Bank-firm credit network in Japan. A bipartite analysis and the characterization and time evolution of clusters of credit

L. Marotta, S. Micciché, Y. Fujiwara, H. Iyetomi, H. Aoyama, M. Gallegati, R. N. Mantegna

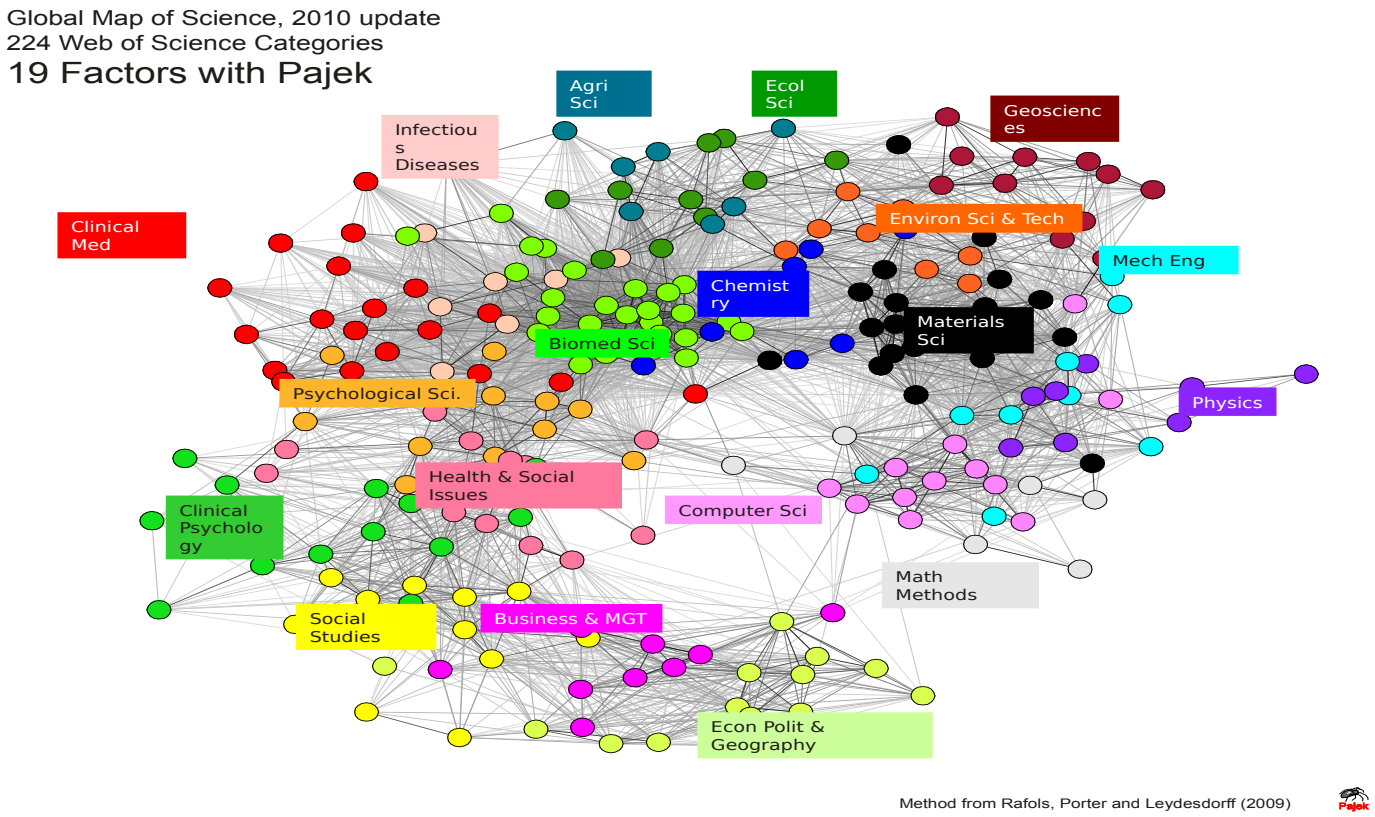
INET WORKSHOP, Ancona, July 5 2013

http::/ocs.unipa.it

Many tanks to:

A praise of Interdisciplinarity

Global science map based on citing similarities among ISI Subject Categories (2010)

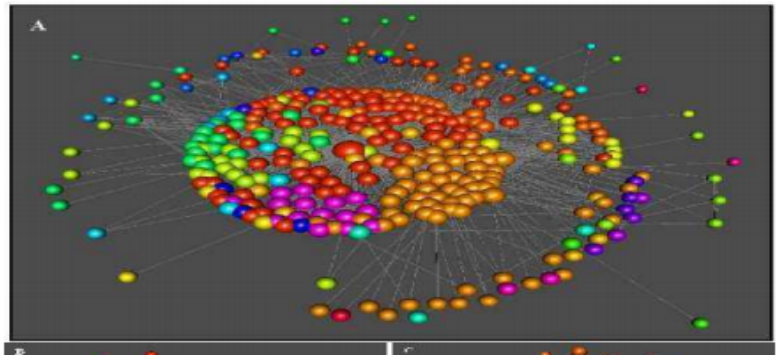


Rafols, Porter and Leydesdorff (2009)

A praise of Interdisciplinarity

The Remarkable Properties of Mythological Social Networks

The social network between characters in Homer's Odyssey is remarkably similar to real social networks today. That suggests the story is based, at least in part, on real events, say researchers



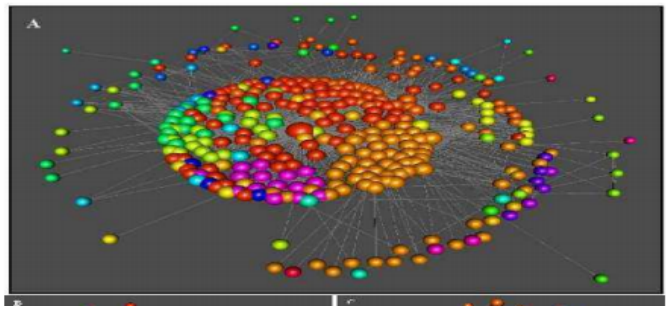
Ten years ago, few people would have heard of a social network. Today, Facebook, Twitter and LinkedIn permeate our lives. They show us how we are linked to each other and how we are

J. P Miranda *et al.*, http://arxiv.org/abs/1306.2537, 2013

A praise of Interdisciplinarity

The Remarkable Properties of Mythological Social Networks

The social network between characters in Homer's Odyssey is remarkably similar to real social networks today. That suggests the story is based, at least in part, on real events, say researchers



Ten years ago, few people would have heard of a social network. Today, Facebook, Twitter and LinkedIn permeate our lives. They show us how we are linked to each other and how we are

Scientific Question

Can we exploit methods and techniques of network theory to show the credit market is a networked market?

Networked Market

Networked Market: a market where relationships among agents are formed in a preferential way rather than randomly, as in the models of labour market described in M.O. Jackson, *Social and economic networks*, chapter 10

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Database Description

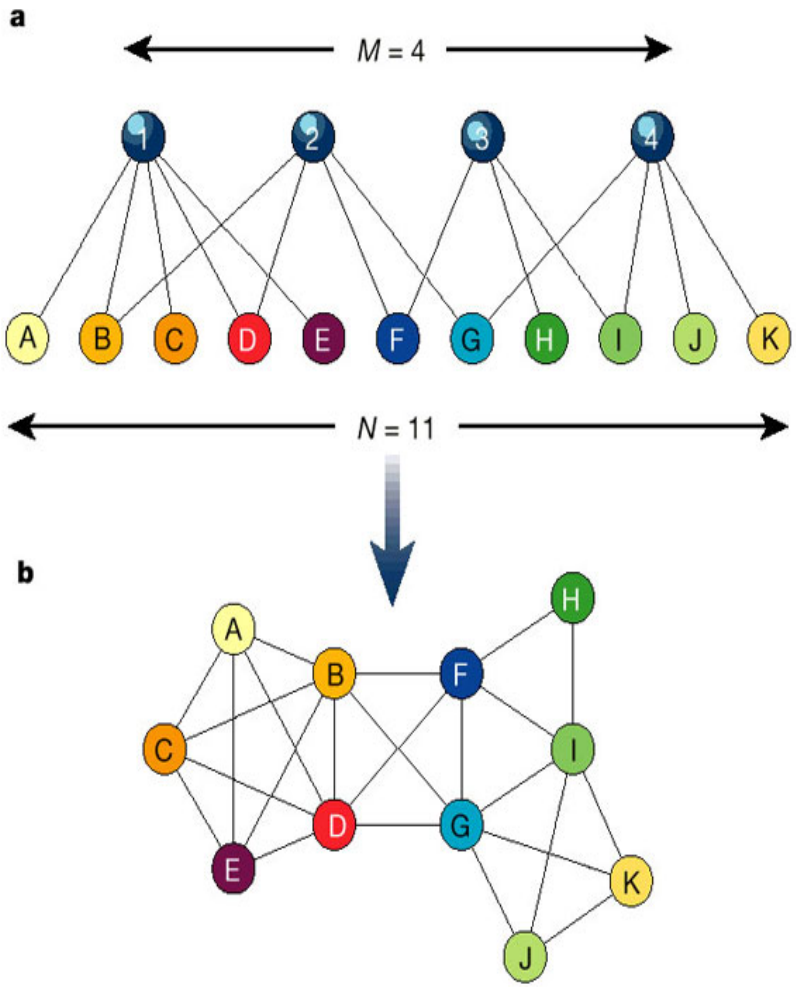
Our dataset is based on a survey of firms quoted in the Japanese stock exchange markets (Tokyo, Osaka, Nagoya, in the order of market size) and spans a period which goes from 1980 to 2012

Since 1996 the dataset includes also over-the-counter (OTC) market and/or on JASDAQ (the present OTC market)

Metadata to identify economic sector and main office prefecture of firms and the type of banks

Bipartite Networks

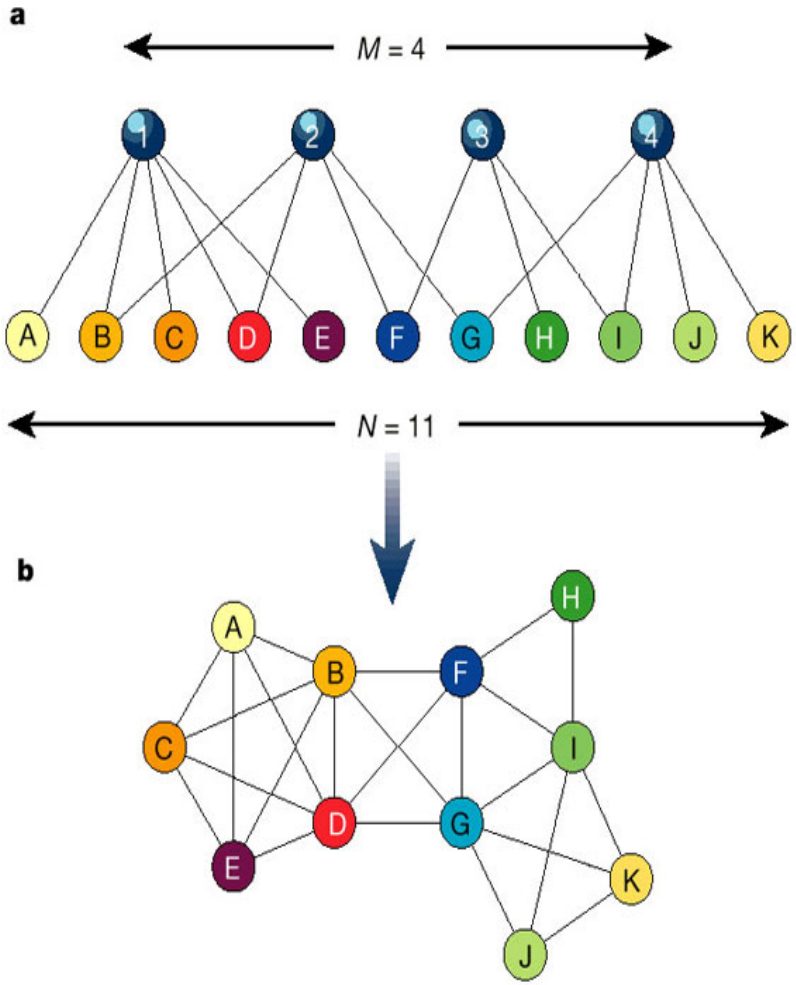
Networks with two disjoint sets of nodes with no links among each other



Bipartite Networks

Networks with two disjoint sets of nodes with no links among each other

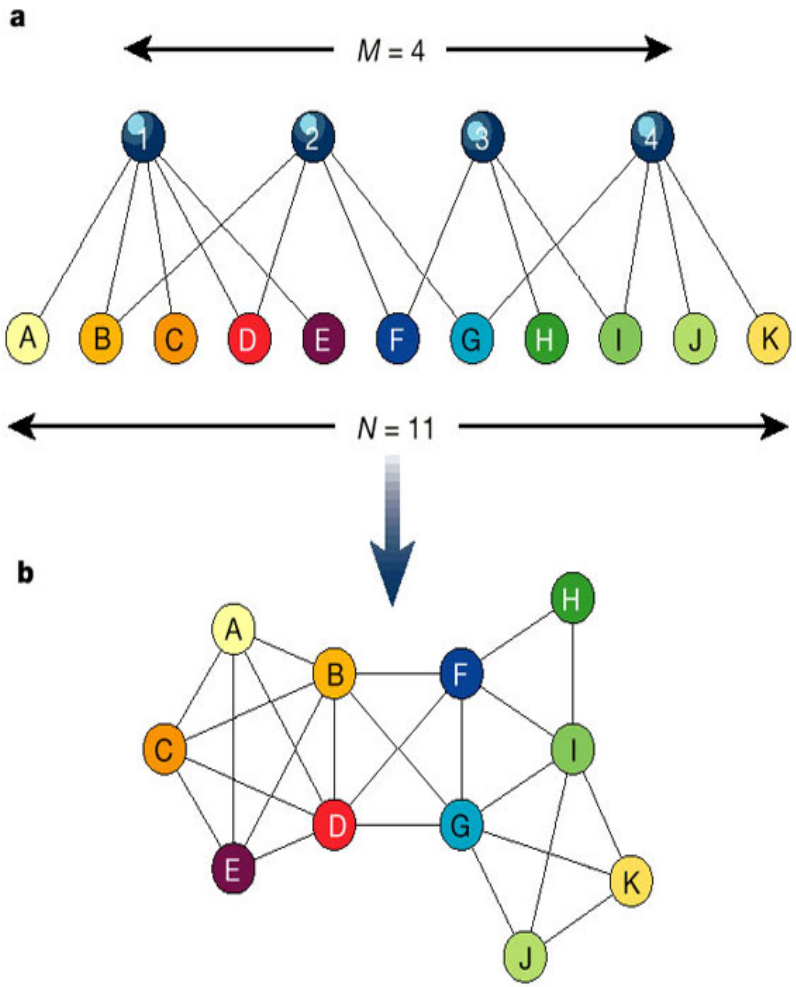
- actors and movies



Bipartite Networks

Networks with two disjoint sets of nodes with no links among each other

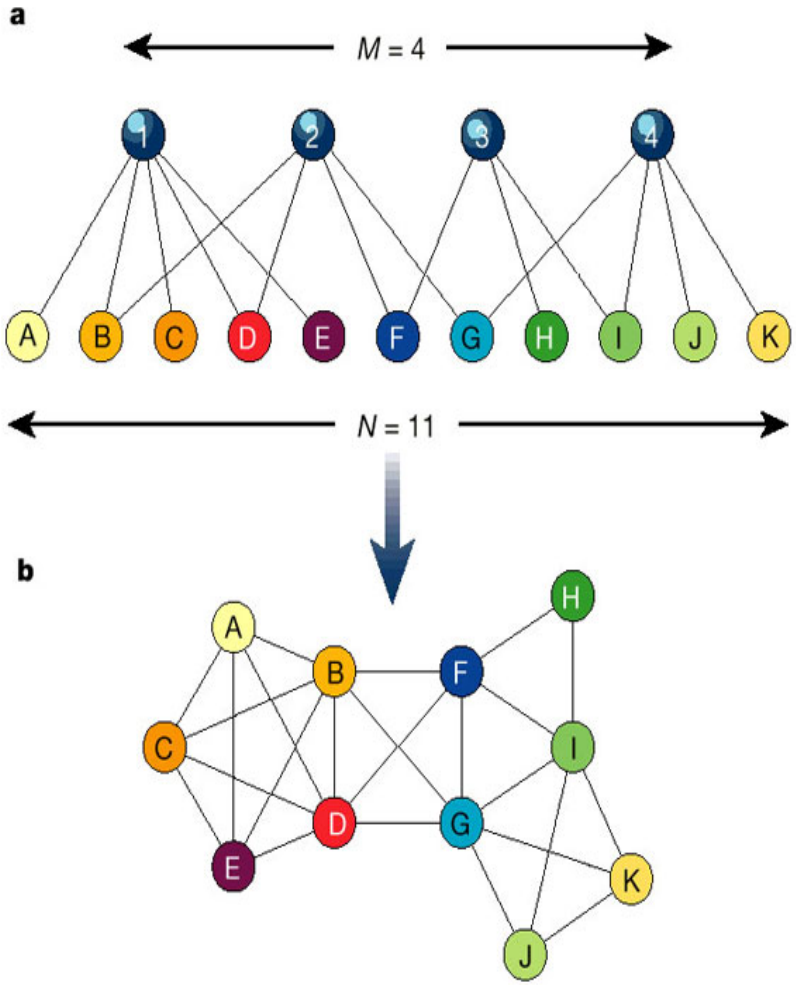
- actors and movies
- authors and papers



Bipartite Networks

Networks with two disjoint sets of nodes with no links among each other

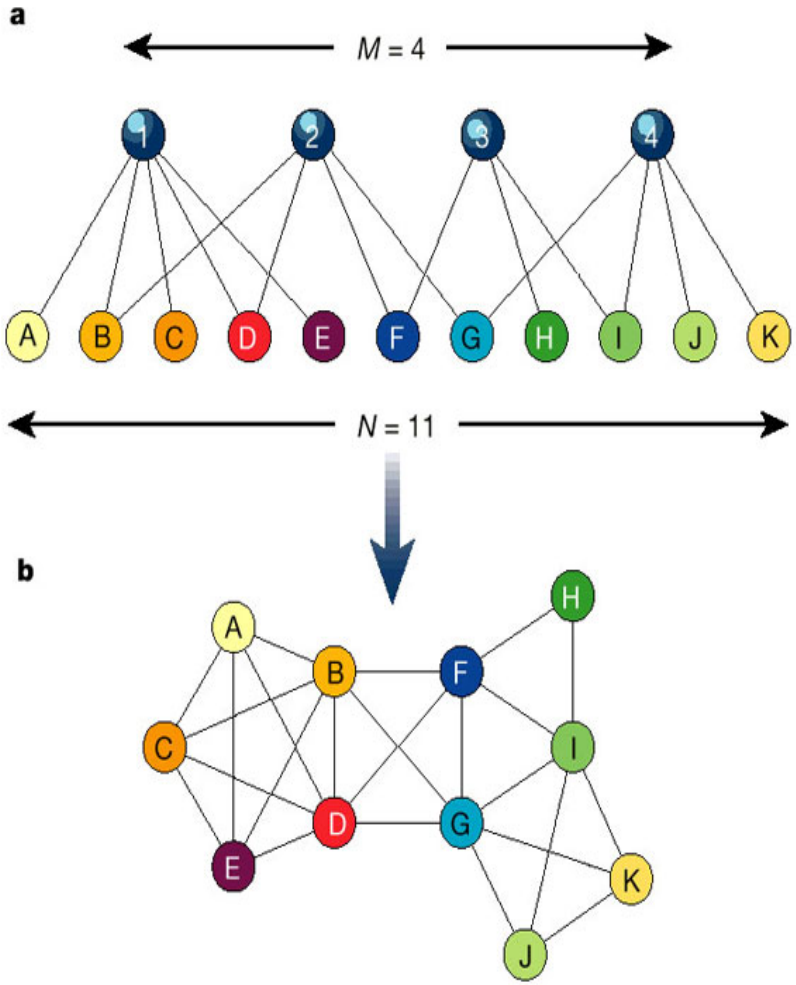
- actors and movies
- authors and papers
- banks and firms



Bipartite Networks

Networks with two disjoint sets of nodes with no links among each other

- actors and movies
- authors and papers
- banks and firms



Projected Network

We can obtain a unipartite network by linking to nodes of the same set if the share a link with a node of the other one. But projection leads to some loss of information.

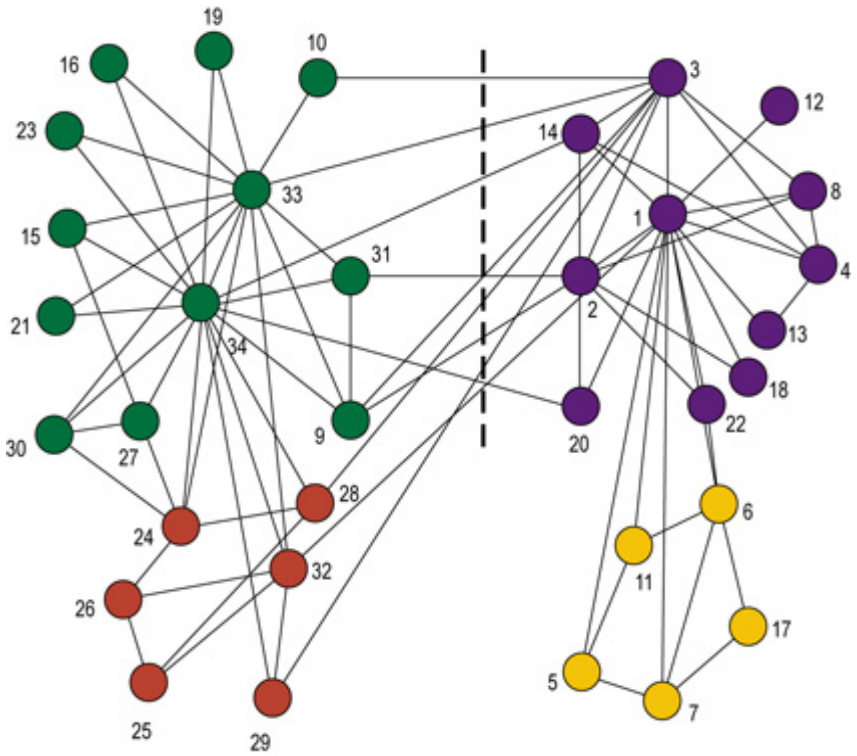
Community detection: what is a community?

The aim of community detection in networks is to identify the communities and, possibly, their hierarchical organization, by only using the information encoded in the network topology

Generic definition

Subsets of vertices with a high density of links within them, joined by a low number of links

Community detections is problematic when $n \ll m$, with n number of vertices and m total number of links.



Zachary's karate club

Modularity as a fitness measure

Partition: a division in clusters of a given network.
Algorithms for detection:

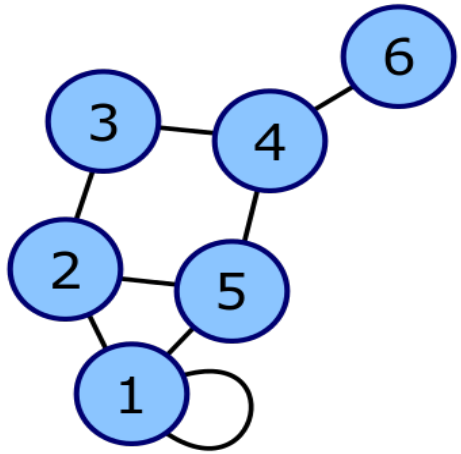
- divisive
- aggregative
- hierarchical

How can we assess the quality of the partitions they deliver?
Modularity(Newman and Girvan, 2006):

$$Q = \frac{1}{2m} \sum_{i,j} (A_{i,j} - P_{i,j}) \delta(g_i, g_j)$$

with $A_{i,j}$ adjacency matrix and $P_{i,j}$ *null model*.
$P_{i,j}$:

- retains some structural property of the original graph (number of edges, vertex degree, etc.)
- does not show community structure



$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Simple graph with its corresponding adjacency matrix

Bipartite Modularity and BRIM algorithm

Modularity for bipartite networks (Barber, 2007)

$$\mathbf{A} = \begin{bmatrix} \mathbf{O}_{p\times p} & \tilde{\mathbf{A}}_{p\times q} \\ \tilde{\mathbf{A}}^{T}_{q\times p} & \mathbf{O}_{q\times q} \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{O}_{p\times p} & \tilde{\mathbf{P}}_{p\times q} \\ \tilde{\mathbf{P}}^{T}_{q\times p} & \mathbf{O}_{q\times q} \end{bmatrix}$$

$$\mathbf{B} = \mathbf{A} - \mathbf{P}$$

with p and q number of vertices belonging to the two set and $\tilde{P}_{ij} = \frac{k_i d_j}{m}$

Bipartite Recursive Induced Modularity

$$Q = \frac{1}{m}\mathrm{Tr}\mathbf{R}^{T}\tilde{\mathbf{B}}\mathbf{T}$$

R and **T** are community index matrix with dimensions $p \times c$ and $q \times c$ respectively, where c is the number of communities. To maximize modularity:

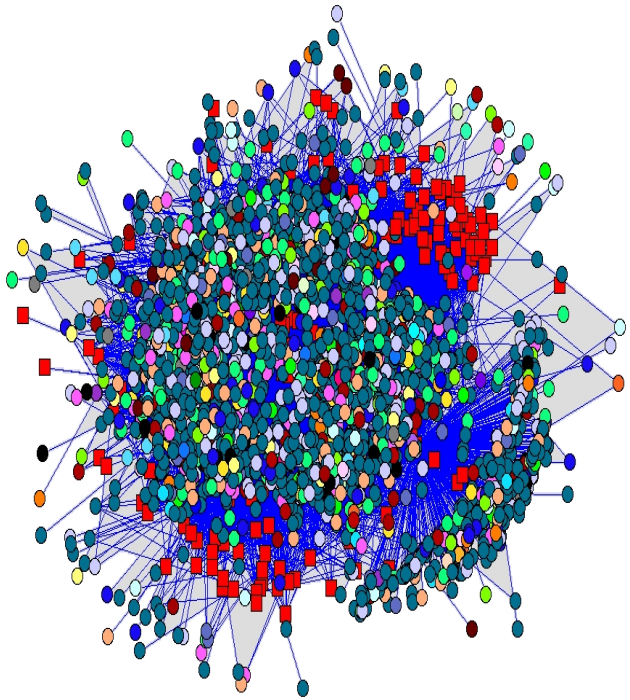
- Choose a random number of initial communities ($1 \leq c \leq \min(p, q)$)
- Randomly assign the nodes belonging to one set to the communities
- Assign the nodes of the other set in order to maximize the matrix product above
- Iterate the procedure until the modularity stops increasing
- Repeat these steps with a different number of initial communities in order to explore the configuration space

Community Index Matrix

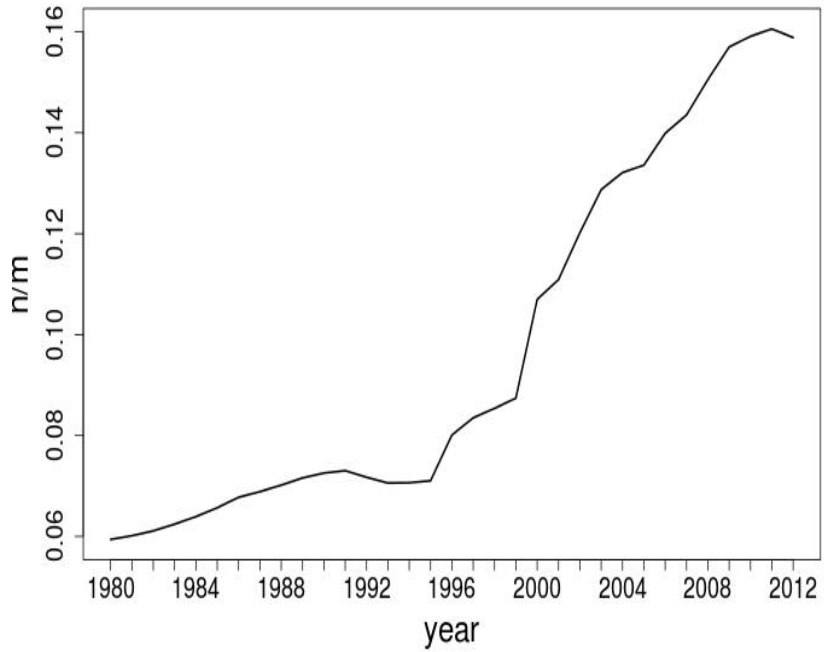
$$\mathbf{T} = \begin{array}{c} \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} c_1 & c_2 & c_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

Sparseness

The Japanese credit network is not *sparse*

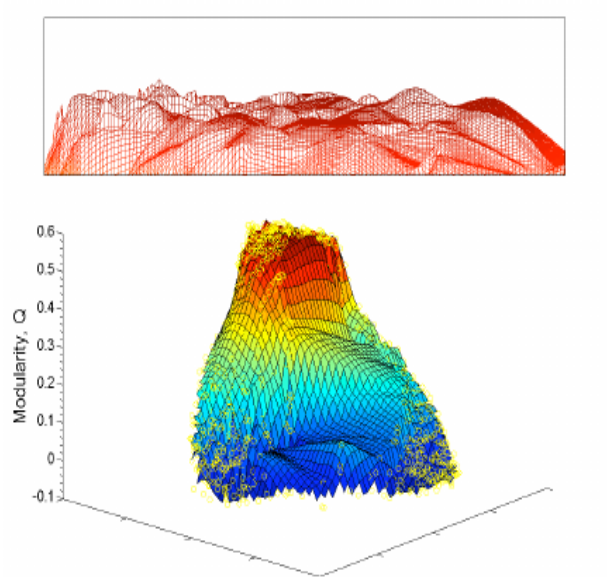


Snapshot of the network in 2011. Squares are banks and circles are firms. Colors identify different prefectures



Evolution of the ratio of the number of nodes n to the total amount of edges m during 1980/2013

The Local Maxima Trap



Modularity landscape for the metabolic network of the spirochete *Treponema Pallidum* as in Good *et al.*, 2009. The global maximum of modularity is surrounded by a plateau of sub-optimal high modularities corresponding to quite different partitions.

Methodology

Execute the BRIM algorithm choosing the best modularity partion over 100 runs

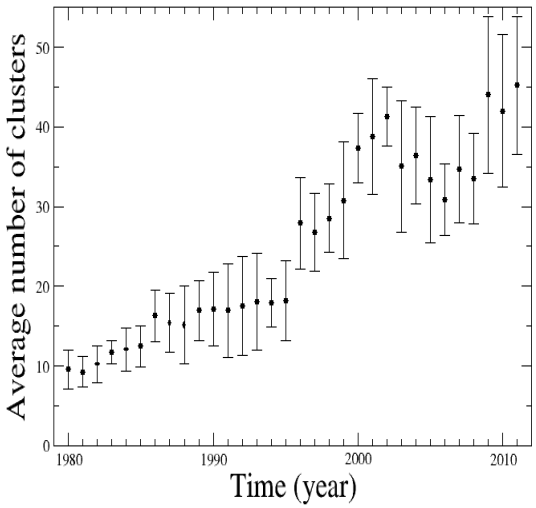
↓

Repeat with different random seeds for 20 times

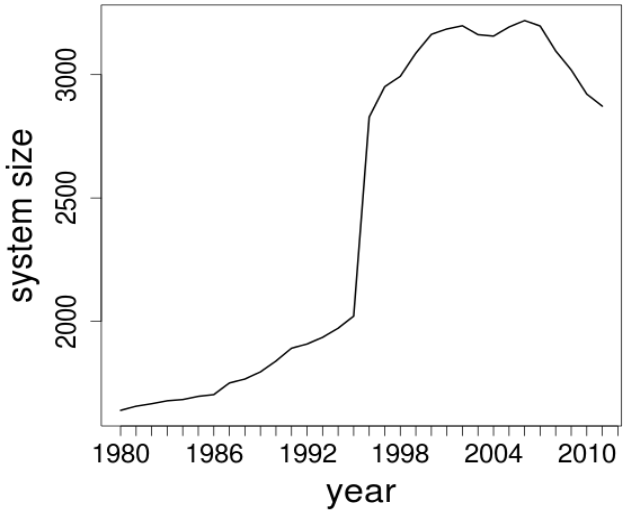
↓

Average the number of final communities obtained in the 20 runs

BRIM Results: Listed + OTC firms

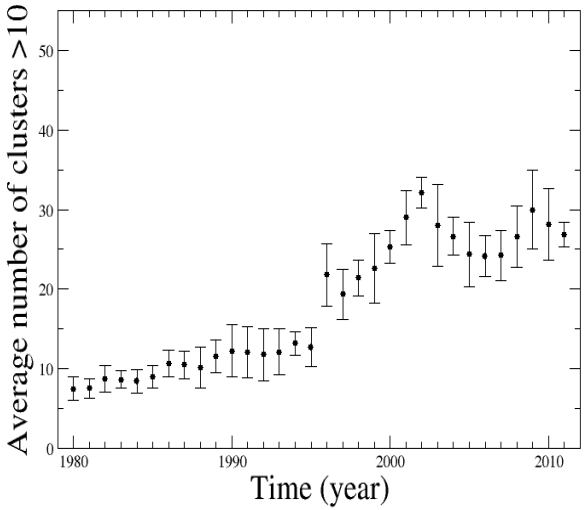


- The number of communities is increasing over times
- The jump in 1996 is due to the OTC firms include by Nikkei
- The high variance in the number of communities is significantly reduced considering only clusters with more than 10 elements

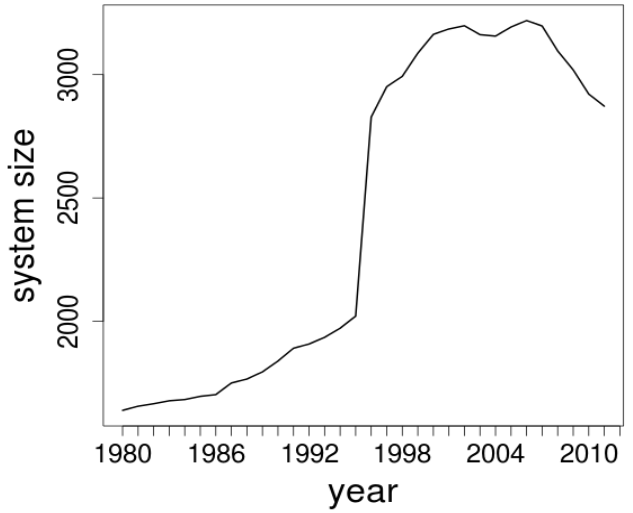


- The number of vertex is increasing over time as well
- From 2000 on is almost constant or even decreasing
- A change in market regulation can account for the behaviour in early 2000s

BRIM Results: Listed + OTC firms



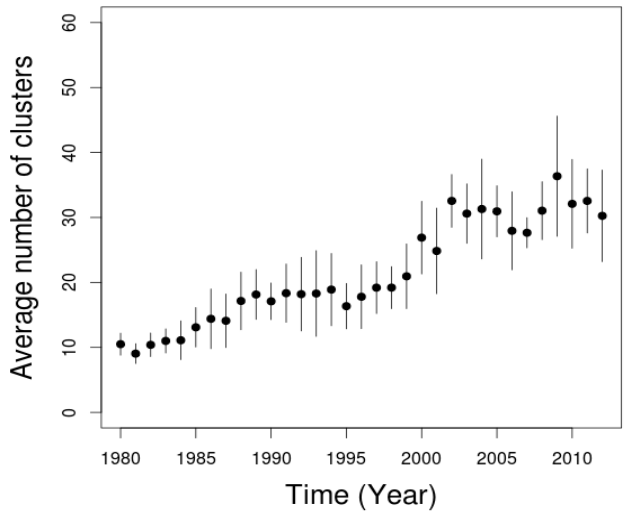
- The number of communities is increasing over times
- The jump in 1996 is due to the OTC firms include by Nikkei
- The high variance in the number of communities is significantly reduced considering only clusters with more than 10 elements

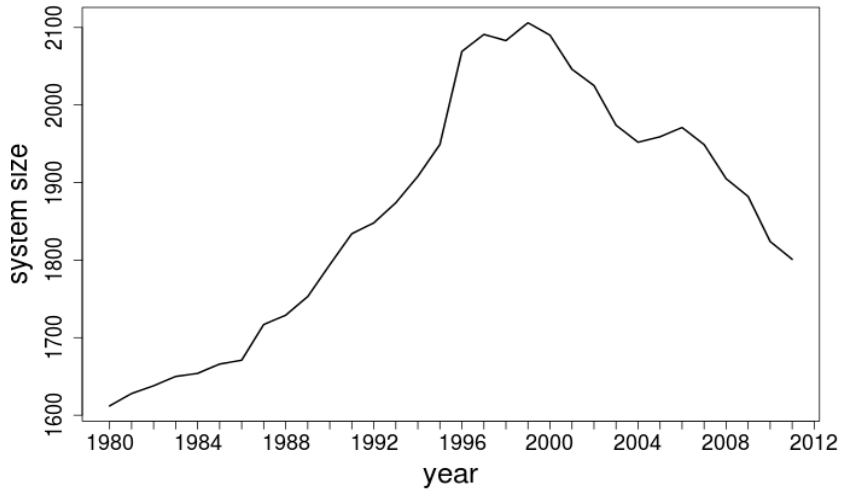


- The number of vertex is increasing over time as well
- From 2000 on is almost constant or even decreasing
- A change in market regulation can account for the behaviour in early 2000s

BRIM Results: Listed

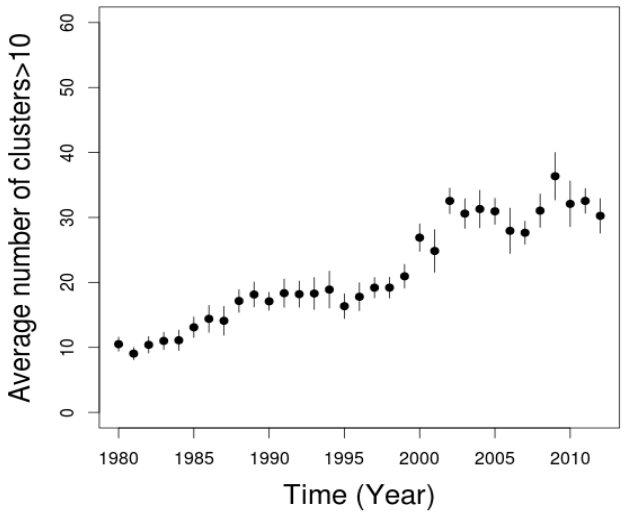
Similar results can be achieved filtering the database to retain only listed firms

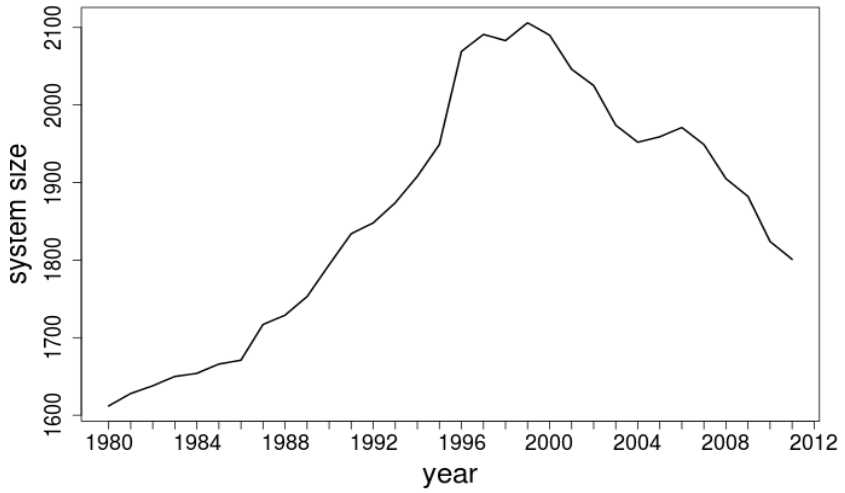




BRIM Results: Listed

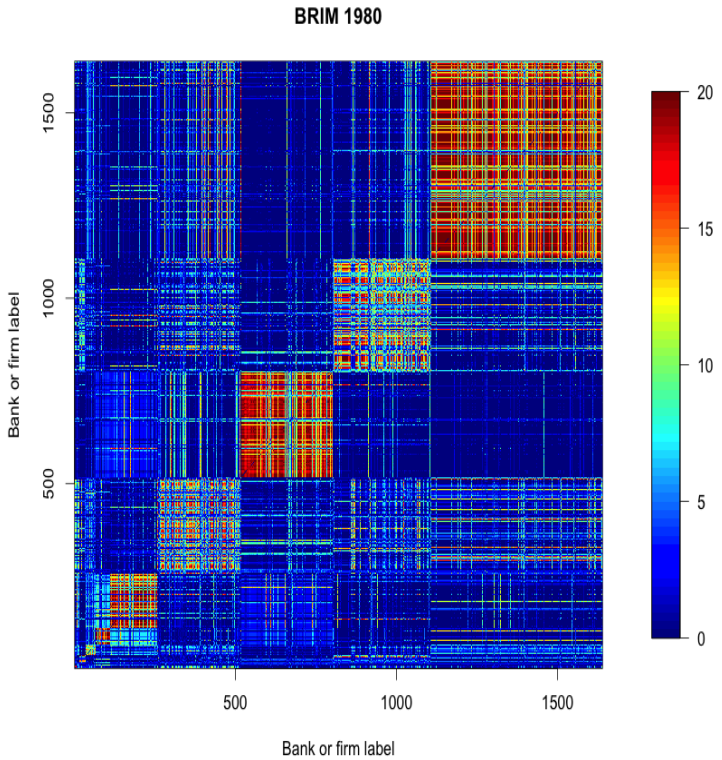
Similar results can be achieved filtering the database to retain only listed firms

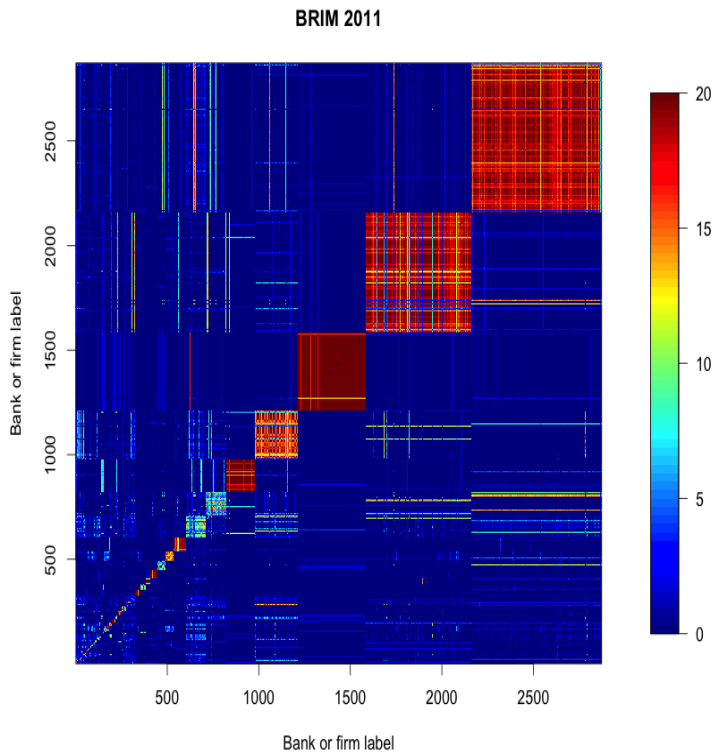




Stability of Best Partitions: Listed + OTC firms

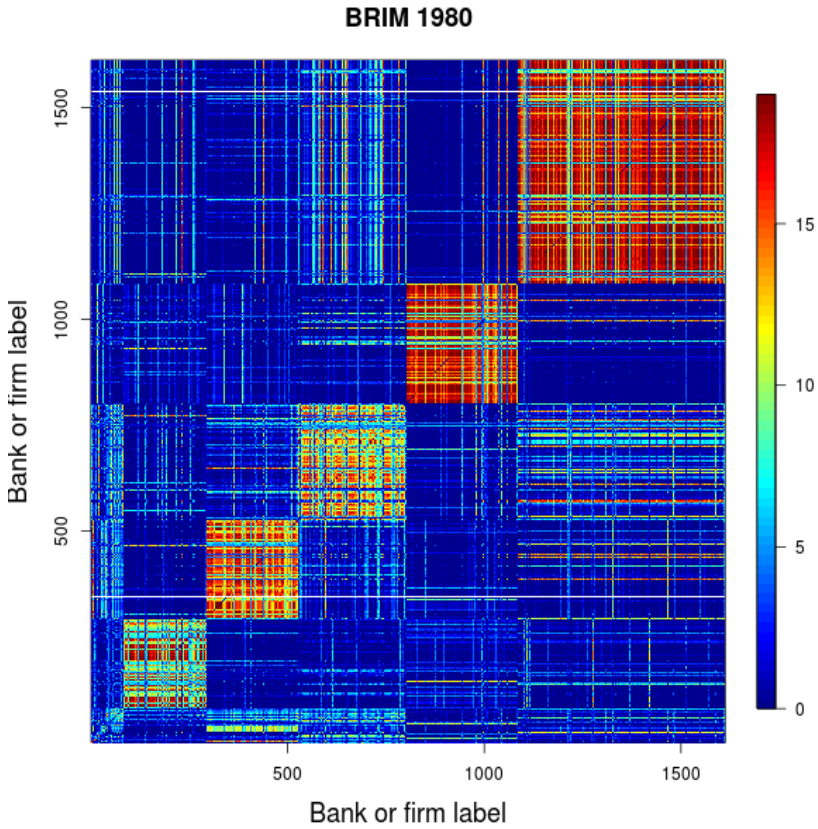
How many times a couple (i, j) belonging to the same community in the best partition is in the same cluster in the other 19 realizations?

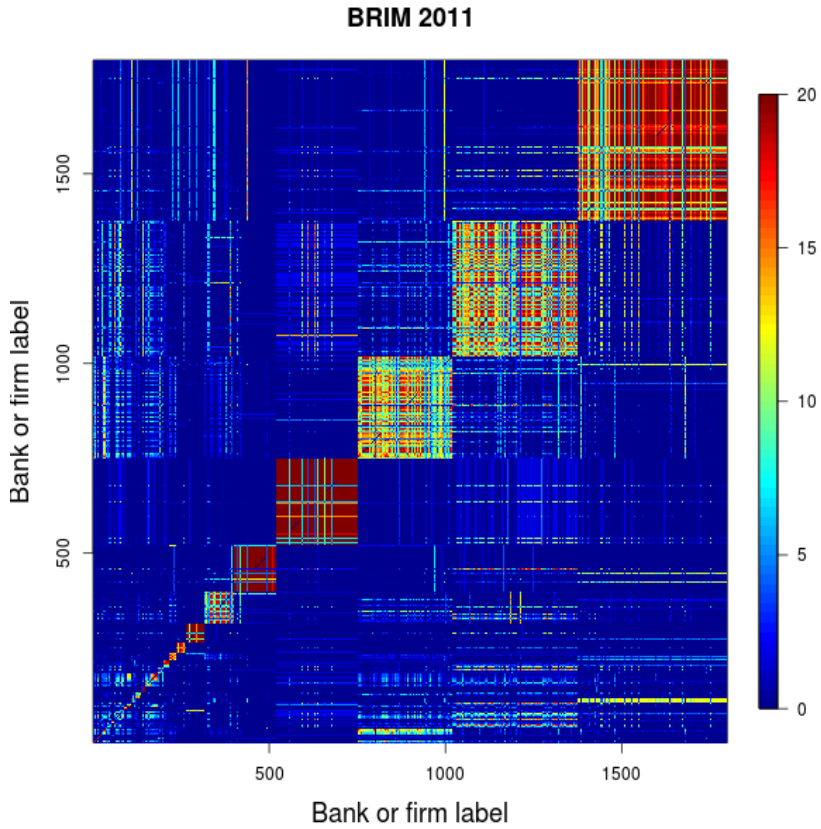




stability improves at the end of the time period

Stability of Best Partitions: Listed firms





Introduction

Are there any guiding principles driving the system self-organization?
Characterization of clusters with a methodology developed in Palermo[1]

Bank Metadata: bank type

- City Banks
- Regional Banks
- Insurance Banks
- ...

Firm Metadata: prefecture and economic sector and subsector

- Tokyo
- Aichi
- Construction
- Credit Leasing
- ...

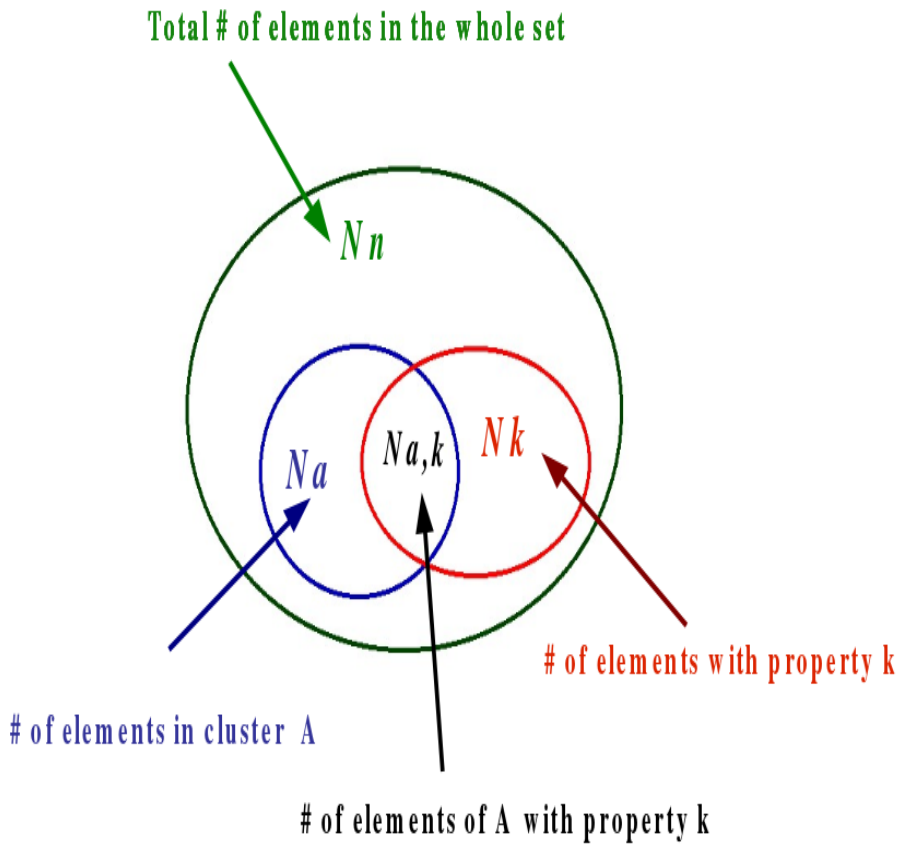
[1] M. Tumminello *et al*, J Stat Mech-Theory Exp P01019 (2011)

Statistical Validation

Suppose we are interested in understanding wheter the frequency of a certain property k (for sake of simplicity, let's k indicate one of the bank type) in one of the community detected is statistically significant. In more precise terms: given

- N_n banks in the whole set
- N_a banks in cluster A
- N_k banks of type k in the whole set
- $N_{a,k}$ banks of type k in cluster A

What is the probability that $N_{a,k}$ are in A by chance?



Statistical Validation: p-value

The probability of $N_{a,k}$ random occurencies in the set A can be calculated exactly using the hypergeometric distribution.

$$P(N_{a,k}|N_n, N_a, N_k) = \frac{\left(\begin{array}{c} N_a \\ N_{a,k} \end{array} \right) \left(\begin{array}{c} N_n - N_a \\ N_k - N_{a,k} \end{array} \right)}{\left(\begin{array}{c} N_n \\ N_k \end{array} \right)}$$

With this probability we can associate a p-value to the presence of $N > N_{a,k}$ banks of type k in cluster A:

$$p = 1 - \sum_{i}^{N_{a,k}} \frac{\left(\begin{array}{c} N_a \\ i \end{array} \right) \left(\begin{array}{c} N_n - N_a \\ N_k - N_{a,k} \end{array} \right)}{\left(\begin{array}{c} N_n \\ N_k \end{array} \right)}$$

Statistical Validation: Bonferroni Threshold

The statistical validation we are performing implies the execution of a multiple testing hypothesis.
Therefore, in setting a statistical threshold to evaluate the p-values we calculate, we need to adopt the so called multiple hypothesis test correction.

Bonferroni correction

by requiring a θ statistical threshold for the single test, the threshold B for the multiple test procedure is set to B$=\theta/T$ where T is the total number of tested hypotheses.

The Bonferroni correction is the most restrictive test correction.

Attributes Over-Expression over Time

We check the over-expression of bank's and firm's attributes

Evolution of the largest cluster

Year	banks	firms	prefecture	sector	bank type
1980	21	512	–	Electric & electronic equipment	city banks
1981	21	510	14	Electric & electronic equipment	city banks
1982	21	534	14	–	city banks
1983	19	515	13 14	Electric & electronic equipment	city banks
1984	17	549	14	Electric & electronic equipment	city banks
1985	18	547	14	Electric & electronic equipment	city banks
1986	19	553	13 14	–	city banks
1987	20	597	14	Electric & electronic equipment	city banks
1988	21	611	14	Electric & electronic equipment	city banks
1989	20	600	14	Electric & electronic equipment	city banks
1990	21	651	–	Electric & electronic equipment	city banks
1991	15	589	–	Electric & electronic equipment	city banks
1992	20	650	13	Electric & electronic equipment	city banks
1993	17	605	13 14	Electric & electronic equipment	city banks
1994	18	694	–	–	city banks
1995	14	668	13	–	city banks
1996	18	1016	13 14	–	city banks
1997	17	1003	13 14	Services	city banks
1998	21	1027	13 14	Services	city banks
1999	14	1104	13 14	Services, Wholesale trade	city banks
2000	12	880	14	Services, Wholesale trade	city banks
2001	12	941	13 27	Services, Wholesale trade	city banks
2002	7	892	13	Services, Wholesale trade	city banks
2003	7	894	13	Services, Wholesale trade	city banks
2004	12	930	13	Services, Wholesale trade	–
2005	8	971	13 14	Services, Wholesale trade	city banks
2006	8	884	13 14	Services, Wholesale trade	–
2007	14	968	13 14	Services, Wholesale trade	–
2008	7	794	13	Services, Wholesale trade	–
2009	11	790	13 14	Services, Wholesale trade	–
2010	10	739	13 14	Services, Wholesale trade	–
2011	9	703	13 14	Services, Wholesale trade	–

13 = Tokyo, 14 = Kanagawa, 27 = Osaka

Attributes Over-Expression over Time

The appearance of new over-expressed economic sectors starting from 1997 hints for a different sector concentration between listed and OTC firms.
Repeating the analysis on the two subnetworks we confirm this intuition

Listed firms

Year	banks	firms	prefecture	sector	bank type
1980	23	512	–	El. & elect. eq.	city banks
1981	19	494	14	El. & elect. eq.	city banks
1982	12	448	–	–	city banks
1983	19	528	14	El. & elect. eq.	city banks
1984	17	535	14	El. & elect. eq.	city banks
1985	16	535	–	–	city banks
1986	22	670	14	–	city banks
1987	16	541	–	–	city banks
1988	18	586	–	–	city banks
1989	18	610	–	–	city banks
1990	22	649	–	El. & elect. eq.	city banks
1991	18	612	13	–	city banks
1992	18	660	–	–	city banks
1993	16	608	–	–	city banks

OTC firms

Year	banks	firms	prefecture	sector	bank type
2000	12	309	13	–	city banks
2001	7	289	13	–	city banks
2002	10	339	13	–	city banks
2003	5	355	13	Services	city banks
2004	8	377	13	Services	city banks
2005	6	372	13	Services	city banks
2006	7	364	13	Services	–
2007	6	371	13	Services	–
2008	5	312	13	Services	city banks
2009	4	2223	13	Services	–
2010	6	259	13 14	Services	–
2011	5	236	–	–	–

Attributes Over-Expression over Time

Year	banks	firms	prefecture	sector	bank type
1980	48	257	–	–	life ins. banks - ins. banks
1981	42	177	–	Util. elect.	life ins. banks - ins. banks
1982	49	194	–	Util. elect.	life ins. banks - ins. banks
1983	47	190	–	Util. elect.	life ins. banks - ins. banks
1984	48	154	–	Util. elect.	life ins. banks - ins. banks
1985	47	167	13	Util. elect.	life ins. banks - ins. banks
1986	50	174	13	Util. elect.	life ins. banks - ins. banks
1987	98	216	13 34 40	Cred. leas.	ins. banks
1988	129	265	13 34 40	Cred. leas.	ins. banks
1989	51	223	–	Cred. leas.	life ins. banks - ins. banks
1990	27	249	–	Rail. trans.	life ins. banks
1991	133	236	40	Cred. leas.	–
1992	108	188	13 15	Cred. leas.	ins. banks
1993	124	245	13	Cred. leas.	ins. banks
1994	55	226	–	Cred. leas.	life ins. banks - ins. banks
1995	47	228	13	Cred. leas.	life ins. banks - ins. banks
1996	–	–	–	–	–
1997	69	209	13	Cred. leas.	life ins. banks - ins. banks
1998	82	236	13	Cred. leas.	ins. banks
1999	–	–	–	–	–
2000	70	136	13	Cred. leas.	ins. banks
2001	103	327	34 37 40	Cred. leas.	ins. banks
2002	58	96	–	Cred. leas.	ins. banks
2003	55	97	–	Cred. leas.	–

Year	banks	firms	prefecture	sector	bank type
1980	93	191	13	Construction	regional banks
1981	91	201	13	Construction	regional banks
1982	89	208	13	Construction	regional banks
1983	81	209	40	Construction	regional banks
1984	91	217	–	–	regional banks
1985	80	201	40	Construction	regional banks
1986	78	189	34 40	Construction	regional banks

Attributes Over-Expression over Time

community detection restricting the database only to a certain type of bank

City Banks and Listed Firms

Year	banks	firms	prefecture	sector
1980	3	237	–	–
1980	3	344	–	–
1980	3	365	13	–
1980	5	367	–	–
1981	1	129	13	–
1981	2	209	–	–
1981	2	218	–	–
1981	4	226	27	–
1981	3	272	–	–
1981	2	277	13	–
1982	1	144	21 23	–
1982	2	199	14	–
1982	3	308	–	–
1982	4	341	–	–
1982	3	351	13	–

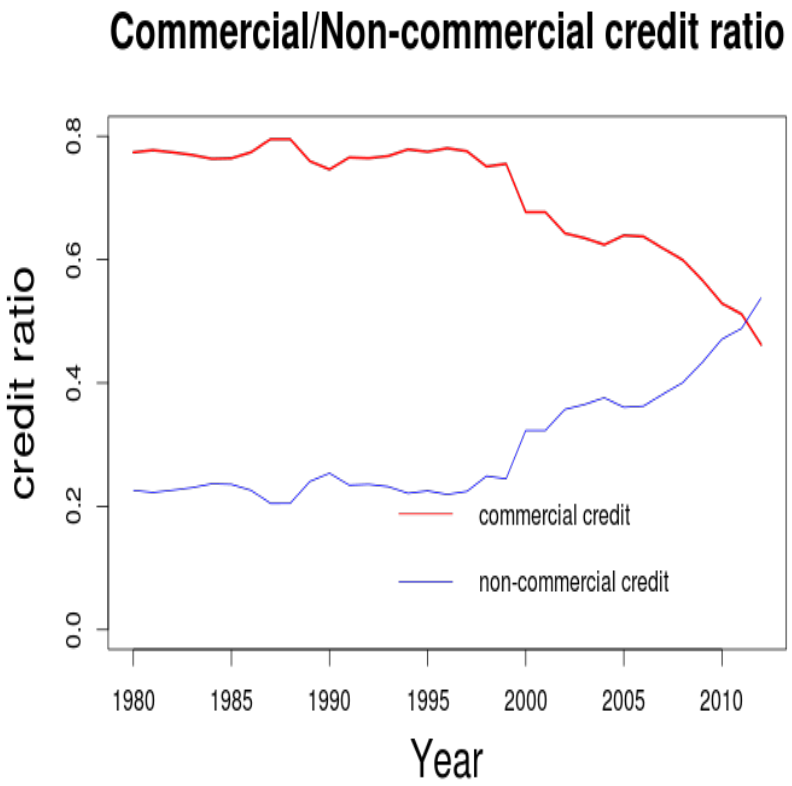
Regional Banks and Listed Firms

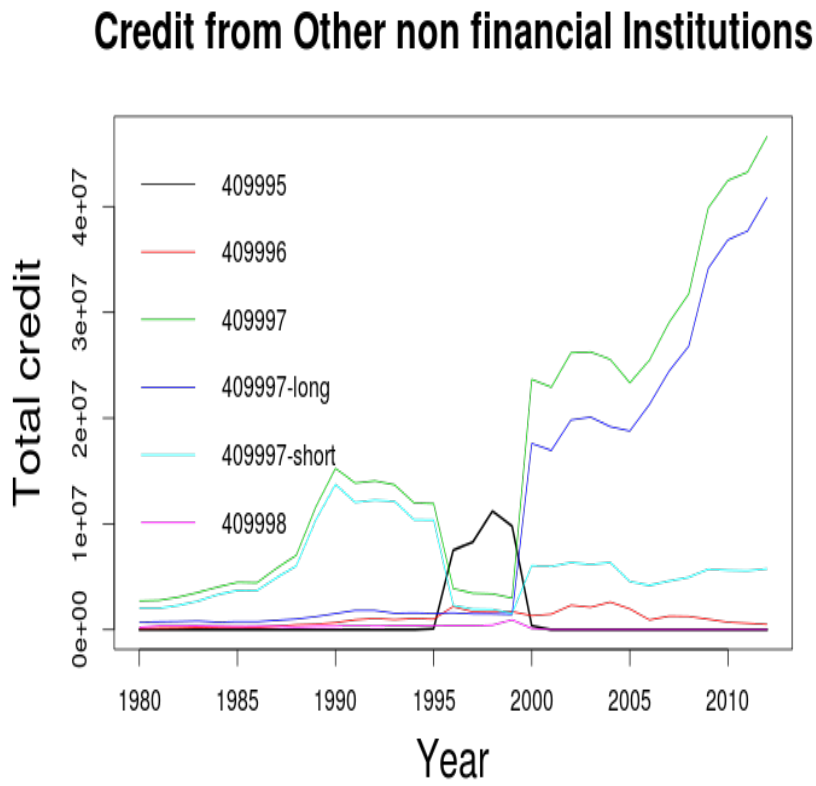
Year	banks	firms	prefecture	sector
1980	35	118	13	–
1980	35	188	34 40	–
1980	33	251	27 28	–
1980	19	274	17 22	–
1980	14	292	12 13 14	Motor Vehicles & Auto Parts
1981	1	3	–	–
1981	26	65	–	–
1981	7	105	22	–
1981	18	152	10 13	–
1981	35	173	34 40	–
1981	17	195	16 17 23	–
1981	21	206	26 27 28	–
1981	10	244	13 14	–
1982	7	87	16 17	–
1982	7	90	10	–
1982	25	131	13 15	–
1982	39	121	40	Construction
1982	14	156	–	–
1982	29	235	21 23 26 27 28	–
1982	12	333	12 13 14 22	–

Japanese Prefectures

1	Hokkaido		北海道
2	Aomori	青森県	
3	Iwate	岩手県	
4	Miyagi	宮城県	
5	Akita	秋田県	
6	Yamagata		山形県
7	Fukushima		福島県
8	Ibaraki	茨城県	
9	Tochigi	栃木県	
10	Gunma	群馬県	
11	Saitama	埼玉県	
12	Chiba	千葉県	
13	Tokyo	東京都	
14	Kanagawa		神奈川県
15	Niigata	新潟県	
16	Toyama	富山県	
17	Ishikawa		石川県
18	Fukui	福井県	
19	Yamanashi		山梨県
20	Nagano	長野県	
21	Gifu	岐阜県	
22	Shizuoka		静岡県
23	Aichi	愛知県	
24	Mie	三重県	
25	Shiga	滋賀県	
26	Kyoto	京都府	
27	Osaka	大阪府	
28	Hyogo	兵庫県	
29	Nara	奈良県	
30	Wakayama		和歌山県
31	Tottori	鳥取県	
32	Shimane	島根県	
33	Okayama	岡山県	
34	Hiroshima		広島県
35	Yamaguchi		山口県
36	Tokushima		徳島県
37	Kagawa	香川県	
38	Ehime	愛媛県	
39	Kochi	高知県	
40	Fukuoka	福岡県	
41	Saga	佐賀県	
42	Nagasaki		長崎県
43	Kumamoto		熊本県
44	Oita	大分県	
45	Miyazaki		宮崎県
46	Kagoshima		鹿児島県
47	Okinawa	沖縄県	

Shadow banking system rises?

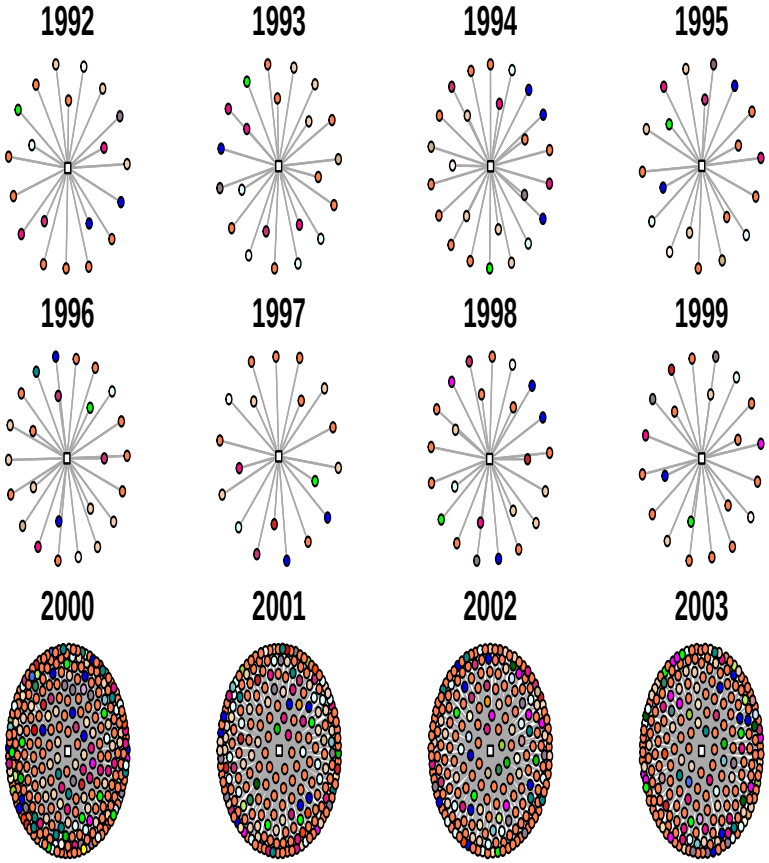


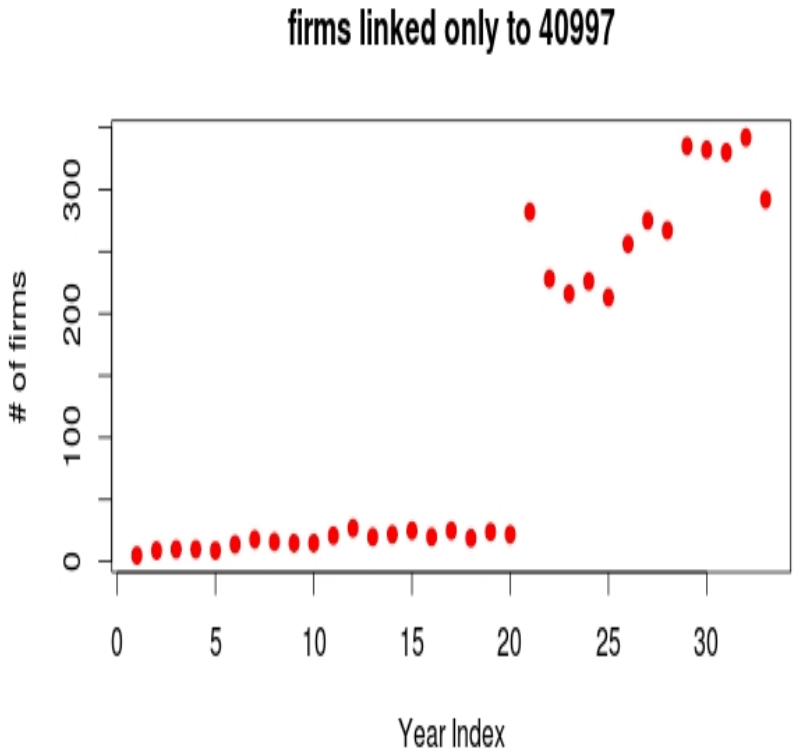


Code 409997 corresponds to **Unkown**

More on 409997

The structure of the network changes abruptly with hundreds of firm starting to link exclusively with this *aggregate* from 2000 on.





Conclusions

- The credit market considered as a bipartite network has a community structure
- A statistical validation of the communities in terms of firm localization and economic sector and bank nature reveals reasonably stable long range patterns
- We believe the above results provide evidence that the credit market is a networked market where these attributes play an important role in determining the probability that a credit relationships will be agreed between a firm and a bank.
- Agent Based Models should embody these features in their agent behavioural rules and market implementation