How to mitigate fraud risk in mobile wallet

A long journey in mobile payment fraud and abuse analytics

Seonmin Kim
Seonmin Kim
Data risk analyst
LINE Corporation

Analyze internal and external data sources to identify anomalies ranging from content abuse to payment fraud.
What is data risk analyst?

Q. A million new users joined LINE PAY through promotions. One month later, we noticed that 20% of these users are not using our services anymore. Why?

- **Business analyst approach**: Let's send a survey so that we can figure it out. Based on the result of survey, we will consider the next step.

- **Data Risk analyst approach**: We think they are fake users who created a number of fake accounts to abuse our promotions. We demonstrate it through various analytical skills and ideas.
LINE platform

**Message**

**Phone call**

**News / Coupon**

**Mobile Wallet**
LINE platform

Message

Phone call

News / Coupon

Mobile Wallet
LINE Mobile wallet

1. Payment by credit cards
2. Payment by balance
3. Money transfer
4. Deposit & Withdrawal
Analytics in action

1. Fraud Risk in a single account
   • Chargeback

2. Fraud Risk using multiple accounts
   • Promotion Fraud
   • Layering pattern

3. Risk based scoring model
   • Transaction based risk scoring
Analytics in action

1. Fraud Risk in a single account
   - Chargeback

2. Fraud Risk using multiple accounts
   - Promotion Fraud
   - Layering pattern

3. Risk based scoring model
   - Transaction based risk scoring

1. CNP (CARD-NOT-PRESENT) FRAUD
   - Expiry date, Card number, Verification code..

2. LOST AND STOLEN CARD FRAUD
   - Physical possession of someone’s card
Analytics in a single account

1. CNP(CARD-NOT-PRESENT) FRAUD

Raw features

<table>
<thead>
<tr>
<th>Features associated with LINE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num of removed card</td>
</tr>
<tr>
<td>Num of devices</td>
</tr>
<tr>
<td>Num of added card</td>
</tr>
</tbody>
</table>

A set of derived features

<table>
<thead>
<tr>
<th>Features derived from LINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendship score</td>
</tr>
<tr>
<td>Service Loyalty score</td>
</tr>
</tbody>
</table>

Aggregated features

<table>
<thead>
<tr>
<th>Aggregated features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of transactions in the last t_p hours</td>
</tr>
<tr>
<td>The unique item codes in the last t_p hours</td>
</tr>
<tr>
<td>Total number of unique merchant codes in the last t_p hours</td>
</tr>
</tbody>
</table>

2. LOST AND STOLEN CARD FRAUD

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>

Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>

Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>

Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>

Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>

Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
</tbody>
</table>
Analytics in action

1. Fraud Risk in a single account
   • Chargeback

2. Fraud Risk using multiple accounts
   • Promotion Fraud
   • Layering pattern

3. Risk based scoring model
   • Transaction based risk scoring

1. Creating fake accounts

2. A number of transactions in a short time

3. Repeatedly buying the same item

4. Purchase expensive products with similar prices

5. Repeatedly buying items from the same merchant category
## Analytics in a single account

### 1. CNP(Card-Not-Present) Fraud

- **Raw features**
  - CNP(CARD-NOT-PRESENT) FRAUD
  - LOST AND STOLEN CARD FRAUD

### 2. Features associated with Transaction

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

### 3. Aggregated features

<table>
<thead>
<tr>
<th>Aggregated features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of transactions in the last $t_p$ hours</td>
</tr>
<tr>
<td>The unique item codes in the last $t_p$ hours</td>
</tr>
<tr>
<td>Total number of unique merchant codes in the last $t_p$ hours</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

### 4. Features derived from LINE

<table>
<thead>
<tr>
<th>Features derived from LINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendship score</td>
</tr>
<tr>
<td>Service Loyalty score</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

### 5. Features associated with LINE ID

<table>
<thead>
<tr>
<th>Features associated with LINE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num of removed card</td>
</tr>
<tr>
<td>Num of devices</td>
</tr>
<tr>
<td>Num of added card</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>
## Analytics in a single account

### 1. CNP(CARD-NOT-PRESENT) FRAUD

- **Features associated with LINE ID**
  - Num of removed card
  - Num of devices
  - Num of added card
  - ...

- **Features derived from LINE**
  - Friendship score
  - Service Loyalty score
  - ...

### 2. LOST AND STOLEN CARD FRAUD

- **Features associated with Transaction**
  - Amount
  - Online/Offline
  - Type of card
  - ...

- **Aggregated features**
  - Total Number of transactions in the last $t_p$ hours
  - The unique item codes in the last $t_p$ hours
  - Total number of unique merchant codes in the last $t_p$ hours
  - ...

---

**LINE**
## Analytics in action

<table>
<thead>
<tr>
<th>1. <strong>CNP (CARD-NOT-PRESENT) FRAUD</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Creating fake accounts</td>
</tr>
<tr>
<td>2. A number of transactions in a short time</td>
</tr>
<tr>
<td>3. Repeatedly buying the same item</td>
</tr>
<tr>
<td>4. Purchase expensive products with similar prices</td>
</tr>
<tr>
<td>5. Repeatedly buying items from the same merchant category</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. <strong>LOST AND STOLEN CARD FRAUD</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CNP (CARD-NOT-PRESENT) FRAUD</td>
</tr>
<tr>
<td>2. LOST AND STOLEN CARD FRAUD</td>
</tr>
</tbody>
</table>
### Features associated with LINE ID

- Num of removed card
- Num of devices
- Num of added card
- ...

### Features associated with Transaction

- Amount
- Online/Offline
- Type of card
- ...

### Aggregated features

- Total Number of transactions in the last $t_p$ hours
- Total sum of amount in the last $t_p$ hours
- Total number of unique merchant codes in the last $t_p$ hours
- ...

#### A set of derived features

- Friendship score
- Service Loyalty score
- ...

**What is data risk analyst?**

- Account
- Subscribed date
- Device info
- IP address
- ....
## Analytics in a single account

### 1. CNP(CARD-NOT-PRESENT) FRAUD

<table>
<thead>
<tr>
<th>Features associated with LINE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num of removed card</td>
</tr>
<tr>
<td>Num of devices</td>
</tr>
<tr>
<td>Num of added card</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Features derived from LINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendship score</td>
</tr>
<tr>
<td>Service Loyalty score</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

### 2. LOST AND STOLEN CARD FRAUD

<table>
<thead>
<tr>
<th>Features associated with Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
</tr>
<tr>
<td>Online/Offline</td>
</tr>
<tr>
<td>Type of card</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aggregated features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of transactions in the last $t_p$ hours</td>
</tr>
<tr>
<td>Total sum of amount in the last $t_p$ hours</td>
</tr>
<tr>
<td>Total number of unique merchant codes in the last $t_p$ hours</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>
Analytics in action

1. Fraud Risk in a single account
   - Chargeback issue

2. Fraud Risk using multiple accounts
   - Promotion Fraud
   - Layering pattern
   - Library: Python
   - Visualization: NetworkX
   - Visualization: Gephi

3. Risk based scoring model
   - Transaction based risk scoring
Analytics of multiple accounts

Community detection

A Network graph

Communities

Modularity Score: 0.94

Meaningless
Analytics of multiple accounts

Community detection

A Network graph

Communities

Community 1

Community 2

Community 3
Analytics of multiple accounts

Definition of suspicious groups.

1. **Community shape**
   - A community where money is being transferred to specific users

2. **Similar Transaction Sequence**
   - Every user in a community has the same (or similar) transaction sequence

3. **Fast transaction Interval**
   - Deposit, payment, and money-transfer occur too fast

4. **Presence of low loyalty users**
   - A community consisting of users with low service loyalty score
Analytics of multiple accounts

Promotion fraud

Layering pattern
Analytics of multiple accounts

Promotion (Marketing) fraud
Analytics of multiple accounts

Promotion (Marketing) fraud

- **Abuser**: A group(s) that seeks monetary benefits from the weaknesses of various promotions
- **Broker**: A user(s) who collects money from abuser groups
- **Banker**: A user(s) who withdraws (spends) the money received from the broker(s)
Analytics of multiple accounts

Promotion (Marketing) fraud

- Community shape
  - Graph shape (Density, Betweenness, Assortativity, ..)
- Presence of low loyalty users
  - Service loyalty Score
- Similar transaction sequence
  - Transaction Similarity
- Short interval of transaction sequences
  - Transaction Interval
Analytics of multiple accounts

Combo Meal (Burger + Drink + Potato fries)

- **John**: Burger → Drink → Fries → Burger → Burger → Fries → Drink → Burger → Fries → Drink → Fries
- **Smith**: Fries → Burger → Drink → Fries → Drink → Fries → Fries → Drink → Burger → Burger → Fries
- **Diana**: Drink → Burger → Fries → Burger → Drink → Fries → Burger → Drink → Burger → Burger
- **Friends community**: BDFBBFDBFDF FBDFDFDBBFBFBDBDBB
Analytics of multiple accounts

Transaction sequence in communities

Deposit = D, Payment = P
Money Transferred = T
Money Received= R
Withdraw = W

C- 1 : [Transaction sequence]

C- 2 : [Transaction sequence]

C- 3 : [Transaction sequence]

C- 4 : [Transaction sequence]
Analytics of multiple accounts

Layering pattern

- Size of community
- Amount of money

Deposit = D
Payment = P
Money Transferred = T
Money Received = R
Withdraw = W

No payment action?

- Action sequence: DTRDTRDTRDTRDTRDTRDTRDTRDTR
  R...DTRDTRDTRDTRDTRDTRDTRDTRDTRDTR
  RDTRDTRDTRDTR...DTRDTRDTRDTRDTRDTRDTR
  DTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTR
  DTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTR
  DTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTRDTR

Broker 3
Bankers
Broker 1
Broker 2

No payment action?
Analytics of multiple accounts

- Key designed features of fraud using multiple accounts
  - Graph shape (Radius, Betweenness, Assortativity, ..)
  - Service loyalty Scoring
  - Transaction Similarity
  - Transaction Interval
Analytics of multiple accounts

- Key designed features of fraud using multiple accounts

✓ **Graph shape**

1. Diameter / Radius
2. Betweenness (Std)
3. Betweenness (Avg)
4. Clustering coefficient
5. Assortativity
6. Degree (Std)
7. Degree (Avg)
Analytics of multiple accounts

- Key designed features of fraud using multiple accounts
  - Graph shape (Radius, Betweenness, Assortativity, ..)
  - Service loyalty Scoring
  - Transaction Similarity
  - Transaction Interval

- Identify suspicious communities
- Verify fraud communities
Analytics in action

1. Fraud Risk in a single account
   • Chargeback issue

2. Fraud Risk using multiple accounts
   • Promotion Fraud
   • Layering pattern

3. Risk based scoring model
   • Transaction based risk scoring
Q. Why do we need a risk based whitelist model?

To divide users in different risk levels for risk management purposes.

- Transaction Risk Scoring
- Risk Intelligence
What is the whitelist model

Pattern recognition on mobile payment

• Who we can trust?
  ① Cycle of deposit
  ② Cycle of payment
  ③ Service loyalty score
  ④ …Etc
Thank you
Rate today’s session

Cyberconflict: A new era of war, sabotage, and fear

9:55am-10:10am Wednesday, March 27, 2019
Location: Ballroom
Secondary topic: Security and Privacy

We’re living in a new era of constant sabotage, misinformation, and fear, in which everyone is a target, and you’re often the collateral damage in a growing conflict among states. From crippling infrastructure to sowing discord and doubt, cyber is now the weapon of choice for democracies, dictators, and terrorists.

David Sanger explains how the rise of cyberweapons has transformed geopolitics like nothing since the invention of the atomic bomb. Moving from the White House Situation Room to the dens of Chinese, Russian, North Korean, and Iranian hackers to the boardrooms of Silicon Valley, David reveals a world coming face-to-face with the perils of technological revolution—a conflict that the United States helped start when it began using cyberweapons against Iranian nuclear plants and North Korean missile launches. But now we find ourselves in a conflict we’re uncertain how to control, as our adversaries exploit vulnerabilities in our hyperconnected nation and we struggle to figure out how to deter these complex, short-of-war attacks.

David Sanger
The New York Times

David E. Sanger is the national security correspondent for the New York Times as well as a national security and political contributor for CNN and a frequent guest on CBS This Morning, Face the Nation, and many PBS shows.