Network relationships and network models in payment systems

Kimmo Soramäki (ECB)
Morten Bech and Jeffrey Arnold (FRBNY)
Walter Beyeler and Robert J. Glass
(Sandia National Laboratories)

Bank of Finland, Helsinki, 24 August 2005
Introduction

• any network can be modeled as a graph with nodes and links between the nodes

• this presentation introduces concepts from graph theory and network science and applies them to describe liquidity flows in payment systems

• it is one of three components in our approach for modelling payment systems: Complex network, Complex behavior, and Adaptation
What can the results be used for?

- to better understand the topology of liquidity flows in payment systems
- to better understand the spill-over effects of liquidity disturbances
- for identifying important banks
- to possibly devise financial fragility indicators on the basis of the topology
- to analyze long-term structural changes and spot the impact of abnormal events
- to generate artificial data
  - we can compare the model’s statistical properties with the one of the actual system
Some terminology

- Graphs are made up by **nodes** and **links** between the nodes.
- Links can be either **directed**, or **undirected**.
- Links can have **weights**.
- A **path** is a sequence of nodes in which each node is linked to the next one (e.g. EDCA is a path of length 3).
- The **degree** of a node is the number of links from (out-degree) or to it (in-degree) from other nodes.
- A **cluster** is a set of nodes that all have links with each other (e.g. ABC).
Basic network models I: Erdos-Renyi

- Erdos-Renyi model (50’s)
- classical random network
- start with N nodes, and connect pairs at random until desired connectivity is reached

The ER model degree distribution follows the Poisson distribution
Basic network models II: Barabasi-Albert

- start with a small number of nodes
- **growth**: at each step add a node and link it to one existing node
- **preferential attachment**: nodes with a higher number of incoming links have a higher likelihood of being selected
- continue until desired number of nodes have been created

The BA model degree distribution follows the power law
Interbank payment networks

• how to define the network depends on the question one wishes to study

Options:

• which payment system participants to include?
  – all, commercial banks, settlement institutions, ...

• what kind of interaction?
  – a payment, exchange of payments, a debt relationship, ...

• how long do we observe the formation of the network?
  – an hour, a day, a week, ...

• how intense should the interaction be
  – certain number, certain value of payments, ...
Network fundamentals used

• we use payment data from Fedwire to illustrate liquidity flows among banks in a payment system
• other large-scale payment systems are likely to exhibit the same properties

• in particular, we build
  – daily networks. If one or more payments are transacted from a bank to another, we establish a directed link from the bank to the other

• we consider only
  – payments between commercial banks,
  – that are not related to overnight funding
Basic statistics

averages for 62 daily networks:

- banks (n) 5,086
- links (l) 85,585
  - possible links, n*(n-1) 25,862,310
  - connectivity, l / n*(n-1) 0.3%
  - reciprocity 21.5%
    (share of two-way links)
- value 1,302 billion
- number 435,533
  - average payment size 3.0 million
Visualising the network

- example random scale-free network of 100 nodes and 680 links
- similar topology as in the liquidity flows
- small core with high flows (red lines)
- large periphery with low flows (black lines)
- visualising larger networks difficult
Components of the created network

GWCC: Giant weakly connected component. All banks in this component can be reached via undirected links from each other.

GSCC: Giant strongly connected component. The core of the network. All banks reachable from any other bank.

GIN: All banks that can reach the GSCC

GOUT: All banks reachable from GSCC

Tendrils: Banks that are not reachable nor reach the GSCC
Degree Distribution

what kind of hierarchy does the network have?

Out Degree Distribution
Fedwire Banks: Directed-Strong Gross Largest Component
1st Quarter 2004

- Log10 of Degree(k)
- Log10 of P(k)

probability of degree 1 ~ 13%
probability of degree 10 ~ 2%
probability of degree 100 ~ 0.02%
average degree = 15

Empirical
ER model (p=0.0015)
BA model (m=15)
The power law distribution

• the slope of the distribution is defined by the co-efficient $\gamma$, $P(k) \sim k^{-\gamma}$
  – for our network $\gamma = 2.1$

• examples of other networks
  – internet, router level $\gamma = 2.4$
  – movie actor collaboration network $\gamma = 2.3$
  – co-authorship network of physicists $\gamma = 2.5$
  – co-authorship network of neuroscientist $\gamma = 2.1$

• networks with a power law degree distribution are called scale-free

• sometimes said to be the new ”normal distribution” as the distribution of many man-made and natural events have this distribution
  – the size of earthquakes, stock market movements, ...
Number of payments on a link

Probability of 1 payment ~ 60%

Distribution of Link Weights (Number of Payments)
Fedwire Banks: Directed-Strong Gross Largest Component
1st Quarter 2004

Average = 5.1

Power law distribution,
$P(n) \sim n^{-2}$

Probability of 10 payments ~ 0.4%

Probability of 100 payments ~ 0.005%
Value of payments on a link

Distribution of Edge Weights (Value)
Fedwire Banks: Directed-Strong Gross Largest Component
1st Quarter 2004

Note: The curve is a normal distribution

value of payments transacted on a link follows a lognormal distribution

average = 15.2 mil.
Average Path Length

Histogram of Average Path Length

Note: The curve is a normal distribution

Mass-Distance Function
Fedwire Banks: Directed-Strong Gross Largest Component
1st Quarter 2004

average path length  2.6
diameter (maximum)  7
Clustering

how are banks connected locally?

Histogram of Clustering Coefficient

many real networks exhibit a high degree of clustering
- internet 0.2-0.3
- co-authorship 0.6-0.8

vs Erdos-Renyi 0.006

Example:

average clustering coefficient = 0.53
Relevance of the numbers?

- We don’t know yet. Some hypotheses:
  - Degree distribution and link weights (power law)
    - most banks irrelevant from financial stability perspective
    - hubs and bridges matter
  - Average path length
    - might be relevant for e.g. gridlocks RTGS systems. The smaller the APL, the quicker a liquidity shortage would spill over to other banks
  - Clustering co-efficient
    - might be relevant in contagion of netting systems, and in liquidity problems when exposures are reinforced by the neighbours
Summary

• payment systems are just one of many similar networks
• the statistics presented here are just scratch the surface
• many questions ahead:
  – what drives the topology?
  – how does the topology relate to liquidity disturbances?
  – what is the topology of liquidity flows in other payment systems?
  – how to best describe the importance of a bank? (PageRank, betweenness centrality, etc)
  – other network statistics? (loops, communities, etc)
auxiliary slides
Other scale-free networks ...

Sources:
Average nearest neighbour degree

average degree of neighbours, i.e. banks to whom one has links to

example:

degree

15

ANND = (15+5+10+50) / 4 = 20

Average Nearest Neighbor Out Degree
Fedwire Banks: Directed-Strong Gross Largest Component
1st Quarter 2004

• average = 617
• compare: average degree = 15

most banks have a very small degree, and connect to banks with high degree
both the number and value of payments on a link increase with the degree of the bank