DIDD
DEPARTEMEN INOVASI & DIGITALISASI DATA

Data at your fingertips.

BANK INDONESIA
BANK SENTRAL REPUBLIK INDONESIA

Unveiling the Interconnectedness of Banks in Payment Systems: Methodology, Utilization, and Data Governance Considerations

17 October 2023

Renardi Ardiya Bimantoro, Mohammad Khoyrul Hidayat, Irfan Sampe
Department of Data Innovation and Digitalization, Department of Statistics – Bank Indonesia
Email: renardi.ardiya@bi.go.id; moh.khoyrul@bi.go.id; irfans@bi.go.id

Background

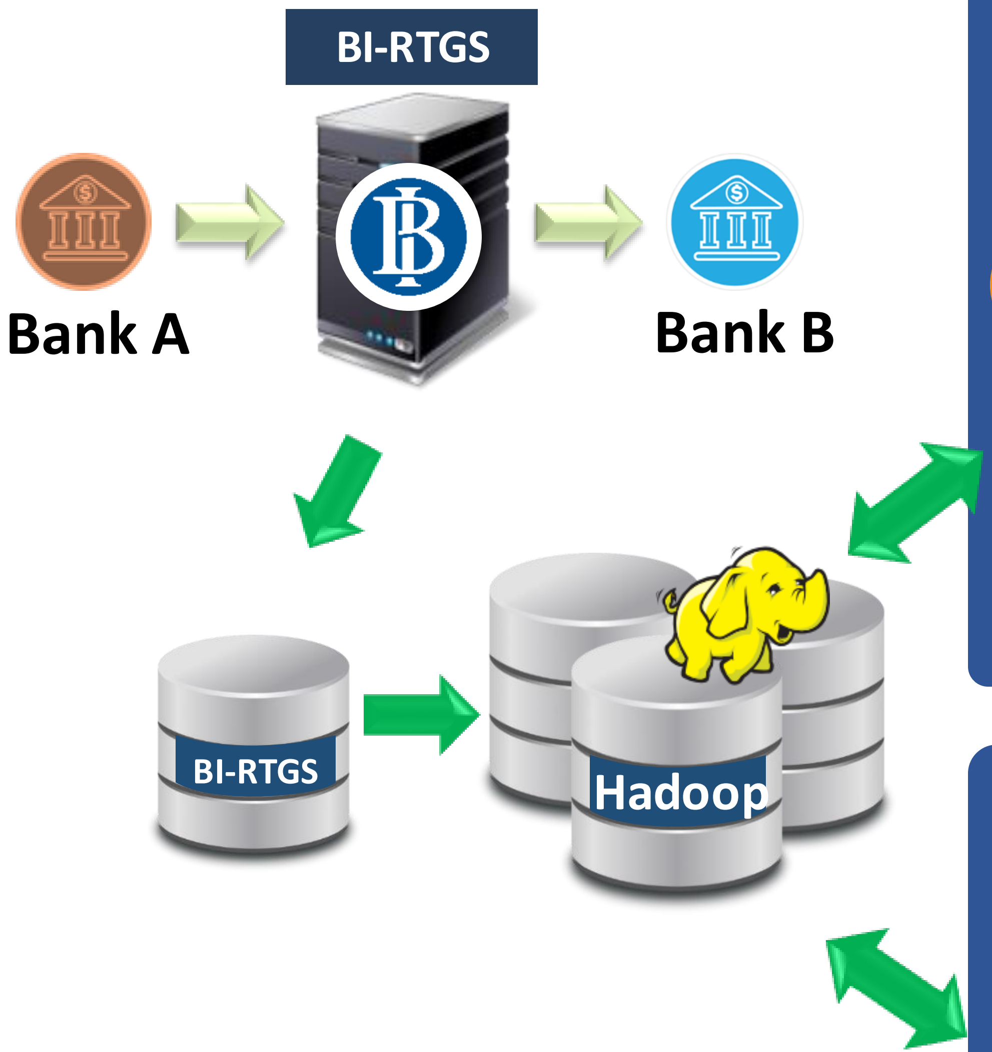
BACKGROUND

- In the payment system, every bank is interconnected with each other. When problems occurs in one bank, they can potentially have ripple effects on other linked institutions. Therefore, it's crucial to identify the interconnected structure of participants in the payment system.
- Leveraging Big Data Analytics through the core-periphery model to visualize the systemic structure (interbank tiering) comprehensively serves to complement the interconnectedness analysis that has been conducted so far.

OBJECTIVE

Discover the interlinked structure of participants in the payment system by applying network analysis method.

DATA SOURCES & METHODOLOGY



- ✓ Total transaction rows: ± 1 million rows per month
- ✓ Number of banks transacting: 117 banks
- ✓ Transaction amount ≥ Rp.100 million or ≥ USD 6,390

Data Preparation

1. Filtering & Mapping

- Filter out non-bank BI-RTGS participants names.
- Map branch unit bank to their parents bank
- Codify the Bank Name using Bank Code

2. Data Aggregation & Matrix Representation

- Aggregate transaction data on a monthly basis.
- Represent transaction data in discrete matrices (connectivity) & continuous matrices (transaction amount/frequency).

Data Extraction

Model *Core - Periphery*

3. Model Asymmetric Discrete

Identify Core & Periphery banks based on connectivity (presence or absence of transactions).

4. Model Asymmetric Continuous

Calculate out-coreness and in-coreness based on transaction amount & frequency.

5. SVM Classification

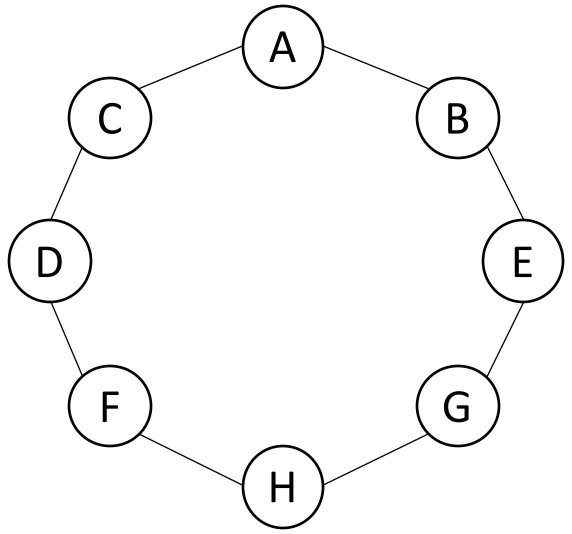
Classification using Support Vector Machine (SVM) based on the results of discrete and continuous models.

Interbank & Payment System Network

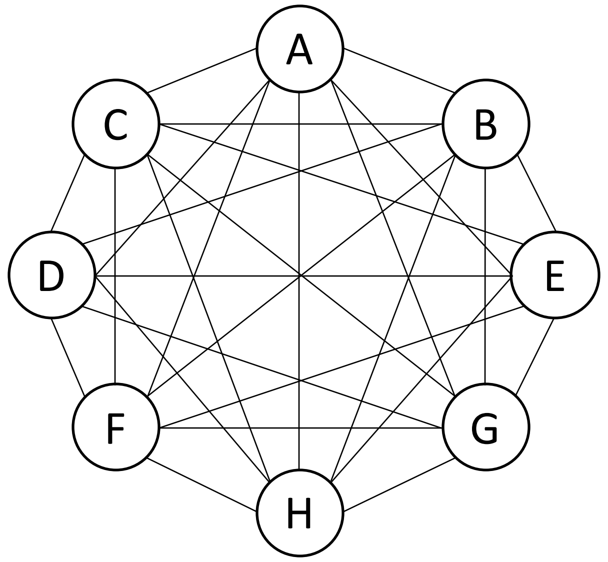
Numerous studies on interbank & payment system networks indicate that the observed structure exhibits a hierarchical pattern, often referred to as interbank tiering.

Interbank tiering describes the interbank structure that emerges when there are interconnected banks also function as intermediaries between non-interconnected banks (Craig & Peter, 2010).

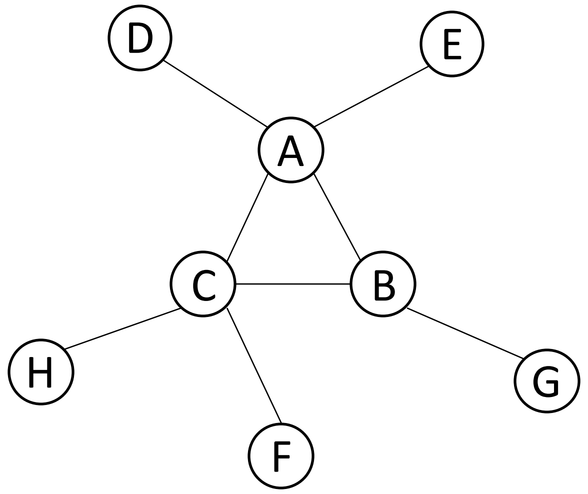
The model often used to identify such network structures is the core-periphery model (Craig & Peter, 2010).



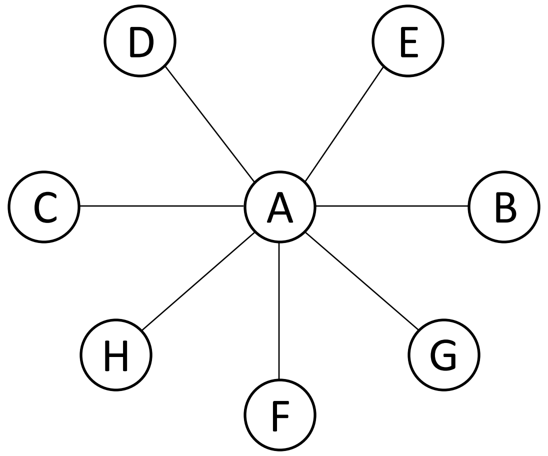
Ring



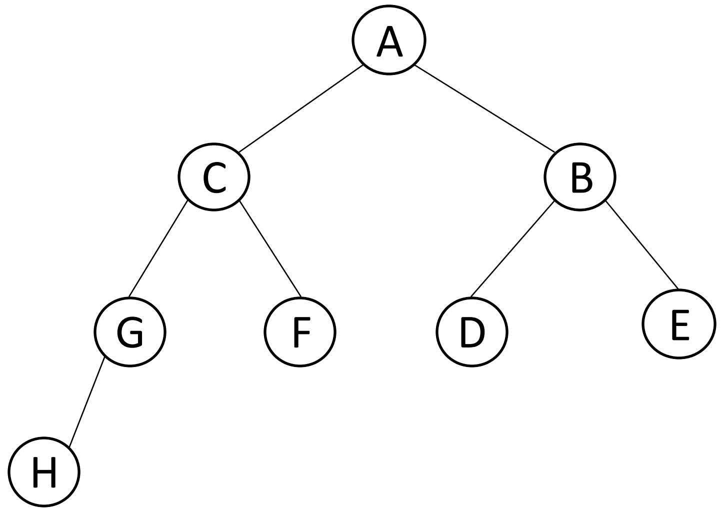
Full Mesh



Core – Periphery



Star



Tree

Core-Periphery Model Type

	A	*B*	*C*	*D*	*E*
A	0	1	1	0	0
B	1	0	0	1	1
C	1	0	0	1	0
D	0	1	1	0	1
E	0	1	0	1	0

Matriks Biner Simetris

Model *Symmetric Discrete* Borgatti dan Everett, 1999	**Model *Symmetric Continuous*** Borgatti dan Everett, 1999
Objective: To classify nodes in an undirected network into core and periphery nodes.	**Objectives:** To determine the coreness value of each node in an undirected network.
Data: Connectivity between nodes in a network (0: not connected, 1: connected).	**Data :** Edge weights in the network with continuous values (e.g., based on transaction amount or frequency).
Model *Asymmetric Discrete* Craig dan Peter, 2010	**Model *Asymmetric Continuous*** Boyd et al., 2010
Objective: To classify nodes in a directed network into core and periphery nodes.	**Objectives:** To determine the coreness value of each node in a directed network
Data : Connectivity between banks in a network (0: not connected, 1: connected).	**Data :** Edge weights in the network with continuous values (e.g., based on transaction amount or frequency).

	A	*B*	*C*	*D*	*E*
A	0	10	5	0	0
B	10	0	0	15	25
C	5	0	0	3	0
D	0	15	3	0	30
E	0	25	0	30	0

Matriks Kontinu Simetris

	A	*B*	*C*	*D*	*E*
A	0	1	1	0	0
B	0	0	0	1	1
C	1	1	0	1	0
D	1	0	0	0	1
E	1	0	1	0	0

Matriks Biner Asimetris

	A	*B*	*C*	*D*	*E*
A	0	10	18	0	0
B	0	0	0	20	1
C	15	10	0	2	18
D	30	0	0	0	10
E	25	0	3	0	0

Matriks Kontinu Asimetris

Model used to identify BI-RTGS *network*

Model Asymmetric Discrete

Using blockmodelling techniques to identify core and periphery banks based on interbank connectivity patterns. (Craig & Peter, 2010)

Blockmodelling is a network analysis method that partitions a network into several groups (also called blocks), where each group consists of nodes that have similar connectivity patterns with other members.

The ideal blockmodel is achieved when the three properties of the model are met, namely:

- ✓ ***Top-tier banks*** (***core***) are interconnected with each other. (CC)
- ✓ ***Lower-tier banks*** (***periphery***) are not interconnected with each other. (PP)
- ✓ ***Top-tier banks (core)*** are connected to (some) ***lower-tier banks (periphery)***. $(CP\ \&\ PC)$

$$\begin{bmatrix} CC & CP \\ PC & PP \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \boldsymbol{RR} \\ \boldsymbol{CR} & \mathbf{0} \end{bmatrix}$$

Blok *CC* (*Core to Core*) : Connectivity between core banks
Blok *PP* (*Periphery to Periphery*) : Connectivity between *periphery* banks
Blok *CP* (*Core to Periphery*) : Connectivity from *core bank* → *periphery bank*
Blok *PC* (*Periphery to Core*) : Connectivity from *periphery bank* → *core bank*

$\boldsymbol{RR}$: *At least* 1 *connection in every row* $\boldsymbol{CR}$: *At least* 1 *connection in every column.*

$$\begin{array}{cc} & \begin{array}{ccc|ccccc} & \text{Core} & & & & \text{Periphery} & & \\ A & B & C & D & E & F & G & H \end{array} \\ \begin{array}{c} \text{Core} \\ \\ \\ \text{Periphery} \\ \\ \\ \\ \end{array} \begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \end{array} & \left[\begin{array}{ccc|ccccc} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \end{array}$$

Example of block model with 3 Cores (A, B, C)

Model Asymmetric Discrete

Let's suppose there are n banks labeled 1,2,..., n.

- The interbank connectivity data (transactions) is stored in matrix N where:

$$N_{ij} = \begin{cases} 1, & if\ there\ is\ a\ transaction\ from\ i\ to\ j \\ 0, & if\ there\ is\ a\ transaction\ rom\ j\ to\ i \end{cases}$$

for $i = 1,2,\ldots n \quad and \quad j = 1,2,\ldots n$

- Define C^* as a set of banks designated as core, and N_{C^*} is the blockmodel of matrix N

Objective Function

$$\arg\min_{C^*} \left[e(N_{C^*}) = \frac{e_{cc} + e_{pp} + (e_{cp} + e_{pc})}{\sum_i \sum_j N_{ij}} \right]$$

Where

$$E = \begin{bmatrix} e_{cc} & e_{cp} \\ e_{pc} & e_{pp} \end{bmatrix} = \begin{bmatrix} c(c-1) - \sum_{i \in C^*} \sum_{j \in C^*} N_{ij} & (n-c) \sum_{i \in C^*} \max\left\{0, 1 - \sum_{j \notin C^*} N_{ij}\right\} \\ (n-c) \sum_{j \in C^*} \max\left\{0, 1 - \sum_{i \notin C^*} N_{ij}\right\} & \sum_{i \notin C^*} \sum_{j \notin C^*} N_{ij} \end{bmatrix}$$

$e(N_{C^*})$ = Total *error* (*blockmodel* N_{C^*})
e_{CC} = *Error* at block *CC*
e_{PP} = *Error* at block *PP*
e_{CP} = *Error* at block *CP*
e_{PC} = *Error* at block *PC*
N_{ij} = Transaction Matrix Element *at* row i and column j

The optimal partition is a set of core banks (C^*) that produces the smallest error value ($e(N_{C^*})$) (closest to the ideal blockmodel).

Model Asymmetric Discrete

$$
\begin{array}{c}
 \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \\ \mathbf{G} \\ \mathbf{H}
\end{array}
\begin{array}{c}
\begin{array}{cccccccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} & \mathbf{E} & \mathbf{F} & \mathbf{G} & \mathbf{H} \end{array} \\
\begin{bmatrix}
0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
\xrightarrow{?}
\begin{array}{c}
 \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \\ \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \\ \mathbf{G} \\ \mathbf{H}
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & & \mathbf{D} & \mathbf{E} & \mathbf{F} & \mathbf{G} & \mathbf{H} \end{array} \\
\begin{bmatrix}
0 & 1 & 1 & | & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & \color{red}{0} & | & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & | & 0 & 0 & 0 & 0 & 1 \\
- & - & - & + & - & - & - & - & - \\
1 & 0 & 0 & | & 0 & 0 & 0 & 0 & \color{red}{1} \\
0 & 1 & 0 & | & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & | & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & | & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & | & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

- ❑ If the number of nodes in a network is n, then there are $2^n - n - 1$ choices in determining which nodes become the core in order to produce an optimal blockmodel (closest to the ideal blockmodel).
- ❑ How to obtain the optimal blockmodel?

- ✓ **Heuristic algorithms can be used to obtain the optimal blockmodel.**
- ✓ **According to Borgatti (1999), the Genetic Algorithm is better for finding the optimal blockmodel as it provides more stable results compared to other algorithms.**

Model Asymmetric Continues

The matrix decomposition method is utilized to measure the 'core' strength of each bank. (Fricke & Lux, 2012)

Pengirim	Penerima	Nominal Transaksi
A	B	143
A	C	95
A	E	61
B	A	61
B	C	70
B	D	72
C	A	66
C	B	86
C	E	50
D	A	20
E	B	29

$$\underbrace{\begin{array}{c c} & \begin{array}{ccccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} & \mathbf{E} \end{array} \\ \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \end{array} & \begin{bmatrix} 0 & 143 & 95 & 0 & 61 \\ 61 & 0 & 70 & 72 & 0 \\ 66 & 86 & 0 & 0 & 50 \\ 20 & 0 & 0 & 0 & 0 \\ 0 & 29 & 0 & 0 & 0 \end{bmatrix} \end{array}}$$

$\boldsymbol{N}$: **Matriks *Adjacency***
nominal transaction data

$$\boldsymbol{u} = \begin{array}{cc} \begin{bmatrix} 0.355 \\ 0.340 \\ 0.232 \\ 0.016 \\ 0.042 \end{bmatrix} & \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \end{array} \end{array}, \boldsymbol{v} = \begin{array}{cc} \begin{bmatrix} 0.212 \\ 0.396 \\ 0.236 \\ 0.130 \\ 0.113 \end{bmatrix} & \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \end{array} \end{array}$$

- $\boldsymbol{u_i}$: ***coreness*** bank i as sender
- $\boldsymbol{v_j}$: ***coreness*** bank j as receiver

The fundamental concept of the **Asymmetric Continuous Model** is to decompose the adjacency matrix into '**out-coreness**' and '**in-coreness**' as follows:

$$\boldsymbol{N} = \begin{cases} \boldsymbol{u}\boldsymbol{v}^T, & off-diagonal \\ 0, & diagonal \end{cases}$$

where $\boldsymbol{u}$: **out-coreness vector** & $\boldsymbol{v}$: **in-coreness vector**.

Model Asymmetric Continues

Suppose there are n banks labeled as $1,2,\dots,n$

- The nominal or frequency of interbank transactions is stored in the matrix N
- where: $N_{ij} = \begin{cases} transaction\ nominal\ or\ frequency, & if\ there\ is\ transaction\ from\ i\ to\ j \\ 0, & if\ there\ is\ no\ transaction\ from\ i\ to\ j \end{cases}$

  for $i = 1,2,\dots n\ \ dan\ \ j = 1,2,\dots n$
- Let u be defined as the out-coreness vector and v as the in-coreness vector.

Objective Function

$$\arg\min_{\boldsymbol{u},\boldsymbol{v}} \sum_{i} \sum_{j\neq i} \left(N_{ij} - \boldsymbol{u}_i \boldsymbol{v}_j^T\right)^2$$

Contoh*

	B	C	H	A	F	D	E	G
B	0	144	59	281	78	114	93	67
C	110	0	34	161	45	65	53	39
H	20	15	0	28	8	12	9	7
A	396	296	120	0	161	234	191	138
F	34	26	10	50	0	20	16	12
D	70	52	21	102	28	0	34	24
E	48	36	14	69	19	28	0	17
G	25	19	8	37	10	15	12	0

Normalisasi

	B	C	H	A	F	D	E	G
B	0	0.144	0.059	0.281	0.078	0.114	0.093	0.067
C	0.11	0	0.034	0.161	0.045	0.065	0.053	0.039
H	0.02	0.015	0	0.028	0.008	0.012	0.009	0.007
A	0.396	0.296	0.12	0	0.161	0.234	0.191	0.138
F	0.034	0.026	0.01	0.05	0	0.02	0.016	0.012
D	0.07	0.052	0.021	0.102	0.028	0	0.034	0.024
E	0.048	0.036	0.014	0.069	0.019	0.028	0	0.017
G	0.025	0.019	0.008	0.037	0.01	0.015	0.012	0

$$\boldsymbol{u} = \begin{array}{c} \begin{bmatrix} 0.85 \\ 0.42 \\ 0.24 \\ 0.24 \\ 0.15 \\ 0.10 \\ 0.07 \\ 0.05 \end{bmatrix} \end{array} \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \\ \mathbf{G} \\ \mathbf{H} \end{array}, \boldsymbol{v} = \begin{array}{c} \begin{bmatrix} 0.68 \\ 0.46 \\ 0.35 \\ 0.27 \\ 0.22 \\ 0.19 \\ 0.16 \\ 0.14 \end{bmatrix} \end{array} \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \\ \mathbf{G} \\ \mathbf{H} \end{array}$$

$$u_A v_B = (0.85)(0.46) = 0.391$$

$$u_B v_A = (0.42)(0.68) = 0.285$$

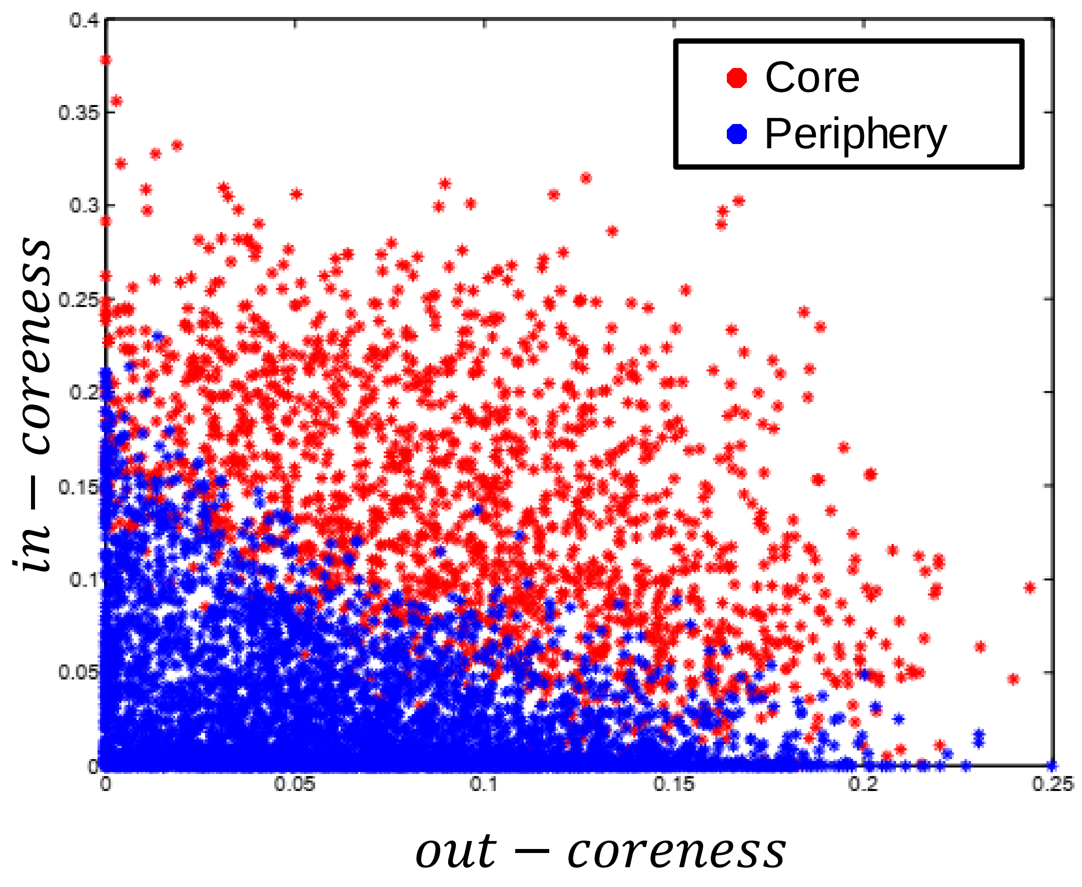
Model *Asymmetric Discrete*

- **Objective:** Classify banks into core and periphery.
- **Criteria:** Only considers the interbank connections in a network. **(0: not connected, 1: connected).**

Model *Asymmetric Continuous*

- **Objective:** Measure coreness, indicating the 'core' strength of each bank in a network.
- **Criteria:** Considers both transaction values and interbank connections when determining the coreness value of each bank.

Scatter plot of Coreness

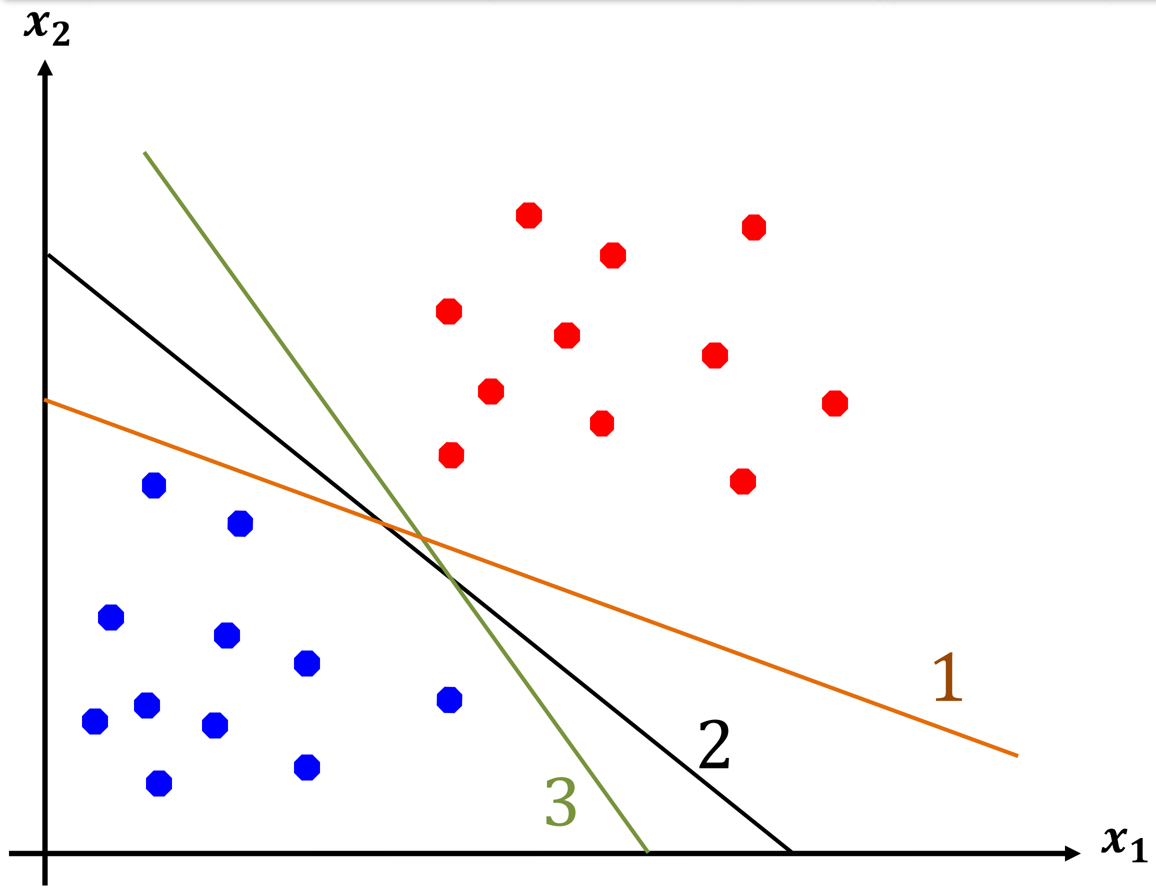


Core and periphery bank characteristic based on coreness value

- **Visualization:** A scatter plot of coreness values can be used to analyze the relationship between the results of the discrete model and the continuous model.
- **Plot Details:** The scatter plot is formed based on out-coreness values (x-axis), in-coreness values (y-axis), and the classification results of the discrete model (color of the dots).
- **Observation:** The out-coreness and in-coreness values of core banks should be significantly higher compared to periphery banks.

Model Asymmetric Continues

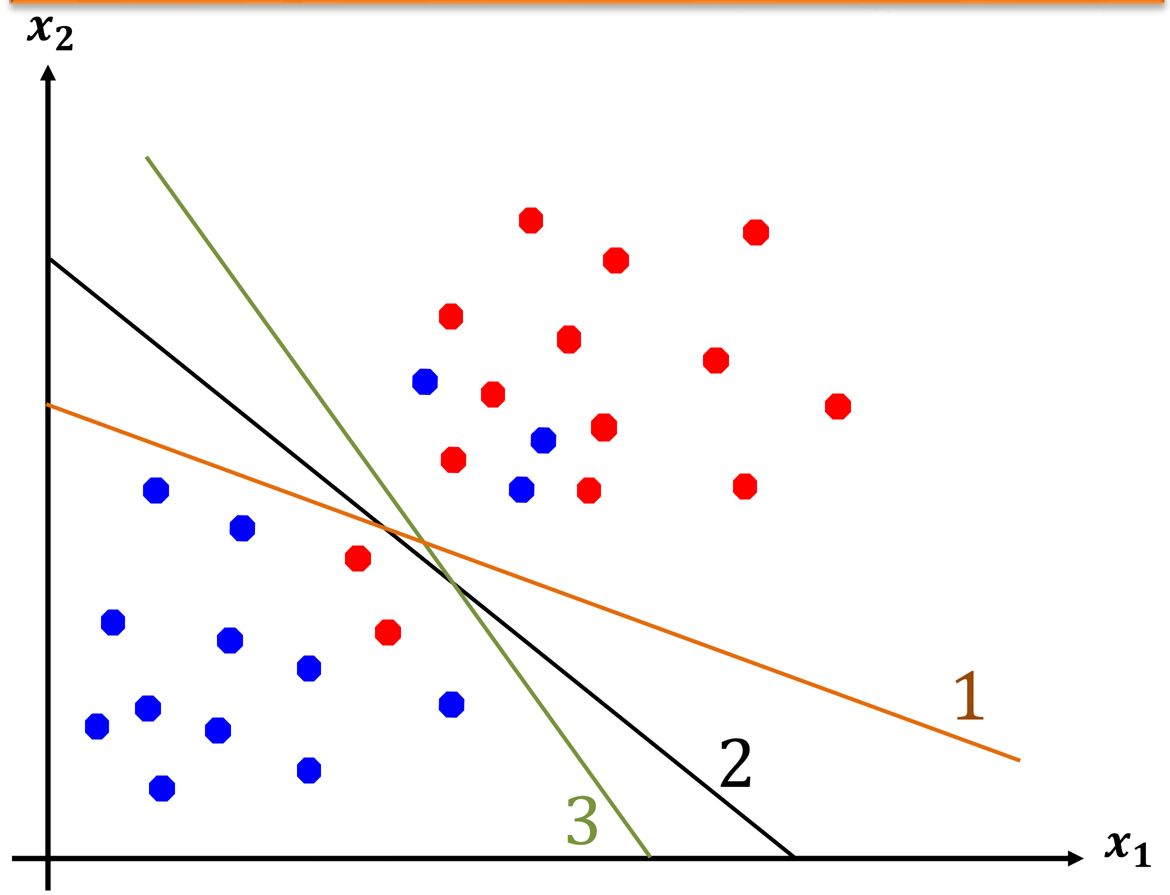
Ideal Cases



- Core
- Periphery

x_1 : *out – coreness*
x_2 : *in – coreness*

Non-Ideal Cases



- A **line equation** can be used as a **boundary that separates the core and periphery regions**.
- There **are several possible line equations**; which equation will be chosen or is the most optimal? **1, 2, or 3?**

Using machine learning algorithm to determine the core and periphery threshold area
Example: *Support Vector Machine* (SVM)

Core-Periphery Classification Using SVM

1 **Model**: Asymmetric Discrete
Algoritma: Genetic Algorithm

Identigy Core & Periphery Bank based on the interlink relationship (have transaction activity or not)

Core | Periphery

$$\begin{array}{c} \\ \text{Bank A} \\ \text{Bank B} \\ \text{Bank C} \\ \text{Bank D} \\ \text{Bank E} \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \left(\begin{array}{ccc|cc} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array}$$

2 **Model**: Asymmetric Continuous

Calculate *out-coreness* dan *in-coreness* based on **transaction nominal**

Bank	in-core	out-core
A	0,54	0,42
B	0,87	0,53
C	0,08	0,11
D	0,06	0,14
E	0,12	0,09

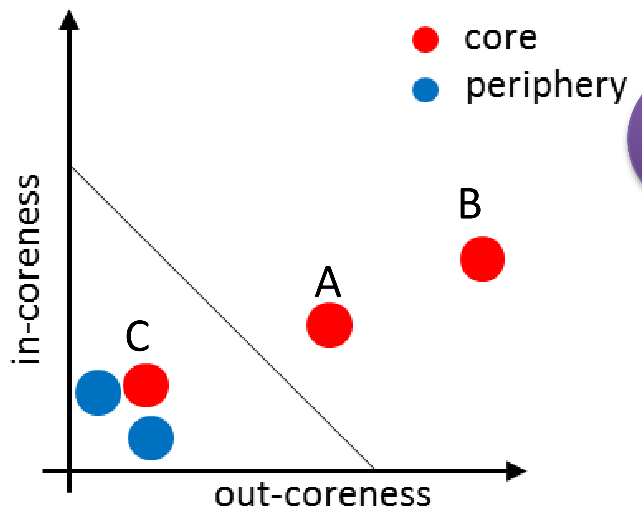
3 **Model**: Asymmetric Continuous

Calculate *out-coreness* dan *in-coreness* based on **transaction frequency**

Bank	in-core	out-core
A	0,75	0,52
B	0,67	0,87
C	0,12	0,01
D	0,08	0,02
E	0,11	0,15

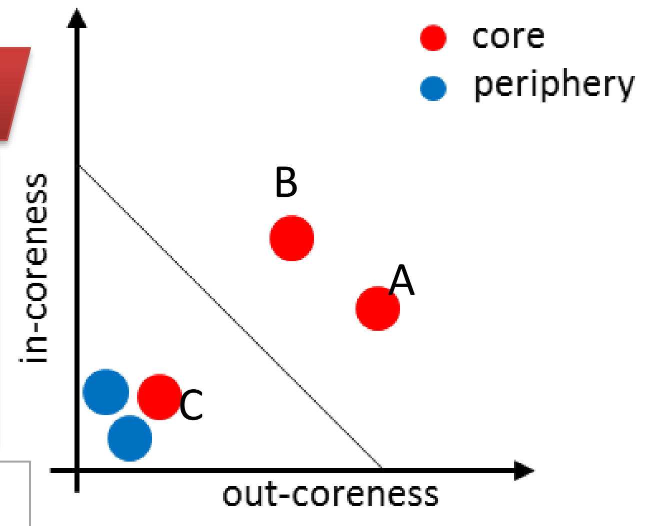
4 **Algoritma**: Support Vector Machine (SVM)

Identify Core & Periphery Bank based on interconnectedness and **transaction nominal**



5 **Algoritma**: Support Vector Machine (SVM)

Identify Bank Core & Periphery based on interconnectedness and **transaction frequency**

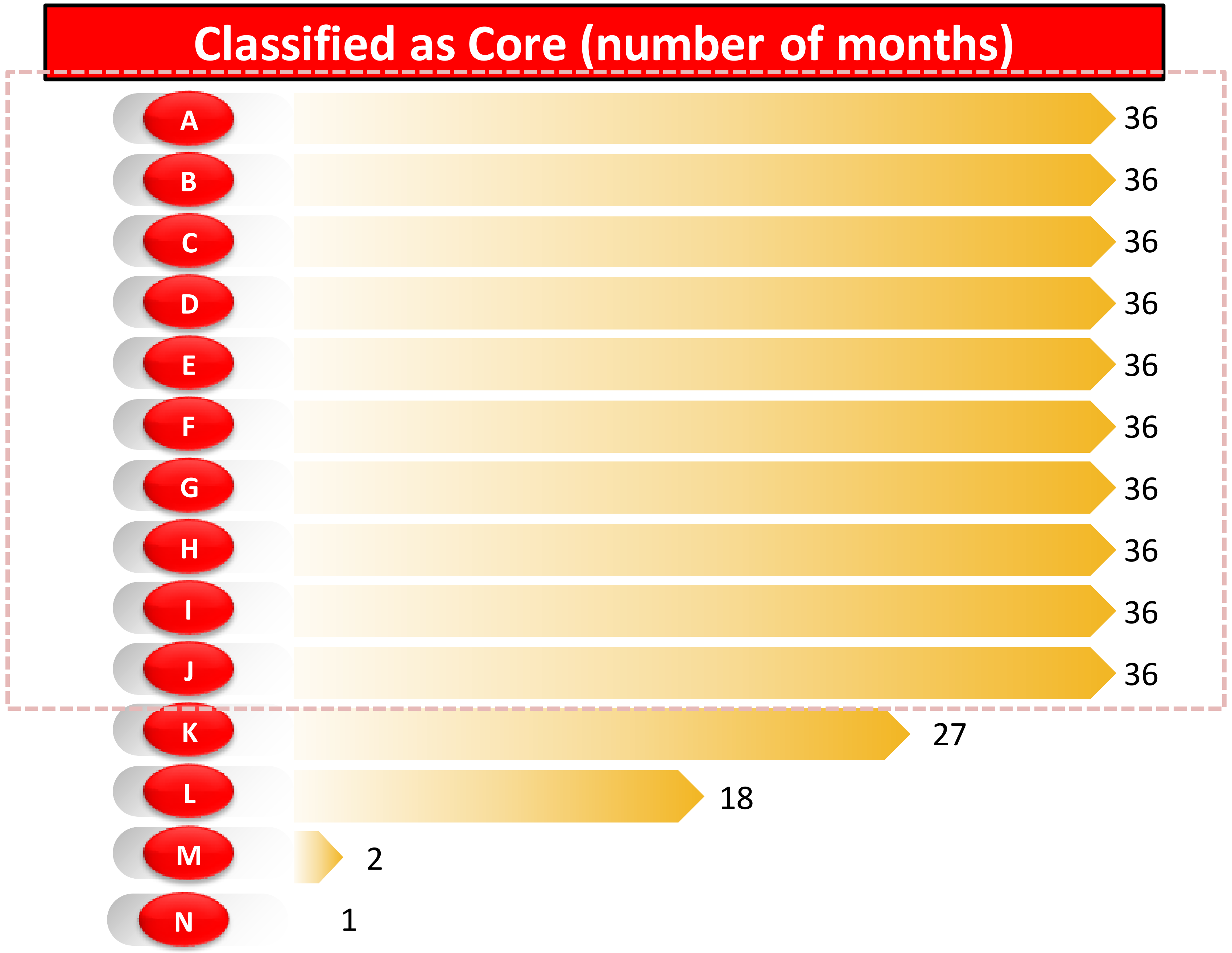


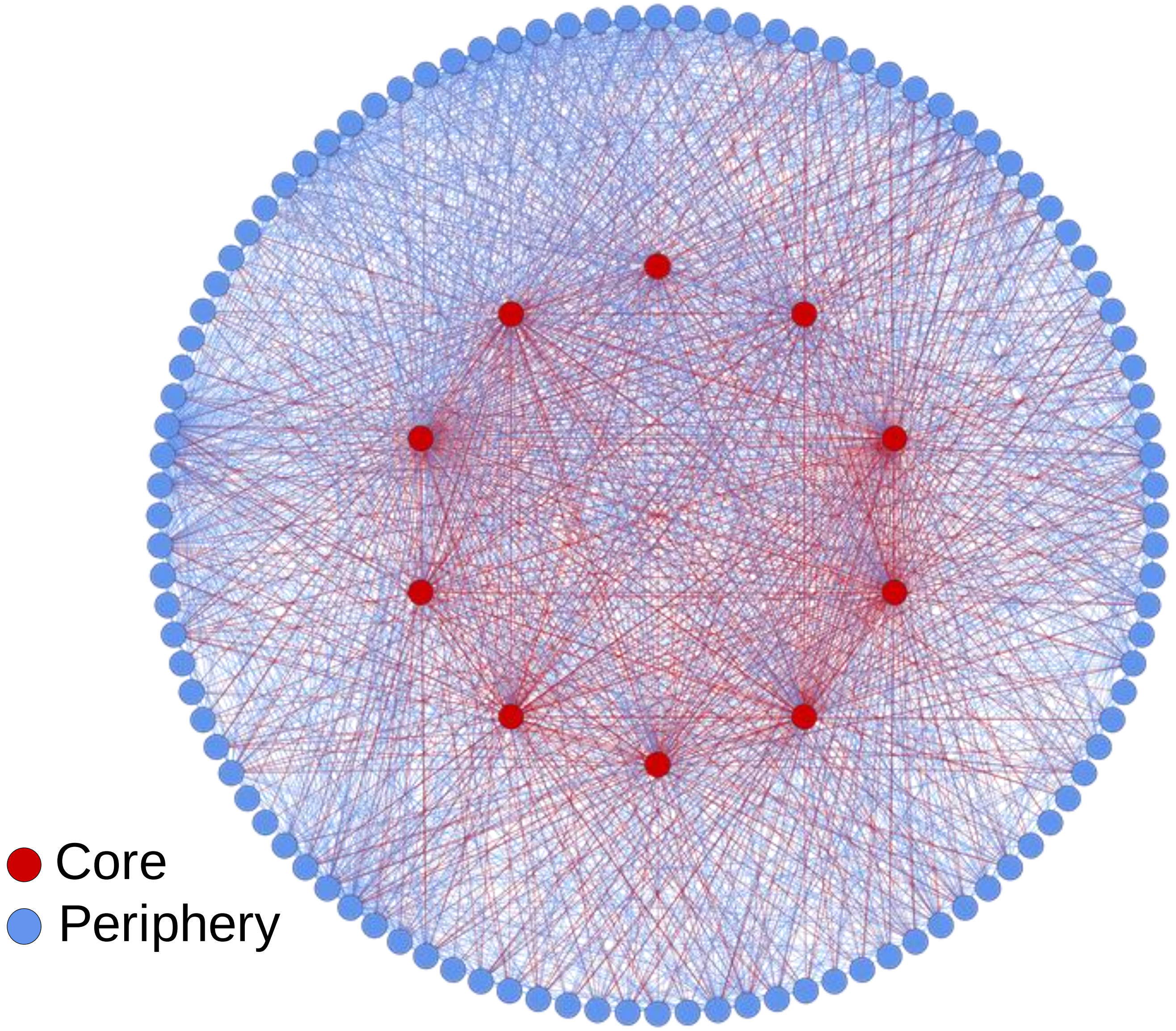
6 Identify Bank Core & Periphery based on the combination of **interconnectedness**, ***frequency*** and **nominal**

Bank A: *Core*
Bank B: *Core*
Bank C: *Periphery*
Bank D: *Periphery*
Bank E: *Periphery*

Core-Periphery Model Result

- The identification of the BI-RTGS network structure with the core-periphery model was conducted every month from January 2016 to December 2018 (36 months).
- Based on the results obtained, a core-periphery structure was identified within the BI-RTGS interbank transactions.
- This is evidenced by the presence of 10 banks that were identified as core banks (out of 117 banks that transacted with each other) throughout the observation period (36 months).

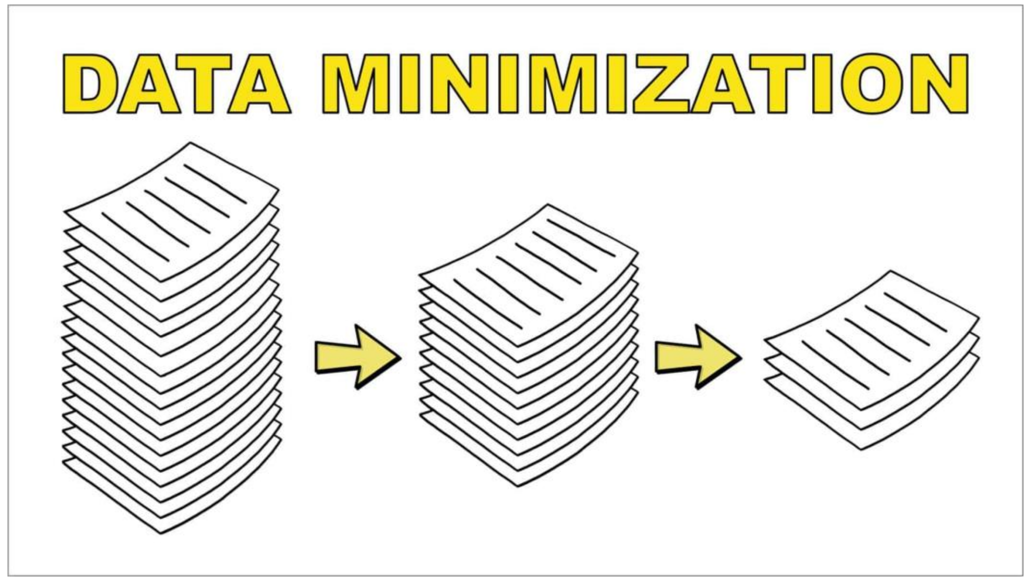


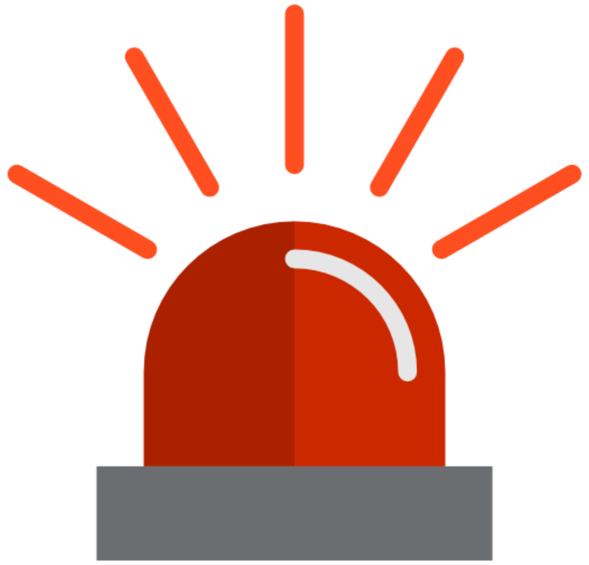


Struktur *Core-Periphery* in BI-RTGS Transaction

Data Governance Considerations

- **Data Minimization:** Only collect and process the minimum amount of data necessary to identify the core-periphery structure. In this case, we only use settlement time, sending/receiving bank, trx status, and amount. This reduces the risk of data breaches and ensures compliance with data protection regulations.
- **Bank Anonymization:** Before processing and analyzing the data, we ensure that all bank identifiers are anonymized. We replace bank names or any other identifiable information with generic IDs or pseudonyms. This ensures that individual banks cannot be directly identified from the data, protecting their privacy and ensuring compliance with data protection standards.
- **User Access Control:** Implement access controls to ensure that only authorized personnel can view and interpret the classification results. This not only ensures data privacy but also prevents potential misuse of the data.
- **Model Usage:** The core-periphery classification model should be used cautiously. For instance, it can be employed as an early warning system to highlight potential risks in the banking network. However, it should not be the sole source for making critical decisions. Always consider the ethical implications of classifying banks, especially if these classifications have significant financial or reputational implications for the banks involved.





Thank You