# The Banking Transactions Dataset and its Comparative Analysis with Scale-free Networks

Akrati Saxena

Co-authors: Yulong Pei, Jan Veldsink, Werner van Ipenburg, George Fletcher, Mykola Pechenizkiy

TU/e, Eindhoven, The Netherlands

March 25, 2022

# Table of Contents

# Dataset

This project is in collaboration with the Coöperatieve Rabobank U.A. (https://www.rabobank.nl/), i.e., a Dutch multinational banking and financial services company.

**Dataset Details:**

- This dataset consists of bank accounts and transactions between them.
- For any pairs of accounts with one or more transactions, we also collected more information, (i) the numbers of transactions between two accounts and (ii) the total amount of money transferred from one account to another over a period of 11 years from 2010 to 2020.
- The data was shared for 1,624,030 bank accounts and 4,127,043 transactions based on (from_account, to_account) pair.
- The dataset is available at request. As per the best of our knowledge, this is the first publicly available dataset of users' banking transactions.

# Dataset

This project is in collaboration with the Coöperatieve Rabobank U.A. (`https://www.rabobank.nl/`), i.e., a Dutch multinational banking and financial services company.

**Dataset Details:**

- This dataset consists of bank accounts and transactions between them.
- For any pairs of accounts with one or more transactions, we also collected more information, (i) the numbers of transactions between two accounts and (ii) the total amount of money transferred from one account to another over a period of 11 years from 2010 to 2020.
- The data was shared for 1,624,030 bank accounts and 4,127,043 transactions based on (from_account, to_account) pair.
- The dataset is available at request. As per the best of our knowledge, this is the first publicly available dataset of users' banking transactions.

# Dataset

This project is in collaboration with the Coöperatieve Rabobank U.A. (https://www.rabobank.nl/), i.e., a Dutch multinational banking and financial services company.

**Dataset Details:**

- This dataset consists of bank accounts and transactions between them.
- For any pairs of accounts with one or more transactions, we also collected more information, (i) the numbers of transactions between two accounts and (ii) the total amount of money transferred from one account to another over a period of 11 years from 2010 to 2020.
- The data was shared for 1,624,030 bank accounts and 4,127,043 transactions based on (from_account, to_account) pair.
- The dataset is available at request. As per the best of our knowledge, this is the first publicly available dataset of users' banking transactions.

# Dataset

This project is in collaboration with the Coöperatieve Rabobank U.A. (`https://www.rabobank.nl/`), i.e., a Dutch multinational banking and financial services company.

**Dataset Details:**

- This dataset consists of bank accounts and transactions between them.
- For any pairs of accounts with one or more transactions, we also collected more information, (i) the numbers of transactions between two accounts and (ii) the total amount of money transferred from one account to another over a period of 11 years from 2010 to 2020.
- The data was shared for 1,624,030 bank accounts and 4,127,043 transactions based on (from_account, to_account) pair.
- The dataset is available at request. As per the best of our knowledge, this is the first publicly available dataset of users' banking transactions.

# Networks

- We create the network from this data, having 1,624,030 nodes which are accounts, and 3,823,167 directed edges which represent that the respective users performed one or more transactions.
- Next, we identify the weakly connected components in the network, and it contains 723 connected components, where the largest weakly connected component contains 1,622,173 nodes and 3,821,514 edges. (The information of connected components is: 1 component has 1,622,173 nodes, 3 components have 27, 15, and 13 nodes, and the rest of the components have less than ten nodes.)
- Considering the transaction information, we create two weighted networks:
  1. $G^T$: edge-weight is the total amount of money transferred between two accounts
  2. $G^N$: edge-weight is the total number of transactions between two accounts

# Networks

- We create the network from this data, having 1,624,030 nodes which are accounts, and 3,823,167 directed edges which represent that the respective users performed one or more transactions.
- Next, we identify the weakly connected components in the network, and it contains 723 connected components, where the largest weakly connected component contains 1,622,173 nodes and 3,821,514 edges. (The information of connected components is: 1 component has 1,622,173 nodes, 3 components have 27, 15, and 13 nodes, and the rest of the components have less than ten nodes.)
- Considering the transaction information, we create two weighted networks:
  1. $G^T$: edge-weight is the total amount of money transferred between two accounts
  2. $G^N$: edge-weight is the total number of transactions between two accounts

# Networks

- We create the network from this data, having 1,624,030 nodes which are accounts, and 3,823,167 directed edges which represent that the respective users performed one or more transactions.
- Next, we identify the weakly connected components in the network, and it contains 723 connected components, where the largest weakly connected component contains 1,622,173 nodes and 3,821,514 edges. (The information of connected components is: 1 component has 1,622,173 nodes, 3 components have 27, 15, and 13 nodes, and the rest of the components have less than ten nodes.)
- Considering the transaction information, we create two weighted networks:
  1. $G^T$: edge-weight is the total amount of money transferred between two accounts
  2. $G^N$: edge-weight is the total number of transactions between two accounts

# Networks

- We create the network from this data, having 1,624,030 nodes which are accounts, and 3,823,167 directed edges which represent that the respective users performed one or more transactions.
- Next, we identify the weakly connected components in the network, and it contains 723 connected components, where the largest weakly connected component contains 1,622,173 nodes and 3,821,514 edges. (The information of connected components is: 1 component has 1,622,173 nodes, 3 components have 27, 15, and 13 nodes, and the rest of the components have less than ten nodes.)
- Considering the transaction information, we create two weighted networks:
  1. $G^T$: edge-weight is the total amount of money transferred between two accounts
  2. $G^N$: edge-weight is the total number of transactions between two accounts

# Networks

- We create the network from this data, having 1,624,030 nodes which are accounts, and 3,823,167 directed edges which represent that the respective users performed one or more transactions.
- Next, we identify the weakly connected components in the network, and it contains 723 connected components, where the largest weakly connected component contains 1,622,173 nodes and 3,821,514 edges. (The information of connected components is: 1 component has 1,622,173 nodes, 3 components have 27, 15, and 13 nodes, and the rest of the components have less than ten nodes.)
- Considering the transaction information, we create two weighted networks:
  1. $G^T$: edge-weight is the total amount of money transferred between two accounts
  2. $G^N$: edge-weight is the total number of transactions between two accounts

# Introduction

- We follow the 'cookbook approach' for performing a systematic analysis of basic and advanced network characteristics, as followed in previous works[1].
- The analysis of bank transaction network will be beneficial for several research directions.
  - It will help in understanding the flow of money at the microscopic level as well as how this contributes towards the macroscopic money transaction system.
  - It will improve the downstream tasks in the financial domain with guidance from the topological perspective. Some representative tasks include fraud detection and user classification in transaction networks.
  - It will shed light on financial simulator design. Existing simulators focus on individual customer behavior, while our analysis can provide complementary information about collective behaviors of group users.

[1] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Introduction

- We follow the 'cookbook approach' for performing a systematic analysis of basic and advanced network characteristics, as followed in previous works[1].
- The analysis of bank transaction network will be beneficial for several research directions.
  - It will help in understanding the flow of money at the microscopic level as well as how this contributes towards the macroscopic money transaction system.
  - It will improve the downstream tasks in the financial domain with guidance from the topological perspective. Some representative tasks include fraud detection and user classification in transaction networks.
  - It will shed light on financial simulator design. Existing simulators focus on individual customer behavior, while our analysis can provide complementary information about collective behaviors of group users.

[1] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Introduction

- We follow the 'cookbook approach' for performing a systematic analysis of basic and advanced network characteristics, as followed in previous works[1].
- The analysis of bank transaction network will be beneficial for several research directions.
  - It will help in understanding the flow of money at the microscopic level as well as how this contributes towards the macroscopic money transaction system.
  - It will improve the downstream tasks in the financial domain with guidance from the topological perspective. Some representative tasks include fraud detection and user classification in transaction networks.
  - It will shed light on financial simulator design. Existing simulators focus on individual customer behavior, while our analysis can provide complementary information about collective behaviors of group users.

[1] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Introduction

- We follow the 'cookbook approach' for performing a systematic analysis of basic and advanced network characteristics, as followed in previous works[1].
- The analysis of bank transaction network will be beneficial for several research directions.
  - It will help in understanding the flow of money at the microscopic level as well as how this contributes towards the macroscopic money transaction system.
  - It will improve the downstream tasks in the financial domain with guidance from the topological perspective. Some representative tasks include fraud detection and user classification in transaction networks.
  - It will shed light on financial simulator design. Existing simulators focus on individual customer behavior, while our analysis can provide complementary information about collective behaviors of group users.

[1] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Introduction

- We follow the 'cookbook approach' for performing a systematic analysis of basic and advanced network characteristics, as followed in previous works[1].
- The analysis of bank transaction network will be beneficial for several research directions.
  - It will help in understanding the flow of money at the microscopic level as well as how this contributes towards the macroscopic money transaction system.
  - It will improve the downstream tasks in the financial domain with guidance from the topological perspective. Some representative tasks include fraud detection and user classification in transaction networks.
  - It will shed light on financial simulator design. Existing simulators focus on individual customer behavior, while our analysis can provide complementary information about collective behaviors of group users.

[1] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Basic Network Characteristics

# Sampled Subgraph

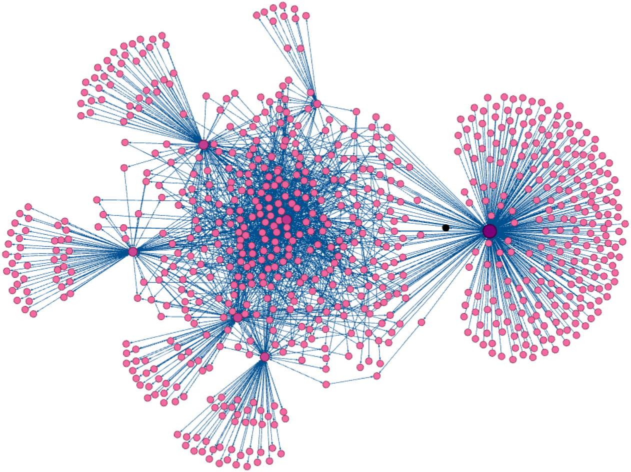


Figure: A sample subgraph extracted from the black node (chosen uniformly at random) till distance $l = 3$.

# l-distant nodes

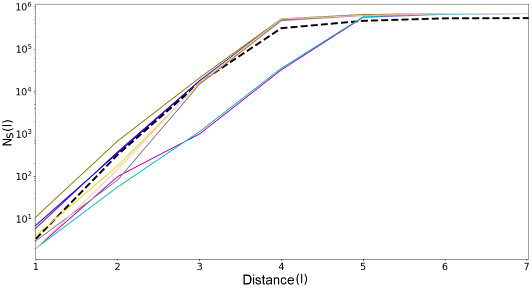
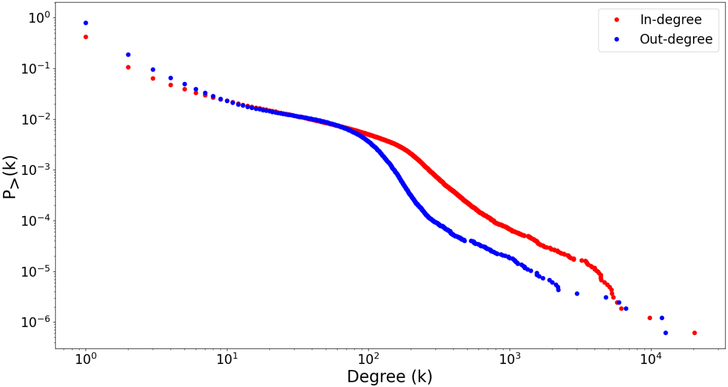



Figure: Average Number of nodes ($N_s(l)$) at distance $l$ as a function of distance ($l$) for uniformly sampled 0.1% source nodes (dashed black lines) and the number of nodes at distance $l$ for some random nodes (solid thin lines).

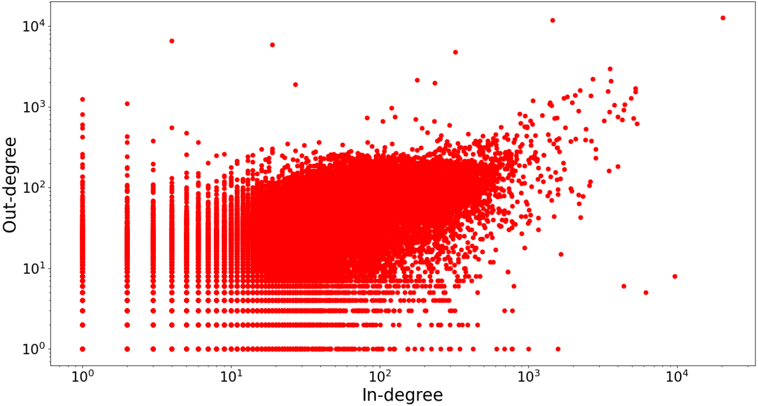
To a good approximation, the curve follows the Boltzman equation $B + (A - B)/(1 + (x/x_0)^p)$, where $A$, $B$, $x_0$, and $p$ are the parameters.

# Cumulative In-degree and out-degree distribution




The cumulative in-degree and out-degree distribution follows the power-law equation $(P(k) = c * (k)^{(-\gamma)})$, i.e., similar to other scale-free networks.

# Out-degree versus In-degree for WLCC




The plot shows that the out-degree is not correlated with in-degree and has the Spearman correlation coefficient $-0.15$.

# Cumulative edge-weight distribution

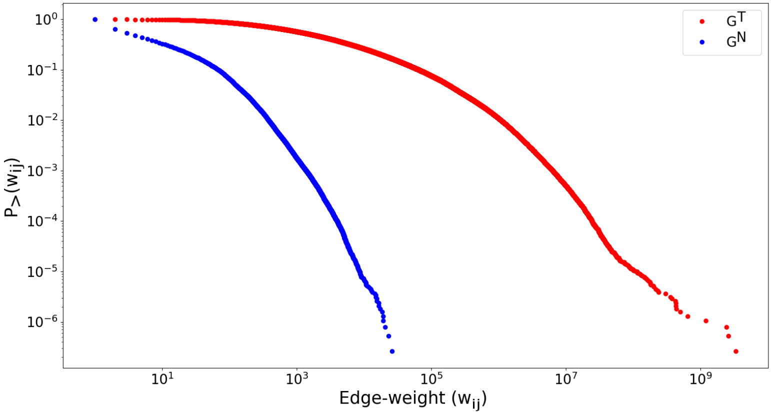



Figure: Cumulative edge-weight distribution for $G^T$ and $G^N$.

# Correlation of edge-weights

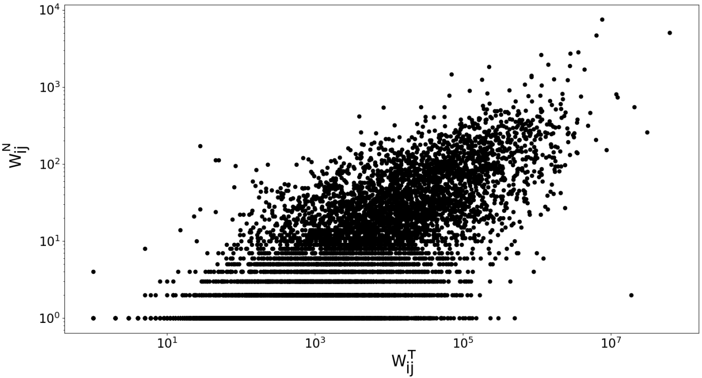
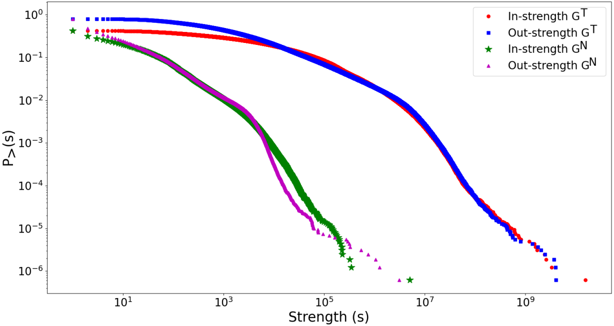



Figure: Correlation of edge-weights for randomly sampled 10000 edges in two networks $G^T$ and $G^N$. The two weights are clearly correlated having Spearman's coefficient 0.71.

# Cumulative In-strength and Out-strength distribution




As expected, both distributions follow the power law. The distributions of in-strength and out-strength are similar because in the process of data collection, we filter out special accounts, such as gas stations, that have much higher in-degrees than out-degrees.

# Out-strength versus In-strength

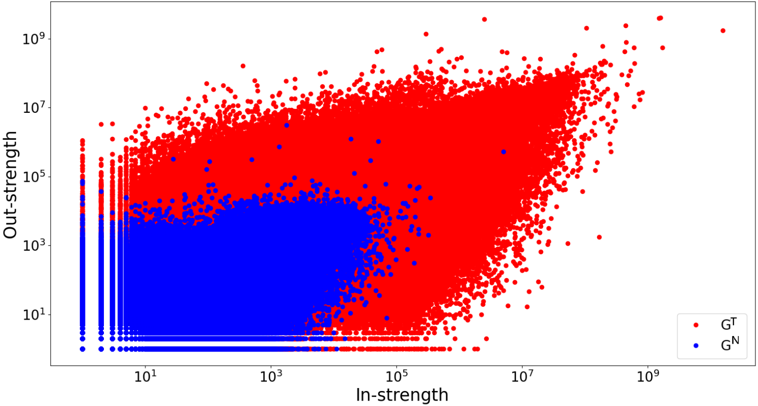
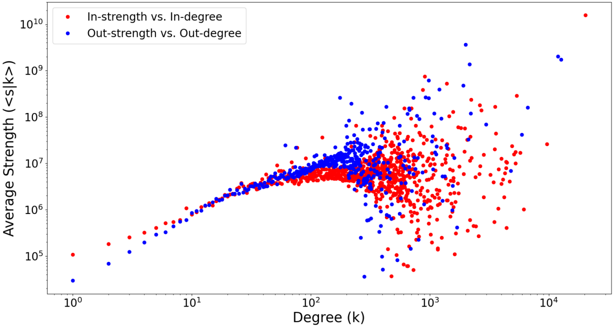



Figure: Out-strength versus In-strength for $G^T$ network in red color and for $G^N$ network in blue color.

# Average Strength vs. Degree

Average In-strength vs. In-degree and average out-strength vs. out-degree for $G^T$.




We observe that the average in/out-strength increases with in/out-degree till a certain range ($\sim 100$), and after that, no clear correlation is observed in $G^T$ network. (Let's discuss its reason)

# Average Strength vs. Degree

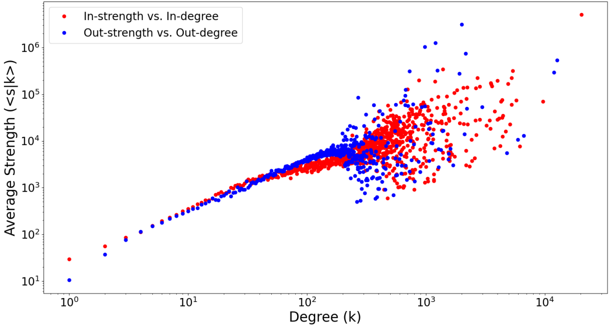



Figure: Average In-strength vs. In-degree and average out-strength vs. out-degree for $G^N$

The average in/out-strength increases with in/out-degree in $G^N$ network and has a higher correlation.

# Neighborhood Analysis

# Assortativity Analysis

| Assortativity | Value |
|---|---|
| In-degree Assortativity | -0.018 |
| Out-degree Assortativity | -0.011 |
| In-strength Assortativity $G^T$ | -0.004 |
| Out-strength Assortativity $G^T$ | -0.005 |
| In-strength Assortativity $G^N$ | -0.005 |
| Out-strength Assortativity $G^N$ | -0.003 |

Neither assortativity nor dis-assortativity is observed.

These results are not in correlation with the assortativity observed in other scale-free networks, such as phone call based communication network[2] or the Internet network[3].

[2] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

[3] Romualdo Pastor-Satorras, Alexei V´azquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. Physical review letters, 87(25):258701, 2001.

# Neighborhood Analysis

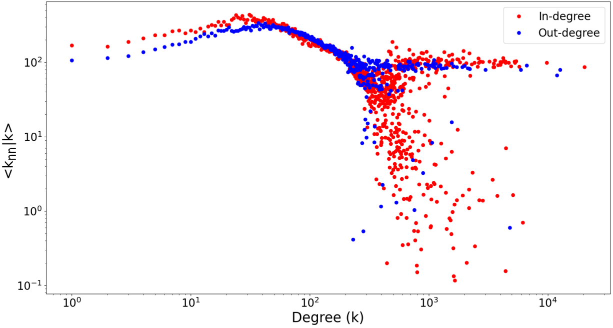



Figure: Average neighbor In-degree versus In-degree and average neighbor out-degree versus out-degree.

We observe that the most of the users have a high average neighbor degree irrespective of in/out-degree of a user.

# Neighborhood Analysis

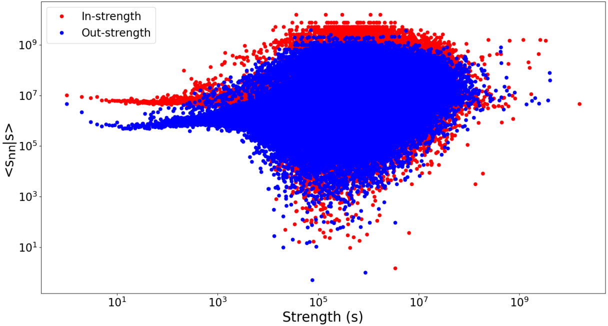



Figure: Average Neighbor In-strength versus In-strength and average neighbor out-strength versus out-strength $G^T$

# Neighborhood Analysis

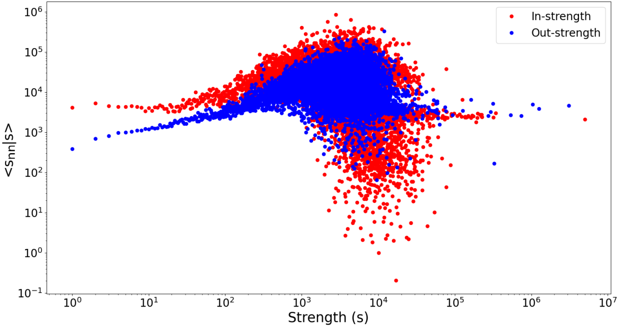
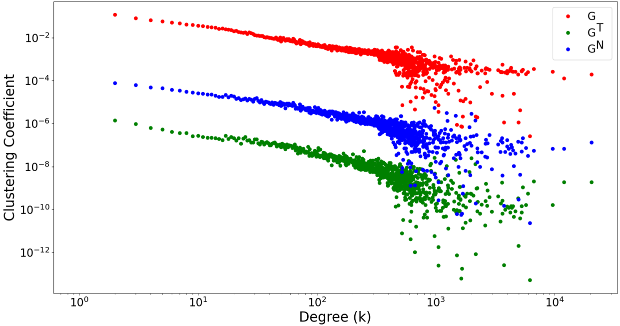



Figure: Average Neighbor In-strength versus In-strength and average neighbor out-strength versus out-strength $G^N$

No correlation is observed due to the basic network characteristics (as explained for Degree).

# Clustering Coefficient Analysis




The clustering coefficient follows the same pattern as observed in other large-scale scale-free unweighted and weighted networks[4].

[4] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. Quantitative Finance, 10(4):389–399, 2010.

# Meso-scale Network Characteristics

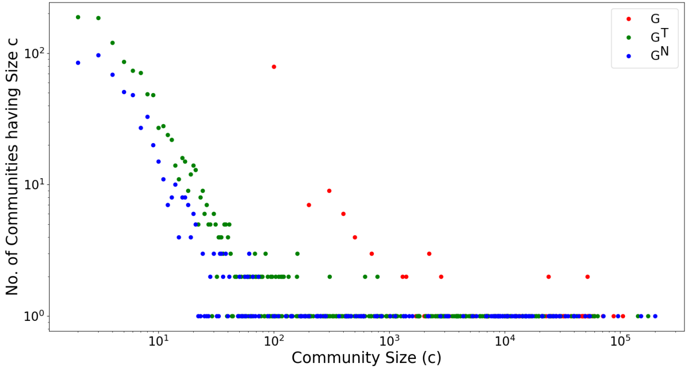
# Clique Analysis

| Order | Empirical Count | ER Expectation |
|---|---|---|
| 1 | 1622173 | 1622173 |
| 2 | 3318903 | 3318903 |
| 3 | 417461 | 11.42 |
| 4 | 34638 | $7.43 \times 10^{-11}$ |
| 5 | 3815 | $9.76 \times 10^{-28}$ |
| 6 | 417 | $2.70 \times 10^{-50}$ |
| 7 | 26 | $1.61 \times 10^{-78}$ |
| 8 | 0 | $1.12 \times 10^{-112}$ |

Table: Number of cliques of order $k = 1, 2, ..., 8$ in the undirected Rabobank network (empirical count) and their expectation values in a corresponding ER network (ER expectation).

# Community Analysis

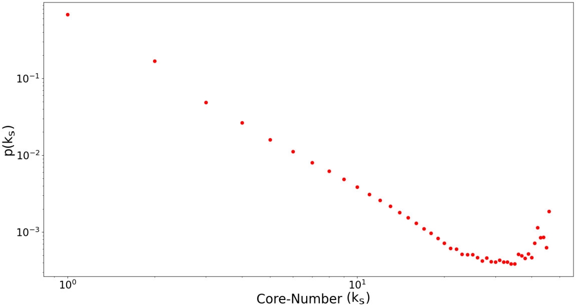
We use the Leiden community detection method[5] that is an extension of the Louvain community detection method to identify well-connected communities in directed weighted/unweighted networks.




[5] Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. From louvain to leiden: guaranteeing well-connected communities. Scientific reports, 9(1):1–12, 2019.

# Core-Periphery Analysis

- Apply K-shell decomposition method[6].
- In network $G$, the highest shell-index is 46 shared by 2,871 nodes that is a very small fraction of all nodes. The results are similar to as observed in other scale-free networks[7].




[6] Stephen B Seidman. Network structure and minimum degree. Social networks, 5(3):269–287, 1983.

[7] Akrati Saxena and SRS Iyengar. Evolving models for meso-scale structures. In 2016 8th International Conference on Communication Systems and Networks (COMSNETS), pages 1–8. IEEE, 2016.

# Transaction network vs. Scale-Free Networks

# Transaction network vs. Scale-Free Networks

The banking transaction network follows the characteristics of scale-free networks; still, it has some clear similarities and differences with other scale-free real-world networks.

- The degree distribution, strength distribution, and edge-weight follow power-law as observed in other social, information, biological, or technological networks[8].
- However, there is no correlation between in-degree vs. out-degree, and in-strength vs. out-strength as observed in other Information networks, such as WWW[9].
- In social and Information (transportation) networks, the strength of nodes increases with degree[10]; however, the banking transaction network has different pattern due to the nature of the network evolution.

[8] Newman, M. E., Barabási, A. L. E., Watts, D. J. (2006). The structure and dynamics of networks. Princeton university press.

[9] Yihong Hu and Daoli Zhu. Empirical analysis of the worldwide maritime transportation network. Physica A: Statistical Mechanics and its Applications, 388(10):2061–2071, 2009.

[10] Barrat, A., Barthélemy, M., Vespignani, A. (2004). Modeling the evolution of weighted networks. Physical review E, 70(6), 066149.

# Transaction network vs. Scale-Free Networks

The banking transaction network follows the characteristics of scale-free networks; still, it has some clear similarities and differences with other scale-free real-world networks.

- The degree distribution, strength distribution, and edge-weight follow power-law as observed in other social, information, biological, or technological networks[8].
- However, there is no correlation between in-degree vs. out-degree, and in-strength vs. out-strength as observed in other Information networks, such as WWW[9].
- In social and Information (transportation) networks, the strength of nodes increases with degree[10]; however, the banking transaction network has different pattern due to the nature of the network evolution.

[8] Newman, M. E., Barabási, A. L. E., Watts, D. J. (2006). The structure and dynamics of networks. Princeton university press.

[9] Yihong Hu and Daoli Zhu. Empirical analysis of the worldwide maritime transportation network. Physica A: Statistical Mechanics and its Applications, 388(10):2061–2071, 2009.

[10] Barrat, A., Barthélemy, M., Vespignani, A. (2004). Modeling the evolution of weighted networks. Physical review E, 70(6), 066149.

# Transaction network vs. Scale-Free Networks

The banking transaction network follows the characteristics of scale-free networks; still, it has some clear similarities and differences with other scale-free real-world networks.

- The degree distribution, strength distribution, and edge-weight follow power-law as observed in other social, information, biological, or technological networks[8].
- However, there is no correlation between in-degree vs. out-degree, and in-strength vs. out-strength as observed in other Information networks, such as WWW[9].
- In social and Information (transportation) networks, the strength of nodes increases with degree[10]; however, the banking transaction network has different pattern due to the nature of the network evolution.

[8] Newman, M. E., Barabási, A. L. E., Watts, D. J. (2006). The structure and dynamics of networks. Princeton university press.

[9] Yihong Hu and Daoli Zhu. Empirical analysis of the worldwide maritime transportation network. Physica A: Statistical Mechanics and its Applications, 388(10):2061–2071, 2009.

[10] Barrat, A., Barthélemy, M., Vespignani, A. (2004). Modeling the evolution of weighted networks. Physical review E, 70(6), 066149.

# Transaction network vs. Scale-Free Networks cont...

- The analysis shows that neighborhood connectivity is different from other kinds of networks.
- The network is neither assortative nor disassortative, as we observe in most of the other scale-free networks including trade and finance networks[11,12]; therefore, no correlation with the neighborhood nodes' degree is observed.
- The evolution of the network is not random and regulated by an underlying evolving mechanism.
- The network has community, and hierarchical structure[13,14] as observed in other scale-free networks.

[11] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. Quantitative Finance, 10(4):389–399, 2010.

[12] Mark EJ Newman. Assortative mixing in networks. Physical review letters, 89(20):208701, 2002.

[13] Alex Arenas, Leon Danon, Albert Diaz-Guilera, Pablo M Gleiser, and Roger Guimera. Community analysis in social networks. The European Physical Journal B, 38(2):373–380, 2004.

[14] Akrati Saxena and SRS Iyengar. Evolving models for meso-scale structures. In 2016 8th International Conference on Communication Systems and Networks (COMSNETS), pages 1–8. IEEE, 2016.

# Transaction network vs. Scale-Free Networks cont...

- The analysis shows that neighborhood connectivity is different from other kinds of networks.
- The network is neither assortative nor disassortative, as we observe in most of the other scale-free networks including trade and finance networks[11][12]; therefore, no correlation with the neighborhood nodes' degree is observed.
- The evolution of the network is not random and regulated by an underlying evolving mechanism.
- The network has community, and hierarchical structure[13][14] as observed in other scale-free networks.

[11] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. Quantitative Finance, 10(4):389–399, 2010.

[12] Mark EJ Newman. Assortative mixing in networks. Physical review letters, 89(20):208701, 2002.

[13] Alex Arenas, Leon Danon, Albert Diaz-Guilera, Pablo M Gleiser, and Roger Guimera. Community analysis in social networks. The European Physical Journal B, 38(2):373–380, 2004.

[14] Akrati Saxena and SRS Iyengar. Evolving models for meso-scale structures. In 2016 8th International Conference on Communication Systems and Networks (COMSNETS), pages 1–8. IEEE, 2016.

# Transaction network vs. Scale-Free Networks cont...

- The analysis shows that neighborhood connectivity is different from other kinds of networks.
- The network is neither assortative nor disassortative, as we observe in most of the other scale-free networks including trade and finance networks[11][12]; therefore, no correlation with the neighborhood nodes' degree is observed.
- The evolution of the network is not random and regulated by an underlying evolving mechanism.
- The network has community, and hierarchical structure[13][14] as observed in other scale-free networks.

[11] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. Quantitative Finance, 10(4):389–399, 2010.

[12] Mark EJ Newman. Assortative mixing in networks. Physical review letters, 89(20):208701, 2002.

[13] Alex Arenas, Leon Danon, Albert Diaz-Guilera, Pablo M Gleiser, and Roger Guimera. Community analysis in social networks. The European Physical Journal B, 38(2):373–380, 2004.

[14] Akrati Saxena and SRS Iyengar. Evolving models for meso-scale structures. In 2016 8th International Conference on Communication Systems and Networks (COMSNETS), pages 1–8. IEEE, 2016.

# Transaction network vs. Scale-Free Networks cont...

- The analysis shows that neighborhood connectivity is different from other kinds of networks.
- The network is neither assortative nor disassortative, as we observe in most of the other scale-free networks including trade and finance networks[11,12]; therefore, no correlation with the neighborhood nodes' degree is observed.
- The evolution of the network is not random and regulated by an underlying evolving mechanism.
- The network has community, and hierarchical structure[13,14] as observed in other scale-free networks.

---

[11] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. International trade and financial integration: a weighted network analysis. Quantitative Finance, 10(4):389–399, 2010.

[12] Mark EJ Newman. Assortative mixing in networks. Physical review letters, 89(20):208701, 2002.

[13] Alex Arenas, Leon Danon, Albert Diaz-Guilera, Pablo M Gleiser, and Roger Guimera. Community analysis in social networks. The European Physical Journal B, 38(2):373–380, 2004.

[14] Akrati Saxena and SRS Iyengar. Evolving models for meso-scale structures. In 2016 8th International Conference on Communication Systems and Networks (COMSNETS), pages 1–8. IEEE, 2016.

# Transaction network vs. Scale-Free Networks cont...

- One another main difference that we observed is the correlation of weak ties with edge-weights.
- In social networks, the weak ties, or also referred to as bridges, are the connections between communities and therefore have a lower edge-weight as the social interaction between people belonging to two different communities is not very strong[15]. However, in the banking transaction network, no such correlation is observed. The transactions that happened between users belonging to different communities might have lower or higher strength based on the transferred amount and the total number of transactions.

[15] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Transaction network vs. Scale-Free Networks cont...

- One another main difference that we observed is the correlation of weak ties with edge-weights.
- In social networks, the weak ties, or also referred to as bridges, are the connections between communities and therefore have a lower edge-weight as the social interaction between people belonging to two different communities is not very strong[15]. However, in the banking transaction network, no such correlation is observed. The transactions that happened between users belonging to different communities might have lower or higher strength based on the transferred amount and the total number of transactions.

[15] Onnela, Jukka-Pekka, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M. Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. "Analysis of a large-scale weighted network of one-to-one human communication." New journal of physics 9, no. 6 (2007): 179.

# Conclusion

- We perform the network analysis from bank transaction records of Rabobank where edge-weights are assigned using the aggregated amount of transactions and the total number of transactions from 2010 to 2020.
- To our knowledge, it is the first intra-bank transaction network studied so far and will be the first bank transaction dataset that is publicly available.
- The analysis of k-cliques and reciprocity shows that the network evolution is not random.
- The network's community structure shows that the users are organized in smaller groups making more transactions between them and fewer transactions outside the group.

# Conclusion

- We perform the network analysis from bank transaction records of Rabobank where edge-weights are assigned using the aggregated amount of transactions and the total number of transactions from 2010 to 2020.
- To our knowledge, it is the first intra-bank transaction network studied so far and will be the first bank transaction dataset that is publicly available.
- The analysis of k-cliques and reciprocity shows that the network evolution is not random.
- The network's community structure shows that the users are organized in smaller groups making more transactions between them and fewer transactions outside the group.

# Conclusion

- We perform the network analysis from bank transaction records of Rabobank where edge-weights are assigned using the aggregated amount of transactions and the total number of transactions from 2010 to 2020.
- To our knowledge, it is the first intra-bank transaction network studied so far and will be the first bank transaction dataset that is publicly available.
- The analysis of k-cliques and reciprocity shows that the network evolution is not random.
- The network's community structure shows that the users are organized in smaller groups making more transactions between them and fewer transactions outside the group.

# Conclusion

- We perform the network analysis from bank transaction records of Rabobank where edge-weights are assigned using the aggregated amount of transactions and the total number of transactions from 2010 to 2020.
- To our knowledge, it is the first intra-bank transaction network studied so far and will be the first bank transaction dataset that is publicly available.
- The analysis of k-cliques and reciprocity shows that the network evolution is not random.
- The network's community structure shows that the users are organized in smaller groups making more transactions between them and fewer transactions outside the group.

# On-going Work

- Labeled Data (Anomalous and Non-anomalous Users)
- How to identify anomalous users?
- Open Questions: The evolving model for banking transaction networks is still an open question, and the above statistical observations mentioning the similarities and differences will help pin down the underlying evolving mechanism.

# On-going Work

- Labeled Data (Anomalous and Non-anomalous Users)
- How to identify anomalous users?
- Open Questions: The evolving model for banking transaction networks is still an open question, and the above statistical observations mentioning the similarities and differences will help pin down the underlying evolving mechanism.

# On-going Work

- Labeled Data (Anomalous and Non-anomalous Users)
- How to identify anomalous users?
- Open Questions: The evolving model for banking transaction networks is still an open question, and the above statistical observations mentioning the similarities and differences will help pin down the underlying evolving mechanism.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Other Interesting Projects

- Fairness in Network Science
  - Fairness-aware link prediction in Social Networks[16][17]
  - Fairness-aware Influence Maximization and Influence Blocking
- Fairness-aware methods in Data Mining
  - Fair Automated Essay Scoring System
  - Fair Batch-Mode Active Learning for Algorithmic Decision Making
  - Fairness Approaches in Reinforcement Learning

[16] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2022). NodeSim: Node Similarity based Network Embedding for Diverse Link Prediction. (Accepted in EPJ Data Science Journal, 2022).

[17] Saxena, A., Fletcher, G., & Pechenizkiy, M. (2021). HM-EIICT: Fairness-aware link prediction in complex networks using community information. Journal of Combinatorial Optimization, 1-18.

# Thank You

Contact: a.saxena@tue.nl
Homepage: www.akratisaxena.com