RUTGERS

Rutgers Business School
Newark and New Brunswick

# Anomaly detection: Transitory Accounts

*Presented by Yongbum Kim*

# Objectives

- To develop and test anomaly detection models that filter the abnormal transactions out of millions of transactions from about ten thousand transitory bank accounts.

- General Constraints
  - Implementable without consuming significant data processing resources
  - General enough to be applied to all the transitory accounts.

# Transitory Account

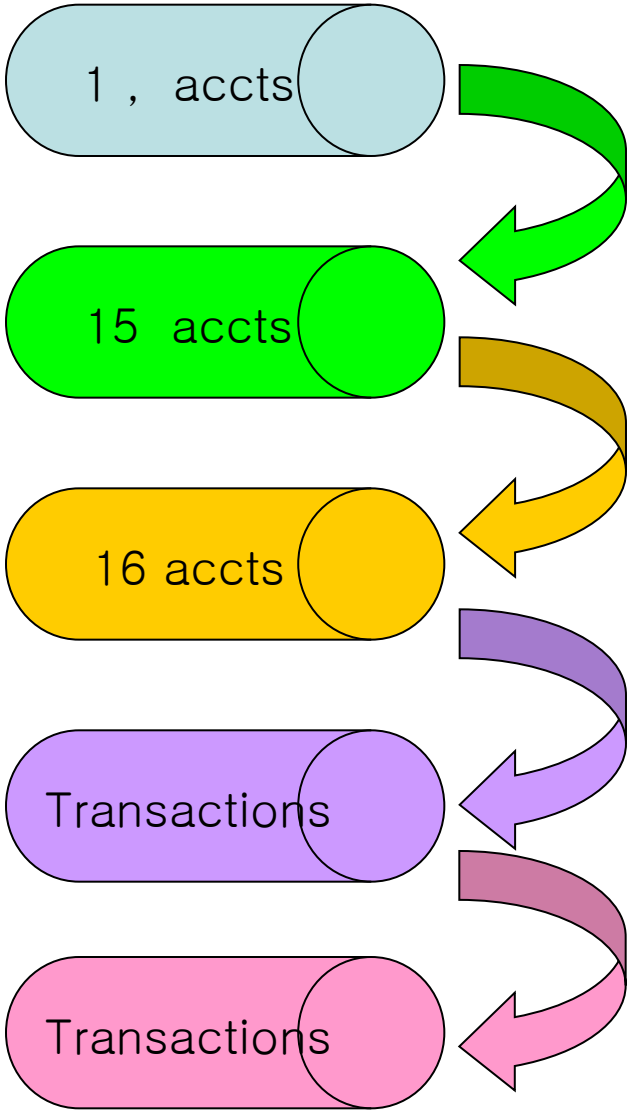
- A bank account that is used to keep money temporarily in the process of wire transfers.
- About ten thousand transitory accounts each of which has about thousands of transactions per month.
- Its transactions must be cleared eventually.
- Any accounts can become transitory accounts.
- A source of employee's frauds

# Challenges

- Data
  - Each account has different characteristics.
    - Amount distribution
    - The number of transactions and transactional days
  - Highly skewed and peaked

- Practical restrictions
  - Limited resource
    - Run at the mainframe level
  - Limited internal auditors
    - The number of alarms should be manageable.
  - Significant amount
    - Only transactions with significant amounts are of interest.

# Level 1-2-3 screening




By survey
= 5 (Finance) + 5 (Consumer)
+ 5 (Insurance)

Level 1 screening
by Analytical Review

Level 2 screening
By P-rules (Kurtosis & Skewness)

Level 3 screening
By detailed rules

# Concerns about the filter generation

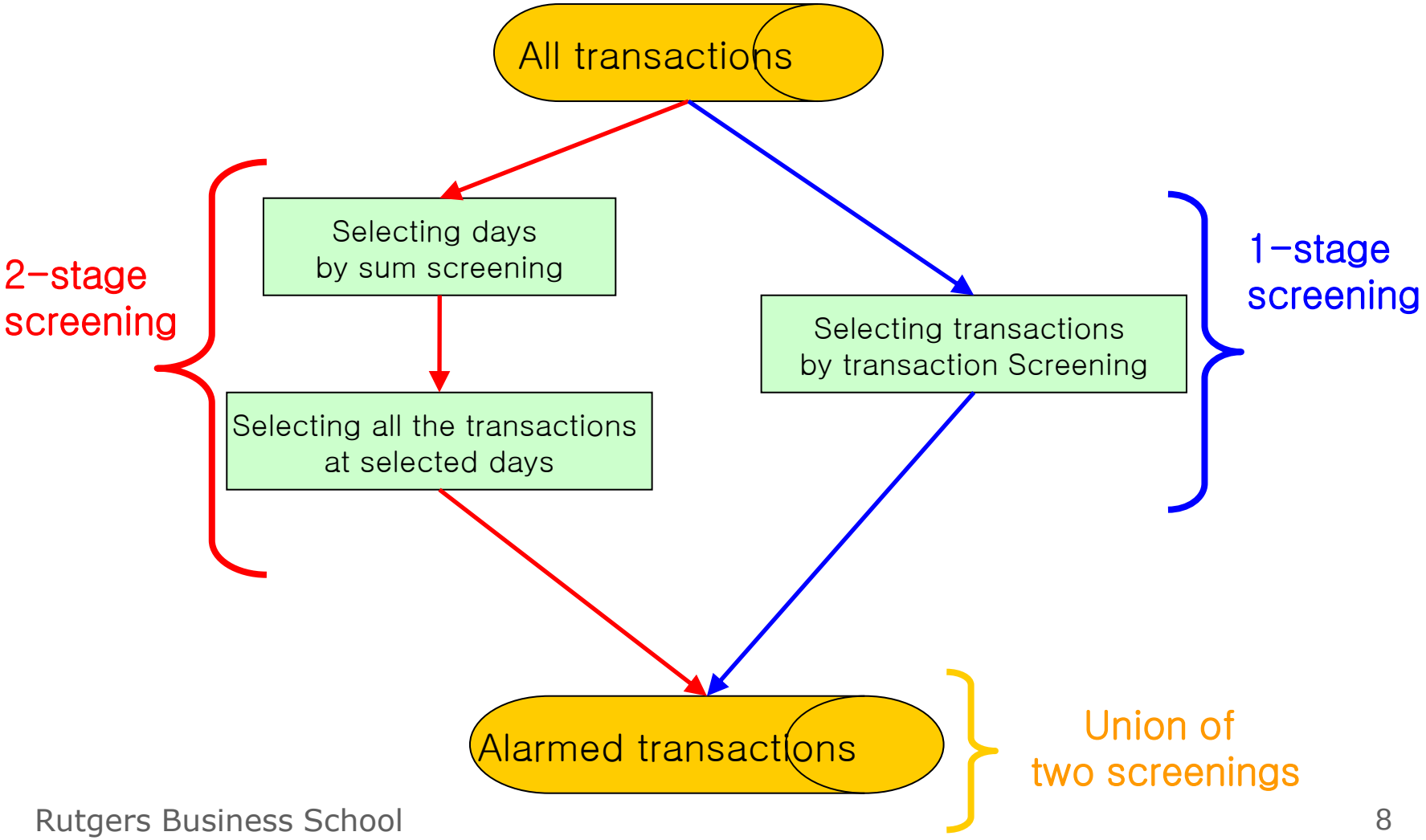
- Implementability
  - Implementable at the mainframe level.
  - Requiring as little computational power as possible.
- Materiality of transaction amounts.
- Not too many alarms

# Data Description

- Summary statistics
  - 16 transitory accounts with 17 variables
  - Data sets after truncation
    - Train Set: 12/2  9~ 2/2 1
    - Test Set:  4Apr2 1 ~ 9Jul2 1

    - Statistics of the Train set is used for level 2 screening.
    - Test set has regularization records on which several filters are based.

# Level 2 Screening

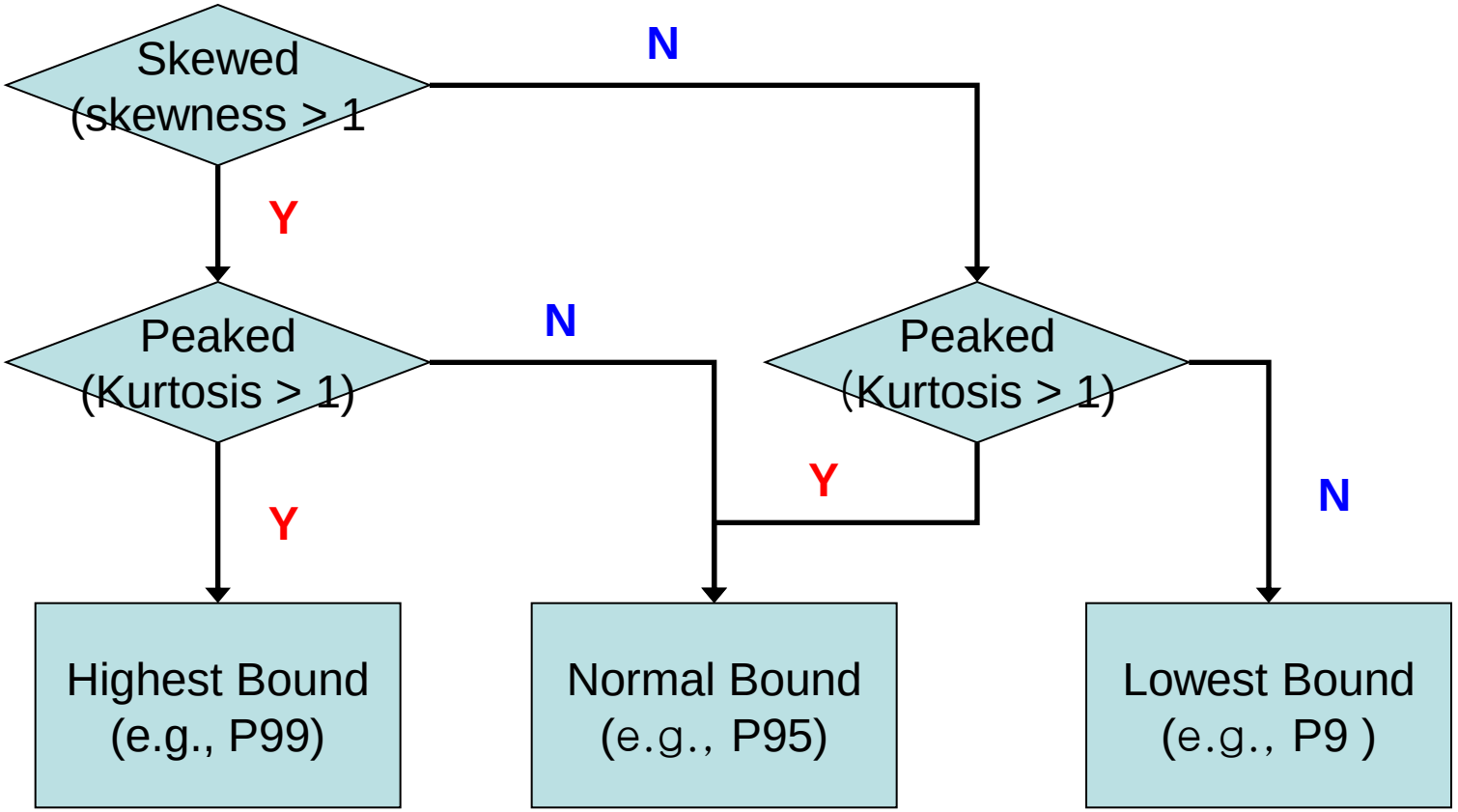


# Level 2 screening: Two things considered…

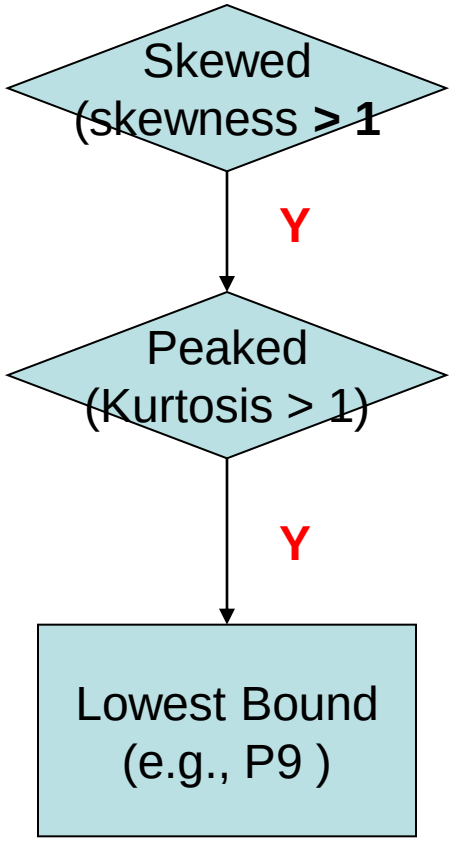
- Two-stage screening
  - Select days that have abnormally high sums
  - Select all the transactions at the screened days

- One-Stage screening
  - Select the transactions whose amounts are abnormally high.

# SK Screening



# An Example (account 5738)




skewness =179.3548 5

Kurtosis = 39133.2558

# Level 3 screening

1. Debit vs. Credit (InDbCr field)

Transactions on the same account are to have the same value either Debit or Credit.

2. Manual entries

Whether a transaction is manually processed

3. Duplicate amounts

Transactions with duplicate amounts for the same account within the same branch at the same day exist.

4. Aging
5. Processed during weekends
6. Wrong dates: balance date < amt date
7. Duplicate Nsue
8. Multiple regularizations

The more regularization occurs, the more likely its purpose is to avoid possible transaction amount check (i.e., splitwire).

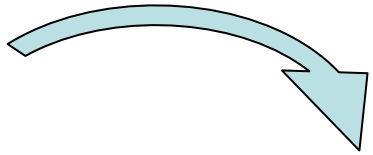
9. Late regularization (first regularization day - the transaction posting day >=30)

Here, the focus is the first regularization.

10. Manual/uncertain regularization
11. Regularized during weekends
12. Multiple alarms: Group by LANVFCDPCB, LANVFDTLANC, LANVFCDUNID

# Level 3 screening

| score | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| cnt_transactions | 281 | 2 2 | 37 | 8 | 1 |



| DbCr | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| wrongDate | | | | | | | | | | |
| cdfunc_lanco | | | | | | | 1 | 1 | 1 | 1 |
| nDup_within | | | | | 1 | 1 | | | | |
| outstandOver3 | | | | 1 | | | | | | |
| amt_wk | | | | | | | | | | |
| bal_wk | | | | | | | | | | |
| Dup_Nsue | | | | | | | | | | |
| multi_regul | | 1 | 1 | | | | | | | 1 |
| manual_regul | 1 | 1 | 1 | | | 1 | | 1 | 1 | 1 |
| regul_wk | 1 | | 1 | 1 | 1 | | 1 | | 1 | 1 |
| FirstRegulafter3 | | | | 1 | | | | | | |
| Score | 2 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | 3 | 4 |
| cnt | 16 | 4 | 5 | 1 | 7 | 2 | 6 | 2 | 2 | 1 |

# Where we are…

- 46 wires were flagged for further investigation.
- Although certain filters alarm transaction for the population, they do not for the level 2-screened one.
  - This may indicate that the level 2 screening need less strict thresholds.
- Several accounts do not have as many transactions as before.
  - This long-term pattern might affect the model's effectiveness.

# How many alarms are practical?

- 3 transactions/hour can be investigated by an internal auditor.
  - 24 transactions/day (=3 * 8 hours)
  - If there is 1 alarm/account each day, approximately 417 auditors in full charge are necessary for their investigation.
  - If 10 auditors are available, one alarm is maximum for every 42 accounts.
  - This indicates that screening need to narrow down the scope more.

# Questions & Suggestions?