded: `1a8ebb56a246a22b8dc164049c703e92`
- Solved the issue later: `6008779d978132785979f29361c18bbe`

Extra inspections for top 1% of each currency. In case of very large amounts TW can run some manual checks (is new recipient? unusual target_ccy? reference?) to make sure the transfer is legit and user is not victim of cyber fraud

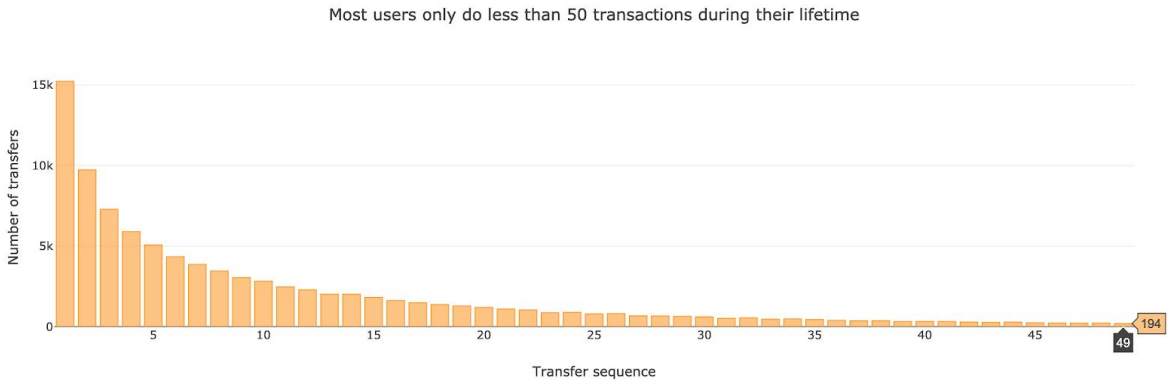
- Based on the handed out dataset 20 000 EUR, USD, GBP looks like a good threshold
- Contact customer, ask for extra proof (e.g. papers about house inherited and sold)
- e.g . `a2cc65be7e9ad89e3761fe3c036531d0`

Regular transfers

It is sometimes easier to spont fraudulent activities if they follow certain patterns. Here are a couple of cases that can be checked in the dataset.

Too frequent

95% of users start less than 50 transactions over their lifetime. After the 50th transfer of a user TW should look into the details to make sure they don't launder money (amounts, destinations, regularity, payment references).



Some users start several attempts within minutes from each other. Most of the transfers never go through and the majority is cancelled before payment made.

Treatment / solution:

- Business customers: ask for invoice, check if the company is legit
- Personal customers: check transfer reference, payment method, recipient details in public registers, government databases

Examples:

- `39ded89a9e1d1a7ee64742c701657716` → a lot of transfers, all of them cancelled before sending the money
- `8917c69f0a728519439d93ba9b49005a` → some kind of test account
- `b78c8ce3f52611c3df21da6b9effe911` → many attempts on the first day, all failed
- `d6fb6f9d8ca89d20d9db824b845159d0` → tons of self-transfers

Unusual attempts

Probably most users use TransferWise according to their habits and lifestyle. Regular users might send money to family, friends, customers or employees, but the recipients are usually concentrated in certain countries, continents and amounts are within certain ranges.
New currencies, countries, unusually high amounts should raise a flag and might indicate cyber fraud or a hacked account. (e.g. `d966e072fab4f783c66d30fa2ed4a723`)

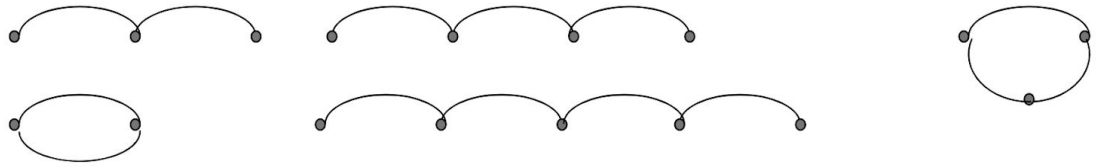
Treatment / solution:

- Email + phone notification
- Delay transfer to give time for customer to notice it.

Chains, circles

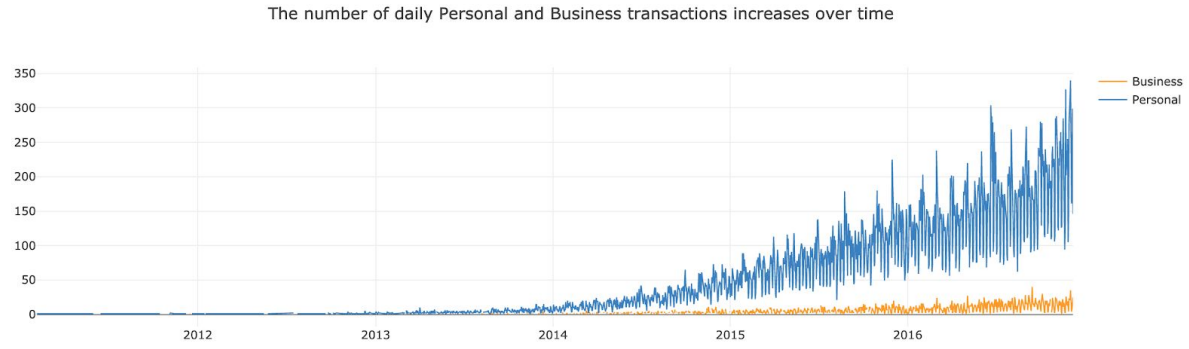
Money laundering and cheating with taxes can take form of organised crime. One common method is that companies invoice each other in a chain or circle, and claim taxes back. The first one doesn't pay taxes, but the rest receives refunds.

Some or all of these transfers can go through TransferWise's systems. The company should cooperate with tax authorities.



Due to the fact that the data is sampled it was unlikely to find these. The longest chain of transfers was of length 3 and looked legit.

Business vs Personal accounts



The number of Business users is much smaller. However, some of them start much more transactions than the average users, and it looks legit.

- `69fd02c4fbd5bfa6533f7a5eac3bd81c`: 9000+ transactions, Business
- `b2afd12d1322929e095bd85468e50a55`: 19000+ transactions from UK to India, all recipients are different, amounts roughly the same -> ask the business about the purpose of txn, Business

And only a handful of Personal accounts start 1000+ transfers

- `fee71f5728308f45cabfd22d4551eca0`: same amount sent multiple times, almost at the same time to the same business → weird, further inspection needed

These are a few cases to consider. Due to the limitations of the available dataset it is hard to assess the risk of a transaction or a customer. Please find the code and the technical part of the analysis on GitHub.