How to mitigate fraud risk in mobile wallet

A long journey in mobile payment fraud and abuse analytics

Seonmin Kim

Speaker info

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*** **:**

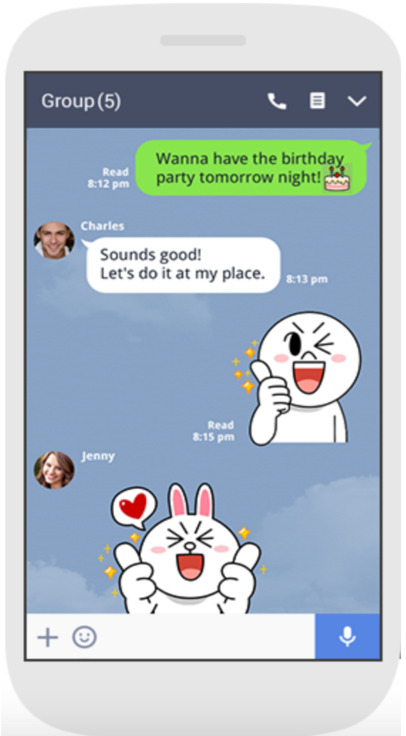
  We think they are fake users who created a number of fake accounts to abuse our promotions.

.	We demonstrate it through various analytical skills and ideas.

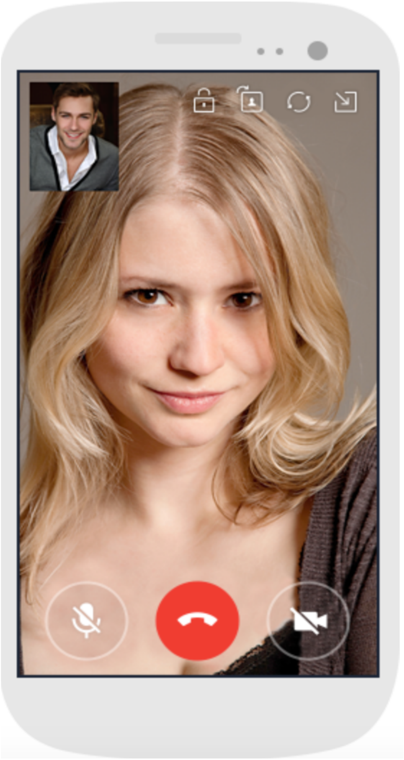
LINE

LINE platform

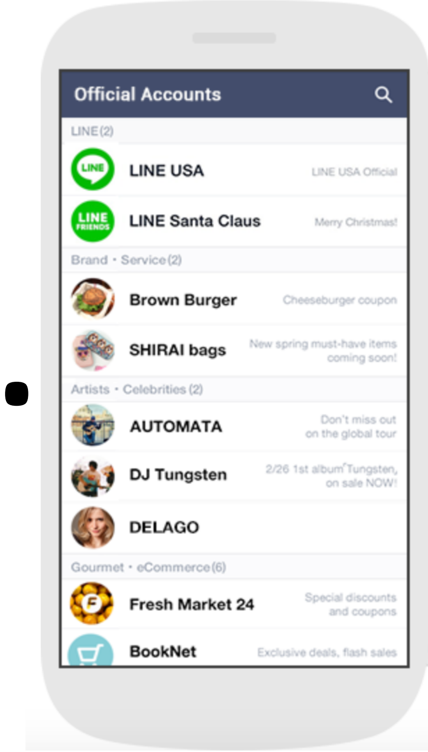
Message



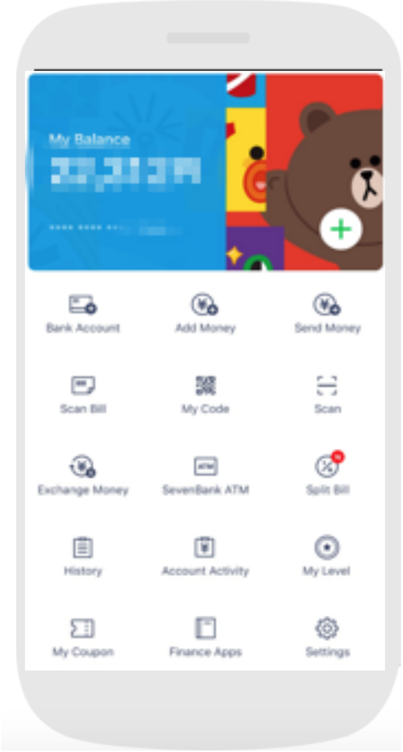
Phone call



News / Coupon

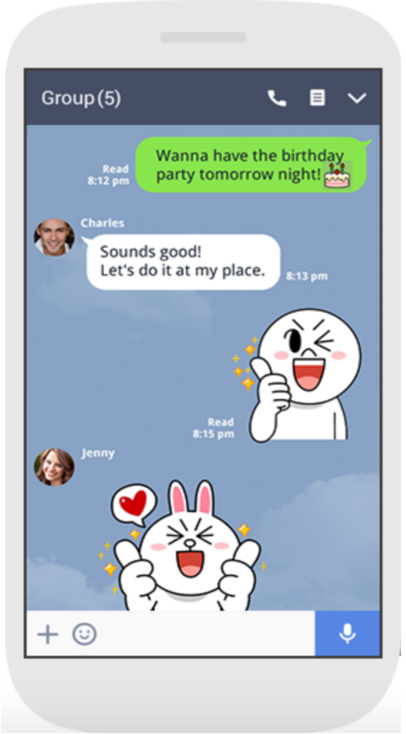


Mobile Wallet

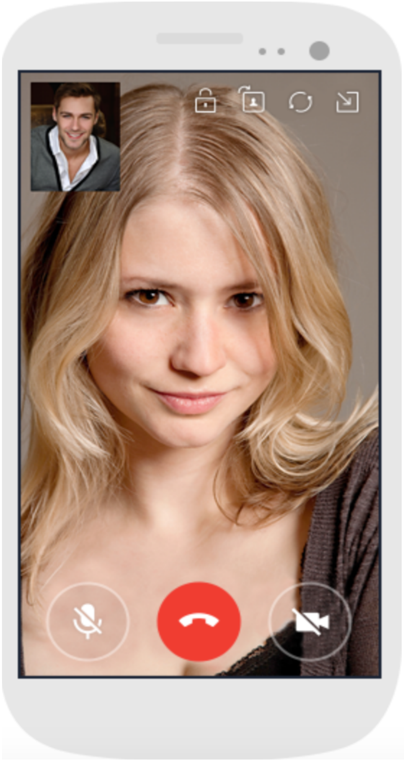


LINE platform

Message

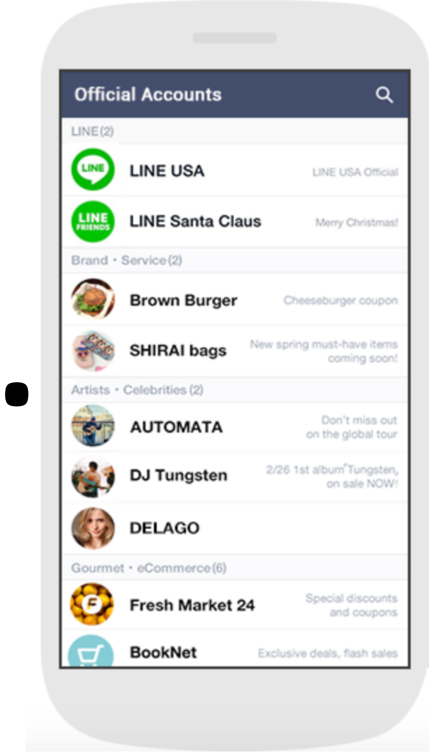


Phone call

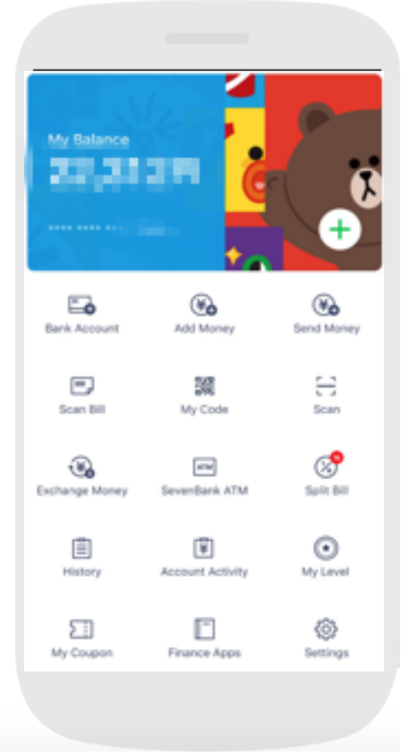


....

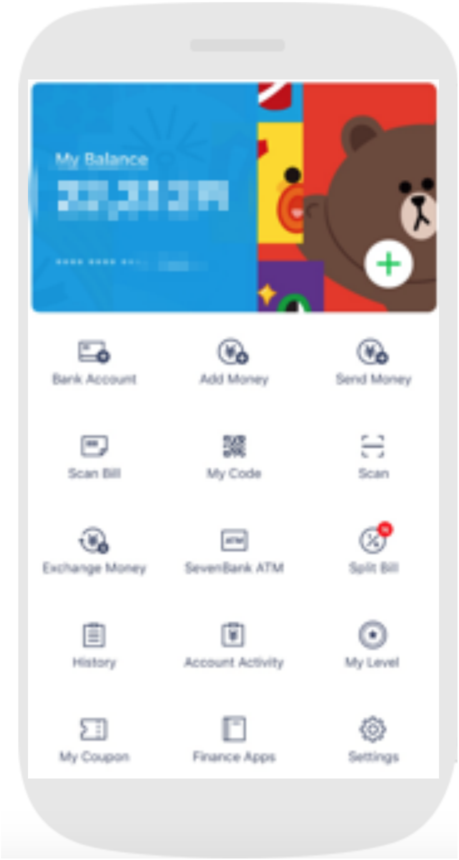
News / Coupon



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

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
…

2. LOST AND STOLEN CARD FRAUD

A set of derived features

Features derived from LINE
Friendship score
Service Loyalty score
…

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
…

Aggregated features

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

LINE

Analytics in a single account

1. CNP(CARD-NOT-PRESENT) FRAUD

2. LOST AND STOLEN CARD FRAUD

Raw features

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
…

☆ ☆ *Features derived from LINE*
Friendship score
Service Loyalty score
…

☆ *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
…

Analytics in a single account

1. CNP(CARD-NOT-PRESENT) FRAUD

2. LOST AND STOLEN CARD FRAUD

Features associated with LINE ID	Features associated with Transaction
Num of removed card	Amount
Num of devices	Online/Offline
Num of added card	Type of card
…	…

☆ ☆ Features derived from LINE	☆ Aggregated features
Friendship score	Total Number of transactions in the last t_p hours
Service Loyalty score	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

1. CNP(CARD-NOT-PRESENT) FRAUD

2. LOST AND STOLEN CARD FRAUD

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~~

LINE

What is data risk analyst?

- **Account**
- **Subscribed date**
- **Device info**
- **IP address**
- **….**

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
…

☆ ☆ ***Features derived from LINE***
Friendship score
Service Loyalty score
…

A set of derived features

☆ ***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
…

Analytics in a single account

1. CNP(CARD-NOT-PRESENT) FRAUD

2. LOST AND STOLEN CARD FRAUD

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
…

☆☆ *Features derived from LINE*
Friendship score
Service Loyalty score
…

☆ *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
…

LINE

Analytics in action

1. **Fraud Risk in a single account**
   - Chargeback issue

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

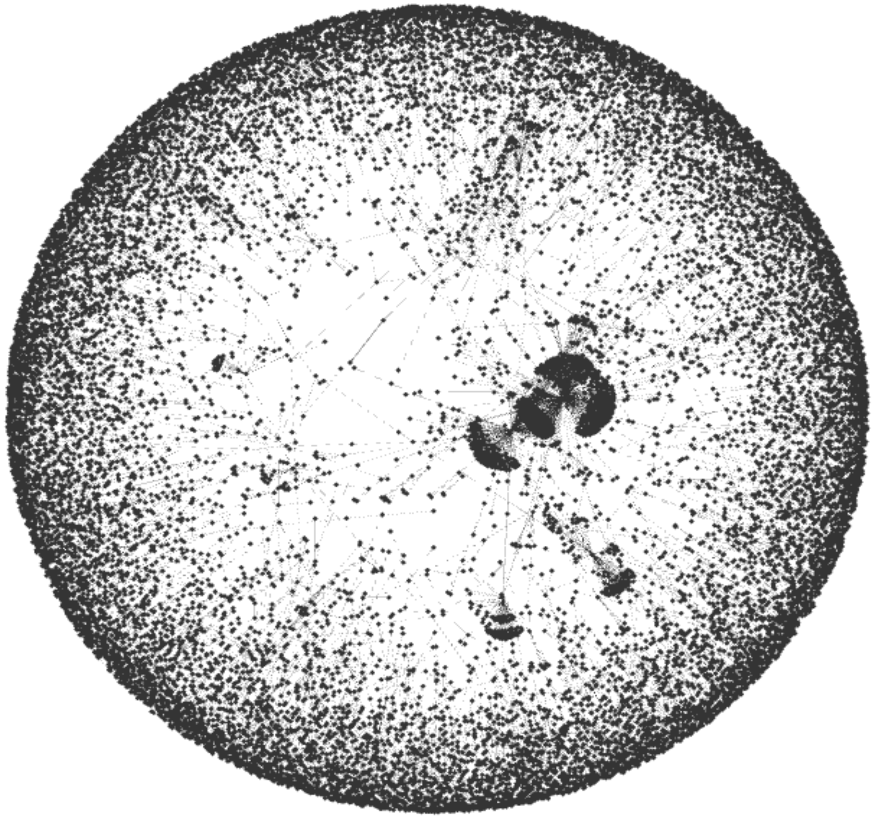
- Library: Python NetworkX
- Visualization: Gephi

3. Risk based scoring model
- Transaction based risk scoring

LINE

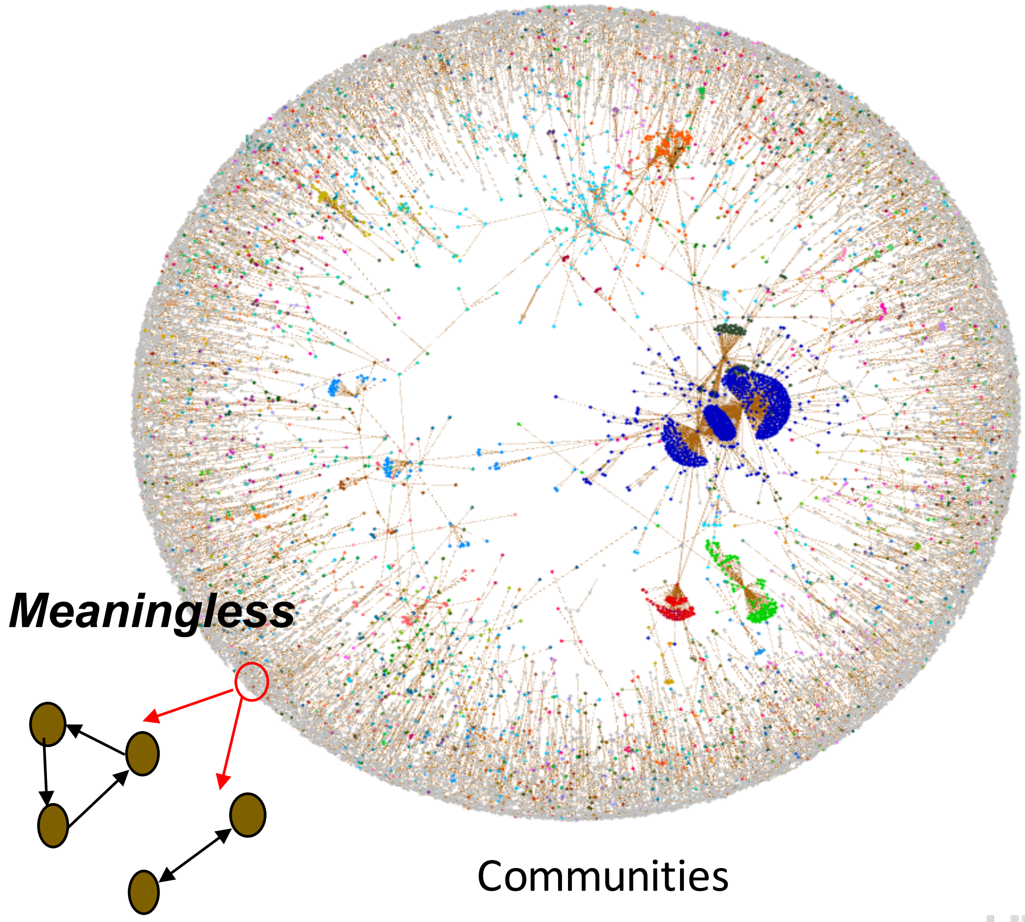
Analytics of multiple accounts

Community detection



A Network graph

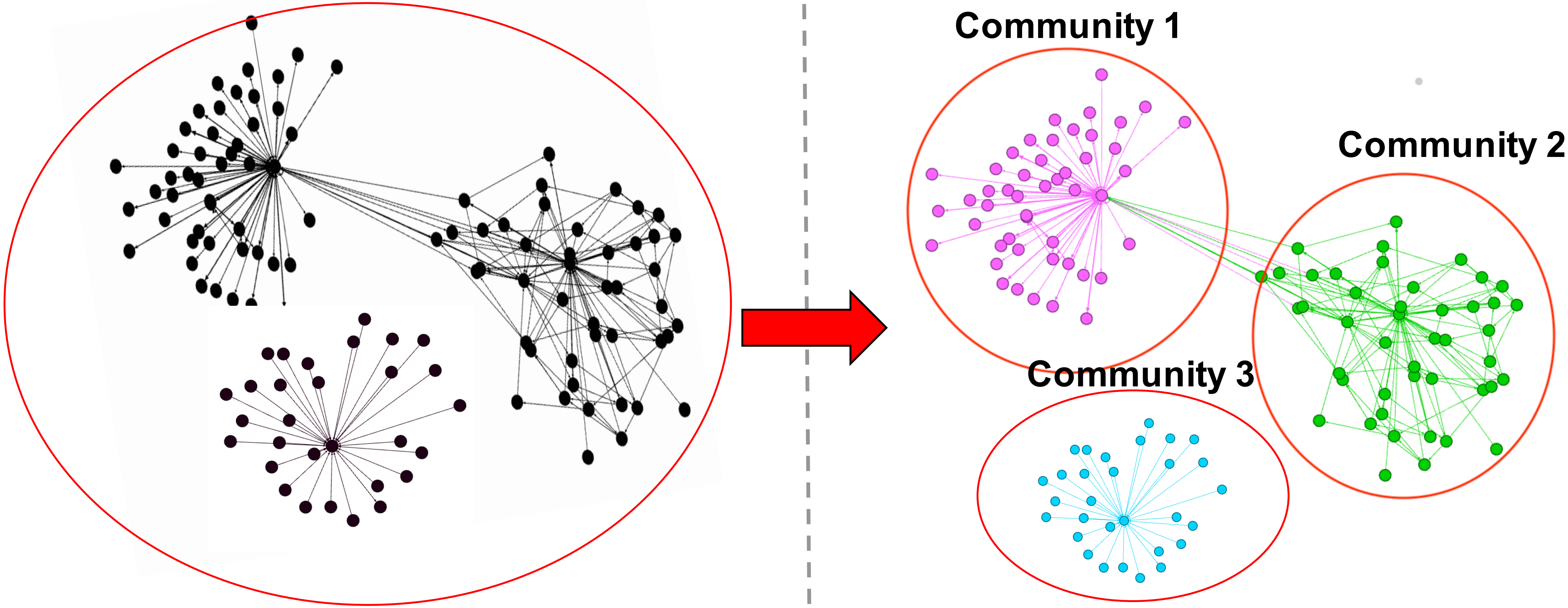
Modularity Score: 0.94



Communities

Analytics of multiple accounts

Community detection



A Network graph

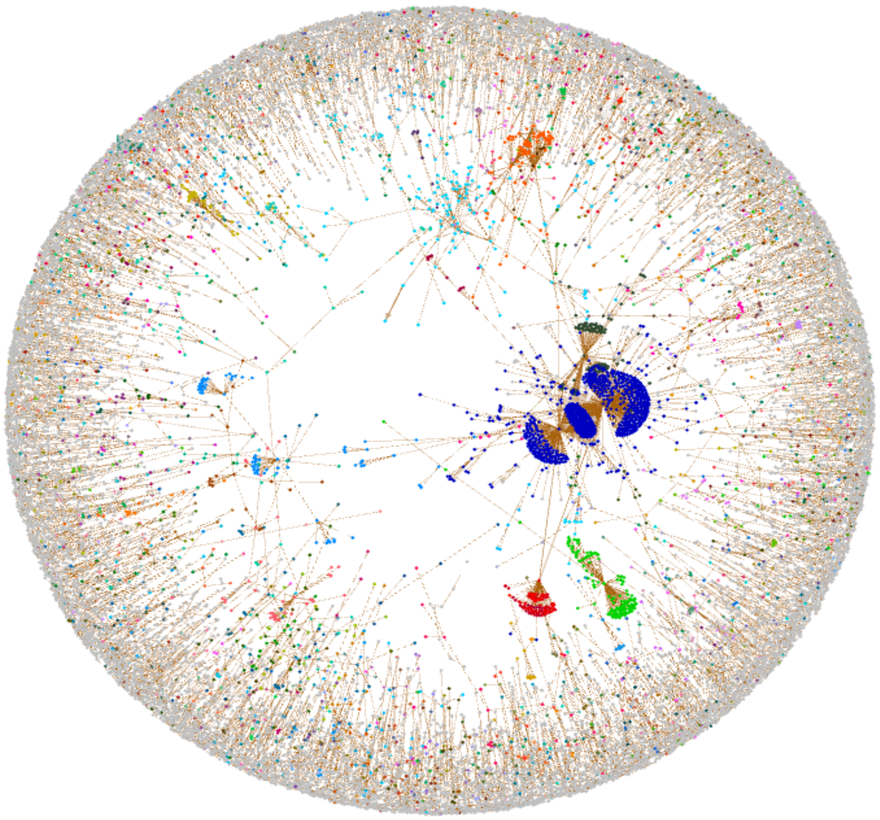
Communities

Analytics of multiple accounts

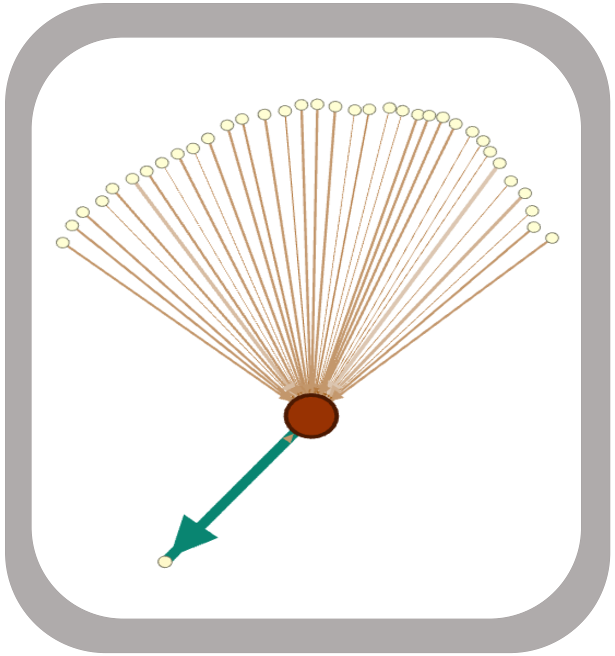
Definition of suspicious groups.

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*

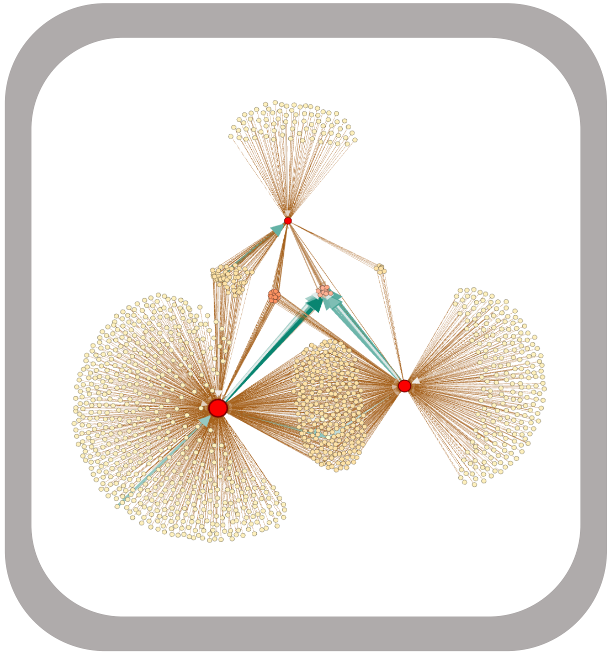
Analytics of multiple accounts



Promotion fraud



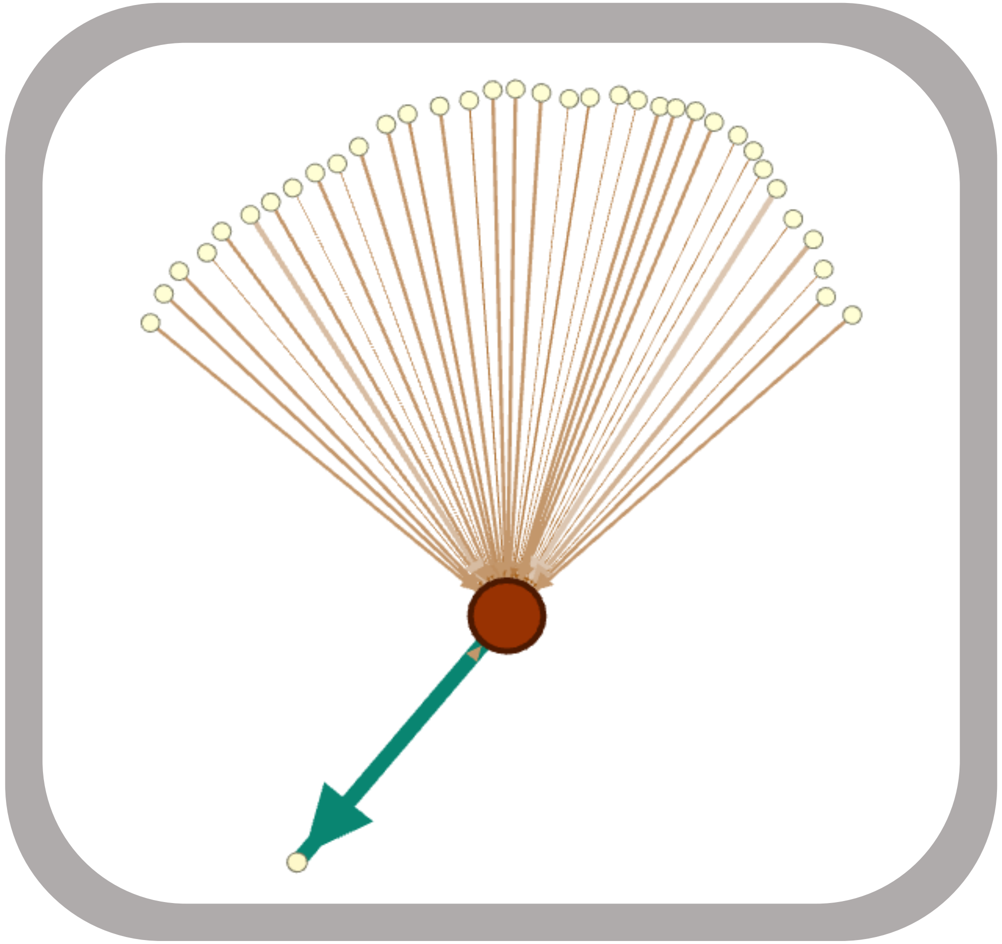
Layering pattern



LINE

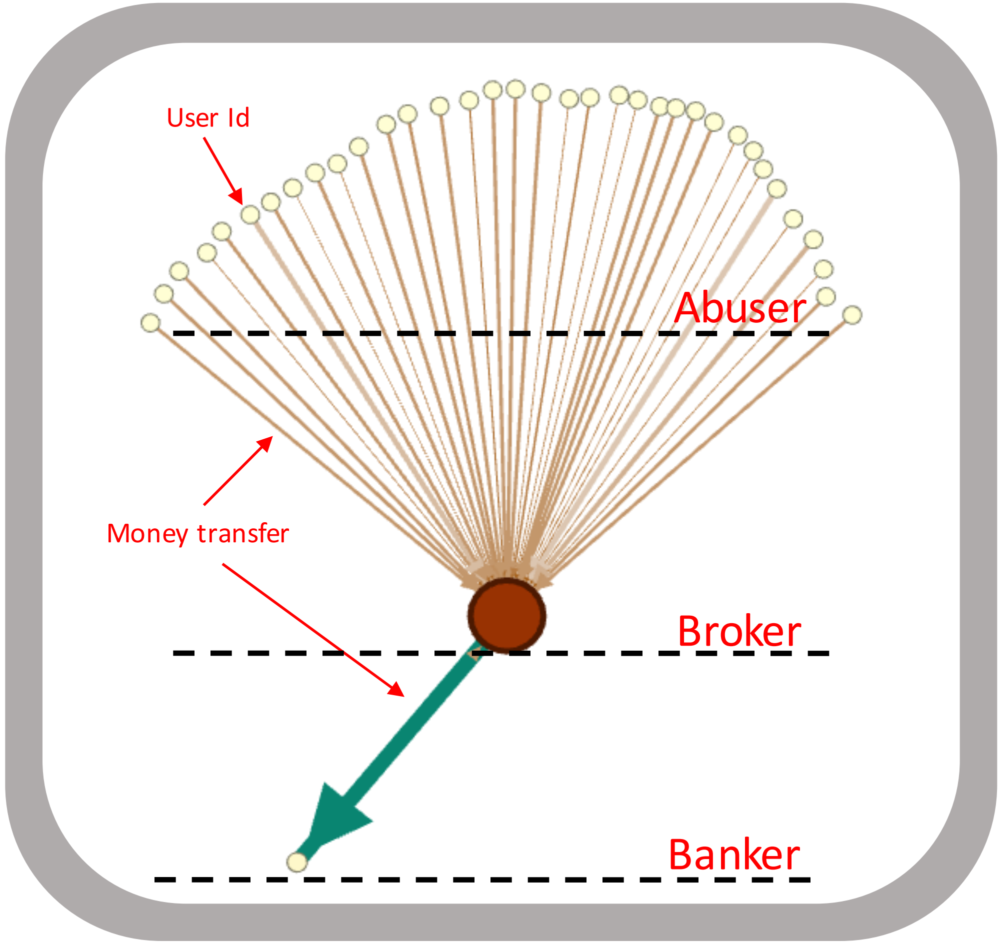
Analytics of multiple accounts

Promotion(Marketing) fraud



Analytics of multiple accounts

Promotion(Marketing) fraud

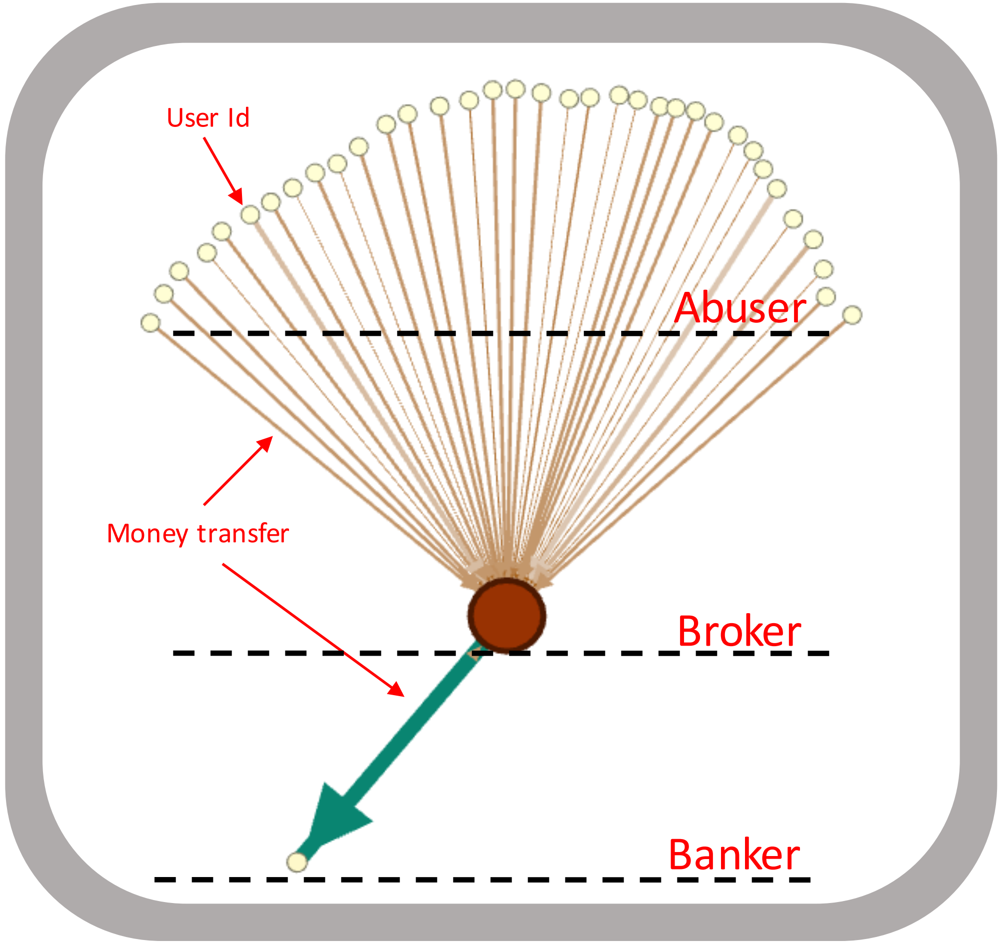


- *Abuser : A group(s) that seeks monetary benefits from the weaknesses of various promotions*
- *Broker : A user(s) who Collects money from abuser groups*
- *Banker : A user(s) who withdraws(spends) the money received from the broker(s)*

LINE

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*

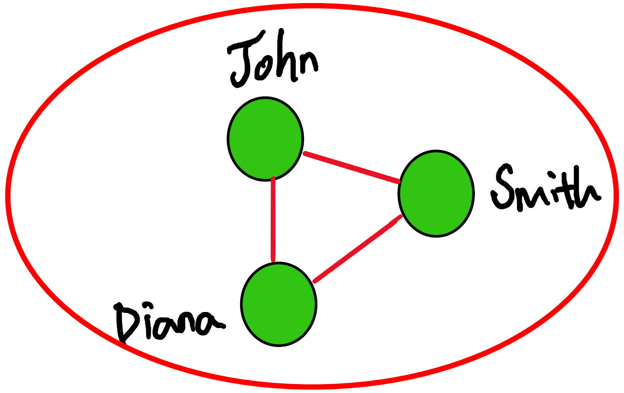
  → ***Transaction Similarity***

- *Short interval of transaction sequences*

  → ***Transaction Interval***

Analytics of multiple accounts

Combo Meal (Burger + Drink + Potato fries)



Friends community

- ***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 FBDFDFFDBBF DBFBDFBDBB*

John Smith Diana

Analytics of multiple accounts

Transaction sequence in communities

Deposit* = *D*, *Payment* = *P

Money Transferred* = *T

Money Received*= *R

Withdraw* = *W

C- 1 : DPPPPPPPPPPPPPPPPPPPDPPPPPTPPPTTPPDPPDPPTPPDPPPDPDTDPPPPPPPPPPDPPPPPPPPPDPTDPTDPTTDPTDPTDPTTDPPPPPPPPPDPTDPTDPTT

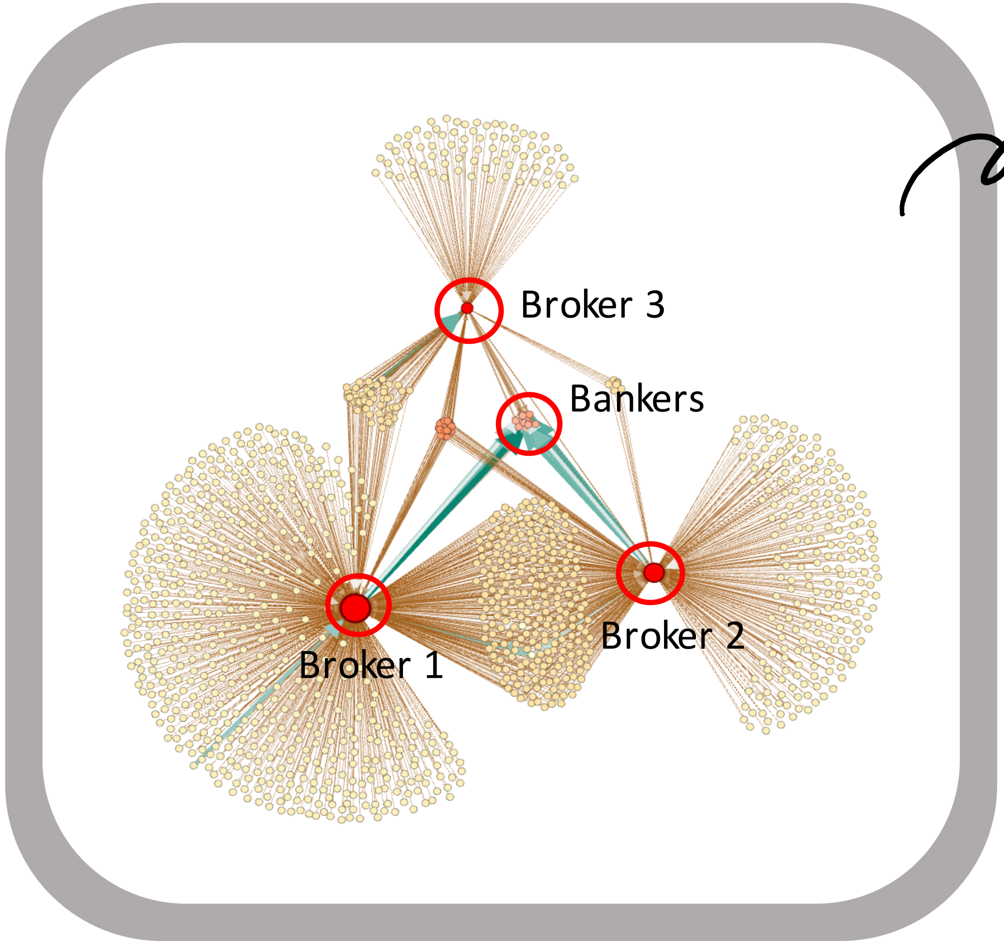
C- 2 : PPPPPPPPPPPPPPDPPPPPTTTPPPTPPPTPTPPPDDPPTPPDPPPDPDTDPPPPPPPPPPDPPPTPPPPPTTTPPPPPPPDPTDPTDPTTTPPPPPPPDPPDPTDPTDPT

C- 3 : DPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPTDPTTDPTDPT

C- 4 : DPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTDPTWWWW

Analytics of multiple accounts

Layering pattern



- *Size of community*
- *Amount of money*

Deposit = D

Payment = P

Money Transferred = T

Money Received= R

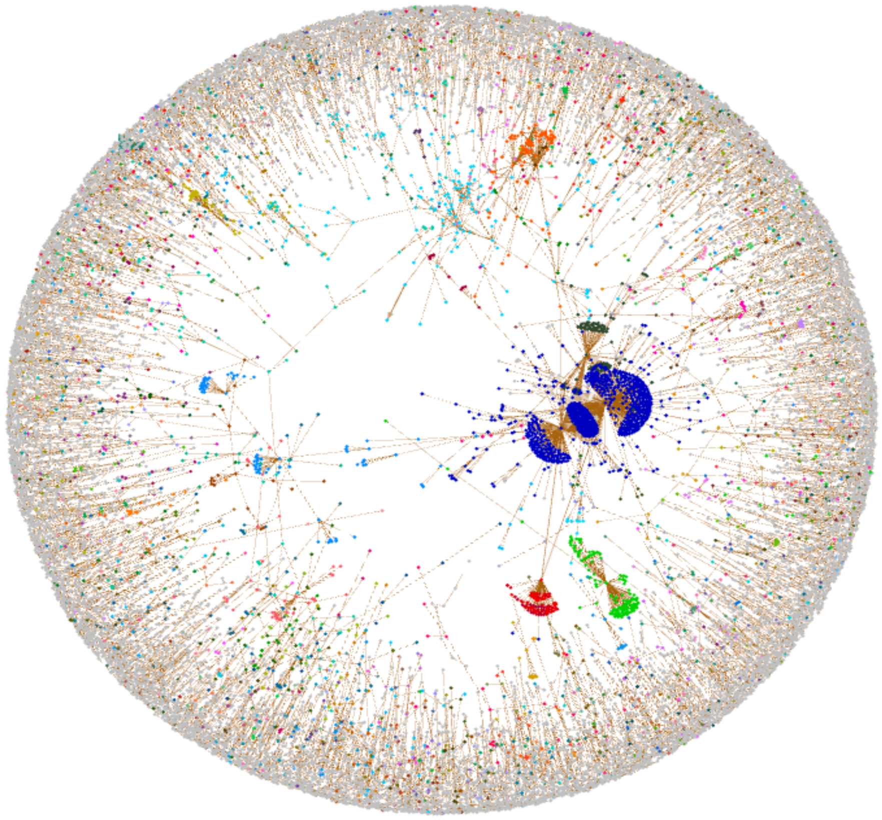
Withdraw = W

No payment action?

- ***Action sequence: DTRDTRDTRDTRDTRDTRDTRDT R…DTRDTRDTRDTRDTRDTRDTR DTRDTRDTRDT RDTRDTRDTRDTR…DTRDTRDTRDTRDTRDTRDTR DTRDTRDTRDTRDTRDTRDTRDTR…DTRDTRDTR DTRDTRDTRWWWWWWWWWWWWWWWWWW***

Analytics of multiple accounts

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



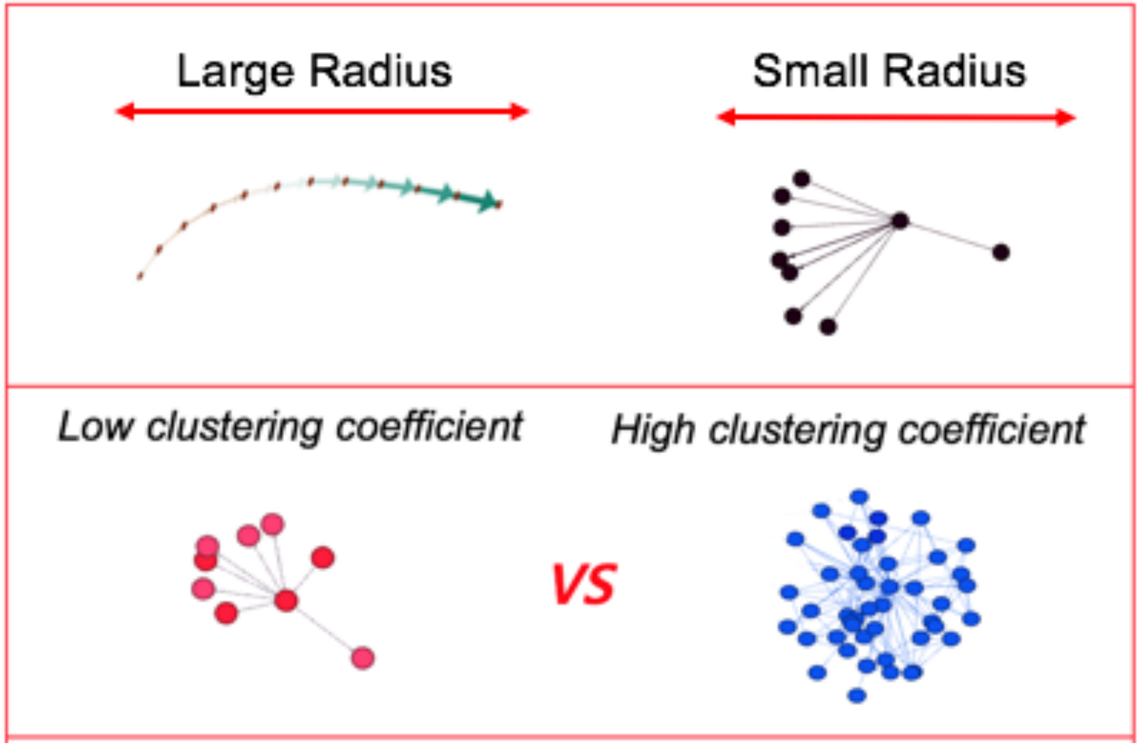
LINE

Analytics of multiple accounts

- *Key designed features of fraud using multiple accounts*

✓ *Graph shape*

1. *Diameter / Radius*
2. *Betweenness (Std)*
3. *Betweeness (Avg)*
4. *Clustering coefficient*
5. *Assortativity*
6. *Degree (Std)*
7. *Dgree (Avg)*



Analytics of multiple accounts

- *Key designed features of fraud using multiple accounts*

  ✓ *Graph shape (Radius, Betweenness, Assortativity, ..)* → *Identify suspicious communities*

  ✓ *Service loyalty Scoring*

  ✓ *Transaction Similarity*

  ✓ *Transaction Interval*

  → *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

Whitelist model

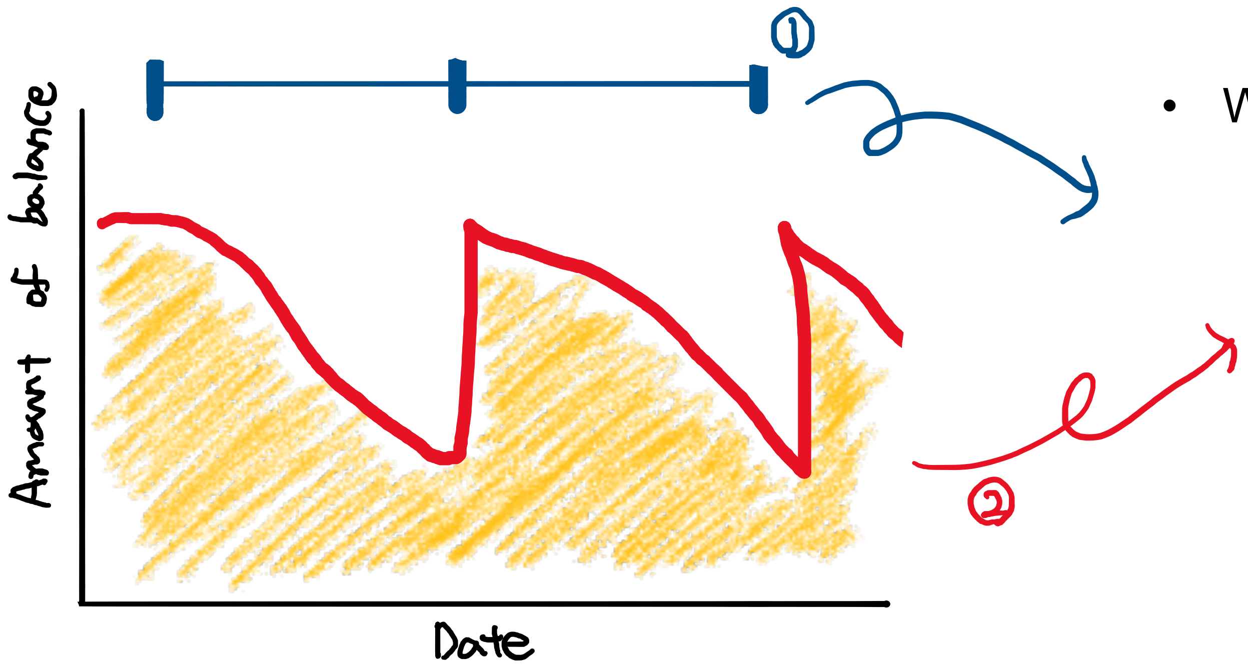
Q. Why do we need a risk based whitelist model?

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

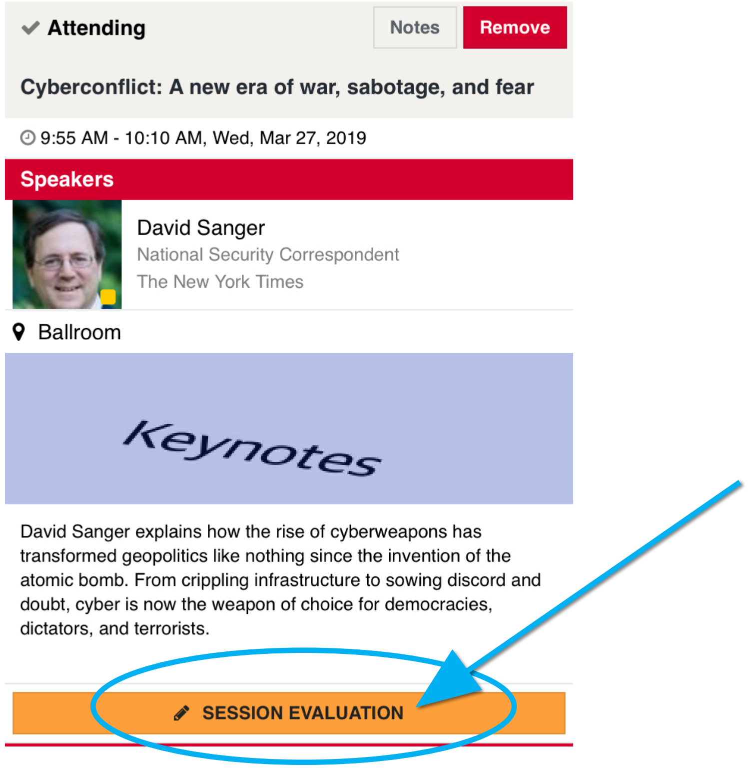


- Who we can trust?
  - ① Cycle of deposit
  - ② Cycle of payment
  - ③ Service loyalty score
  - ④ …Etc

Thank you

Rate today’s session

Session page on conference website



O’Reilly Events App