



Structure and Dynamics of Financial Networks

Panos M. Pardalos

Center for Applied Optimization, ISE Department,
University of Florida, Gainesville, FL 32611
pardalos@ufl.edu

Joint work with V. Boginski and S. Butenko



Introduction

- One of the challenging problems arising in the modern finance is finding efficient ways of *summarizing* and *visualizing* the stock market data that would allow one to obtain useful information about the behavior of the market.
- We consider a method summarizing the stock prices data based on representing the stock market as a *graph*.



Introduction

Graph representation of massive datasets

In many cases, a real-life massive data set can be represented as a large graph with certain attributes associated with its vertices and edges. These attributes may contain specific information characterizing the given application. Studying the structure of this graph is important for understanding the structural properties of the application it represents, as well as for improving storage organization and information retrieval.



Graph Theory Basics

Definitions:

$G = (V, E)$ is a simple undirected graph,
 $V = \{1, 2, \dots, n\}$.

$\overline{G} = (V, \overline{E})$, is the *complement graph* of $G = (V, E)$,
where $\overline{E} = \{(i, j) \mid i, j \in V, i \neq j \text{ and } (i, j) \notin E\}$.

For $S \subseteq V$, $G(S) = (S, E \cap S \times S)$ *the subgraph induced by S .*

Graph Theory Basics

Example:

$V =$

$\{1, 2, 3, 4, 5\}$

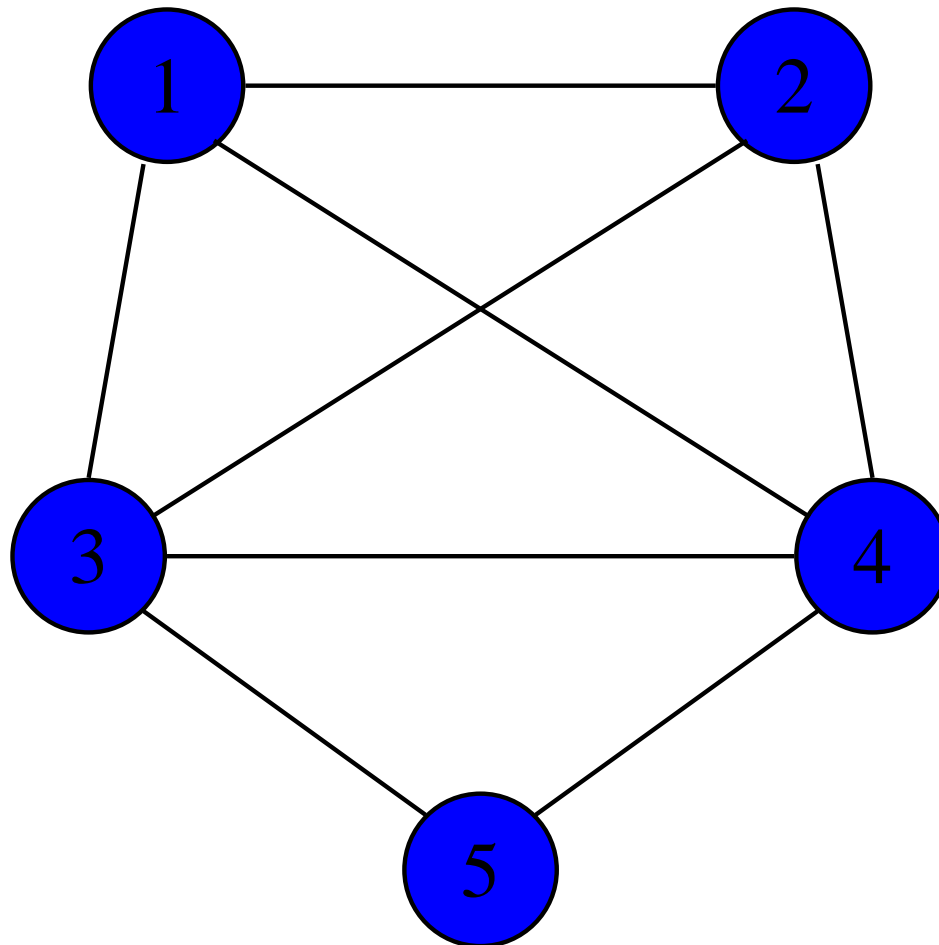
$E =$

$\{(1,2), (1,3),$

$(1,4), (2,3),$

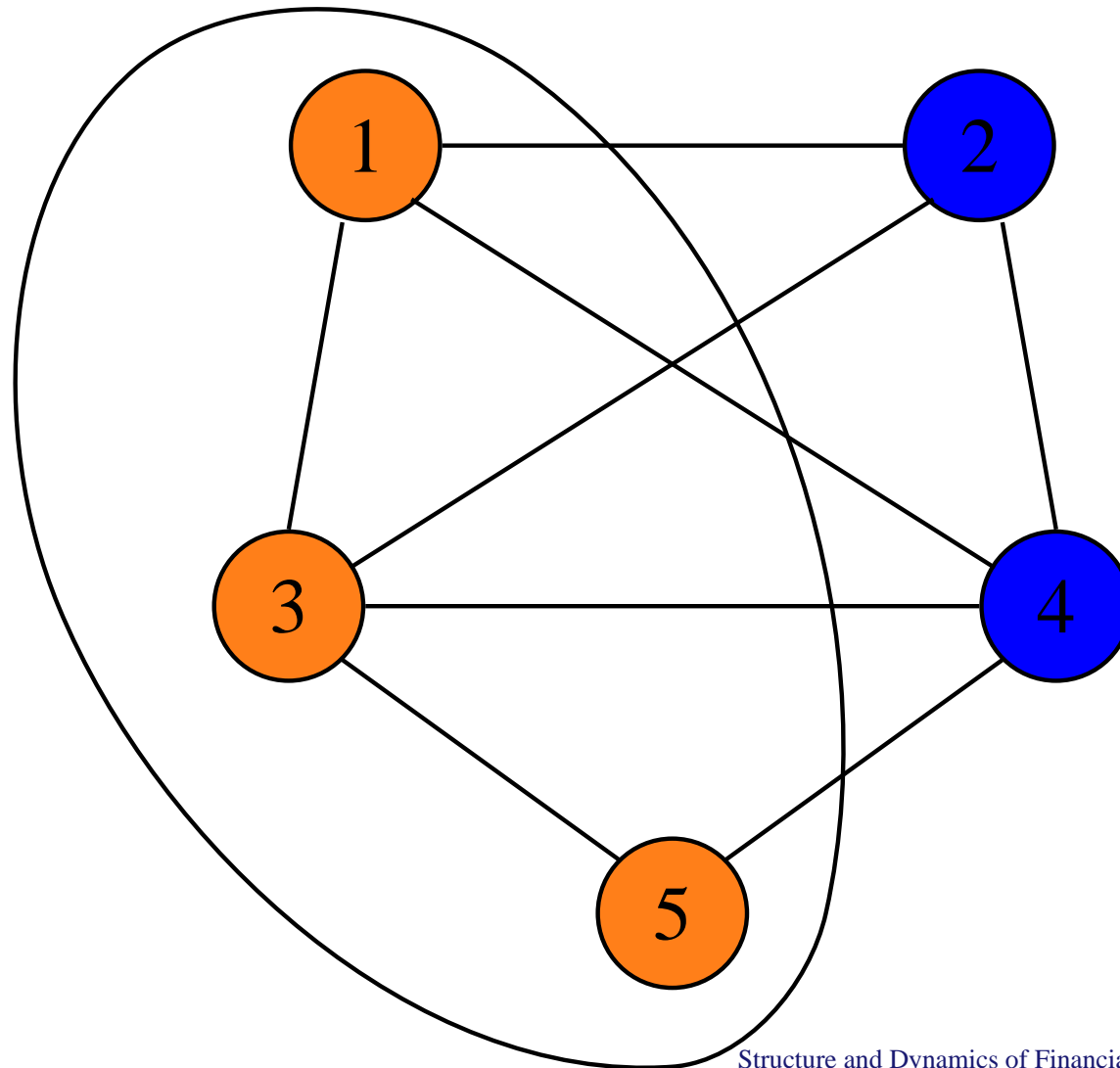
$(2,4), (3,4),$

$(3,5), (4,5)\}$



Graph Theory Basics

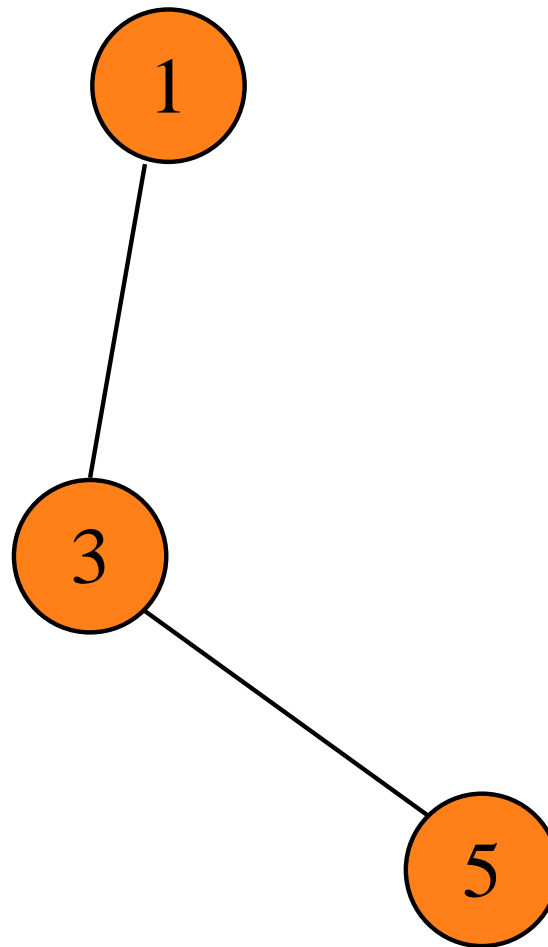
$$S = \{1, 3, 5\}$$



Graph Theory Basics

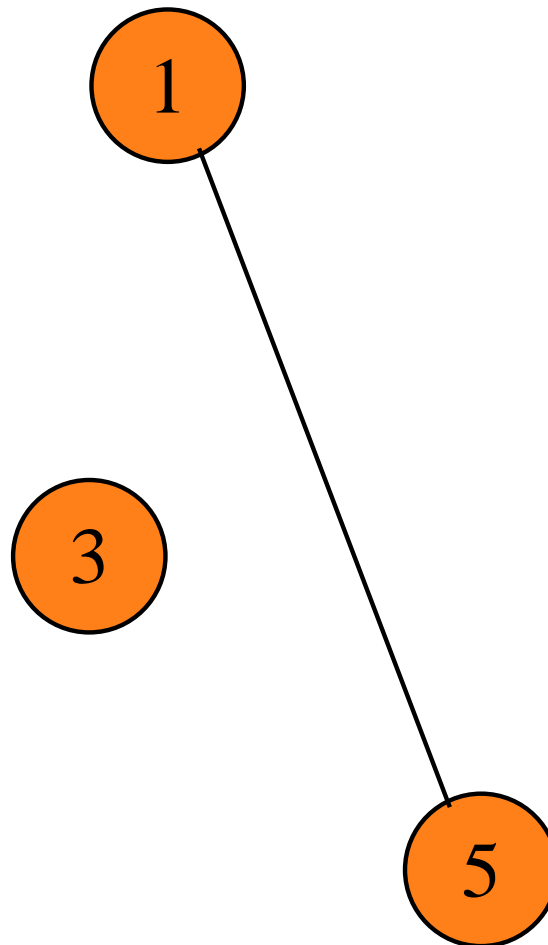
$$S = \{1, 3, 5\}$$

$G(S)$:



Graph Theory Basics

$$S = \{1, 3, 5\}$$
$$\overline{G(S)} :$$





Graph Theory Basics

The *edge density* of the graph is the ratio of the number of edges to the maximum possible number of edges

The *degree* of the vertex is the number of edges emanating from it. For every integer number k one can calculate the number of vertices $n(k)$ with the degree equal to k , and then get the probability that a vertex has the degree k as $P(k) = n(k)/n$, where n is the total number of vertices. The function $P(k)$ is referred to as the *degree distribution* of the graph.



Graph Theory Basics

A subset $C \subseteq V$ is called a *clique* if $G(C)$ is complete, i.e. it has all possible edges.

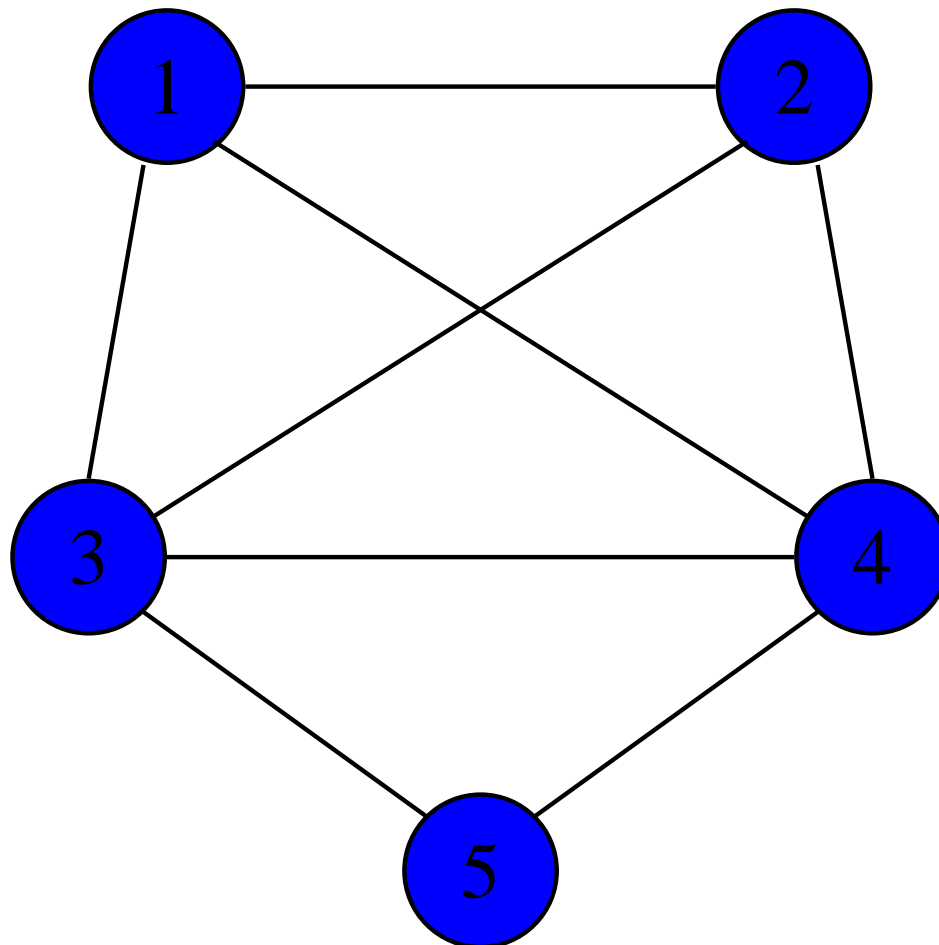
A subset $I \subseteq V$ is called an *independent set* (stable set, vertex packing) if $G(I)$ is empty.

A clique (independent set) is said to be

- *maximal*, if it is not a subset of any larger clique (independent set);
- *maximum*, if there is no larger clique (independent set) in the graph.

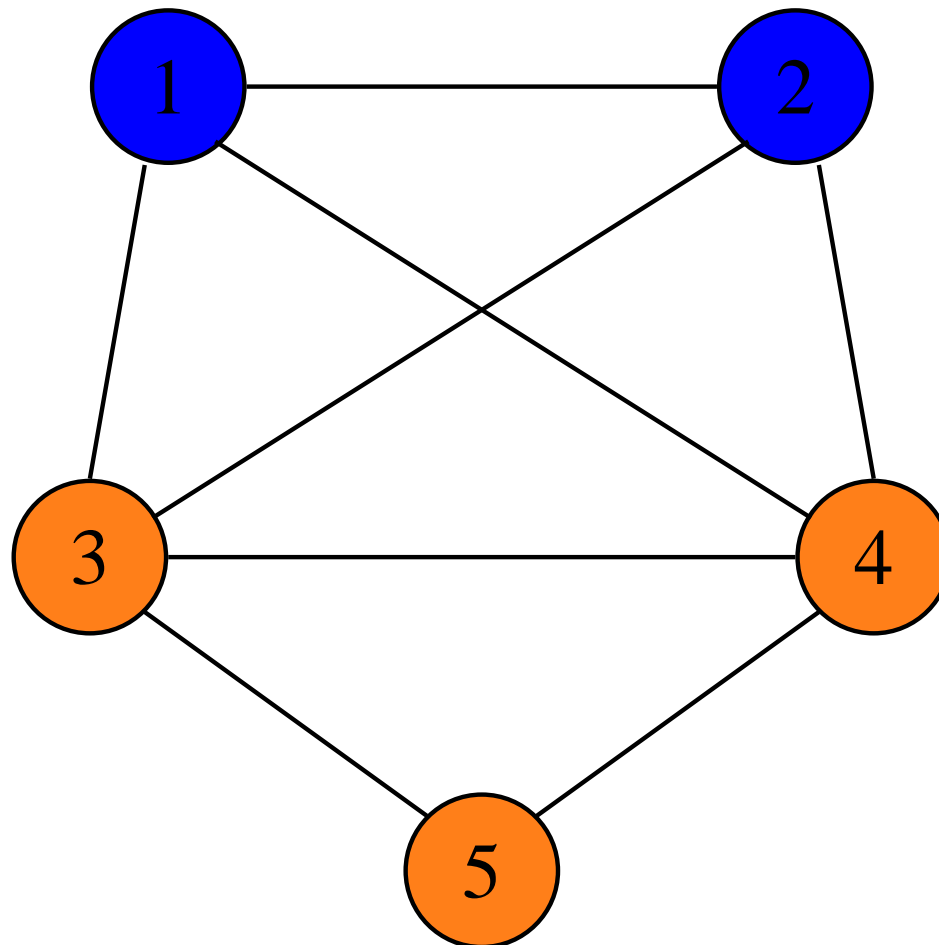
Graph Theory Basics

Example:



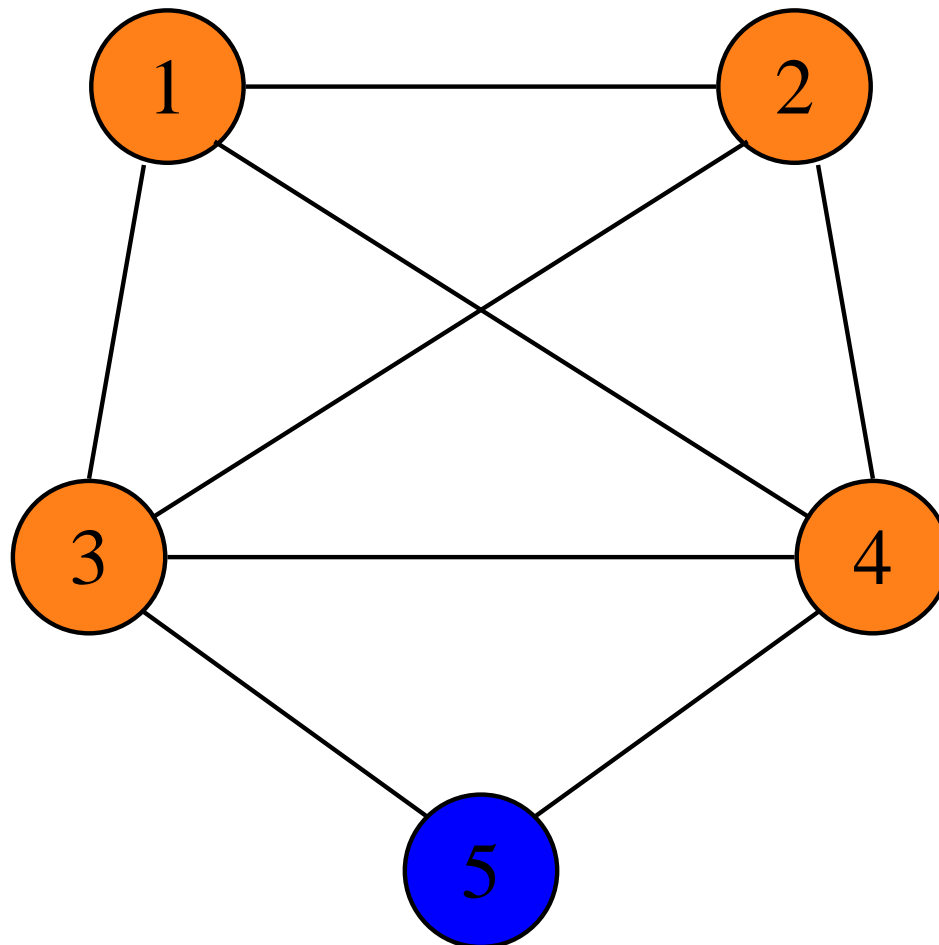
Graph Theory Basics

A maximal
clique:
 $\{3, 4, 5\}$



Graph Theory Basics

The maximum
clique:
 $\{1, 2, 3, 4\}$





Graph Theory Basics

$\alpha(G)$ – the *independence (stability) number* of G .

$\omega(G)$ – the *clique number* of G .

I is a maximum independent set of G



I is a maximum clique of \bar{G}

Maximum clique and maximum independent set problems are NP-hard



Market Graph

A stock is represented by a vertex and two vertices are connected by an edge if the cross-correlation of price fluctuations of the corresponding pair of stocks (calculated for a certain period of time) is above a threshold θ :

$$-1 \leq \theta \leq 1,$$



Market Graph

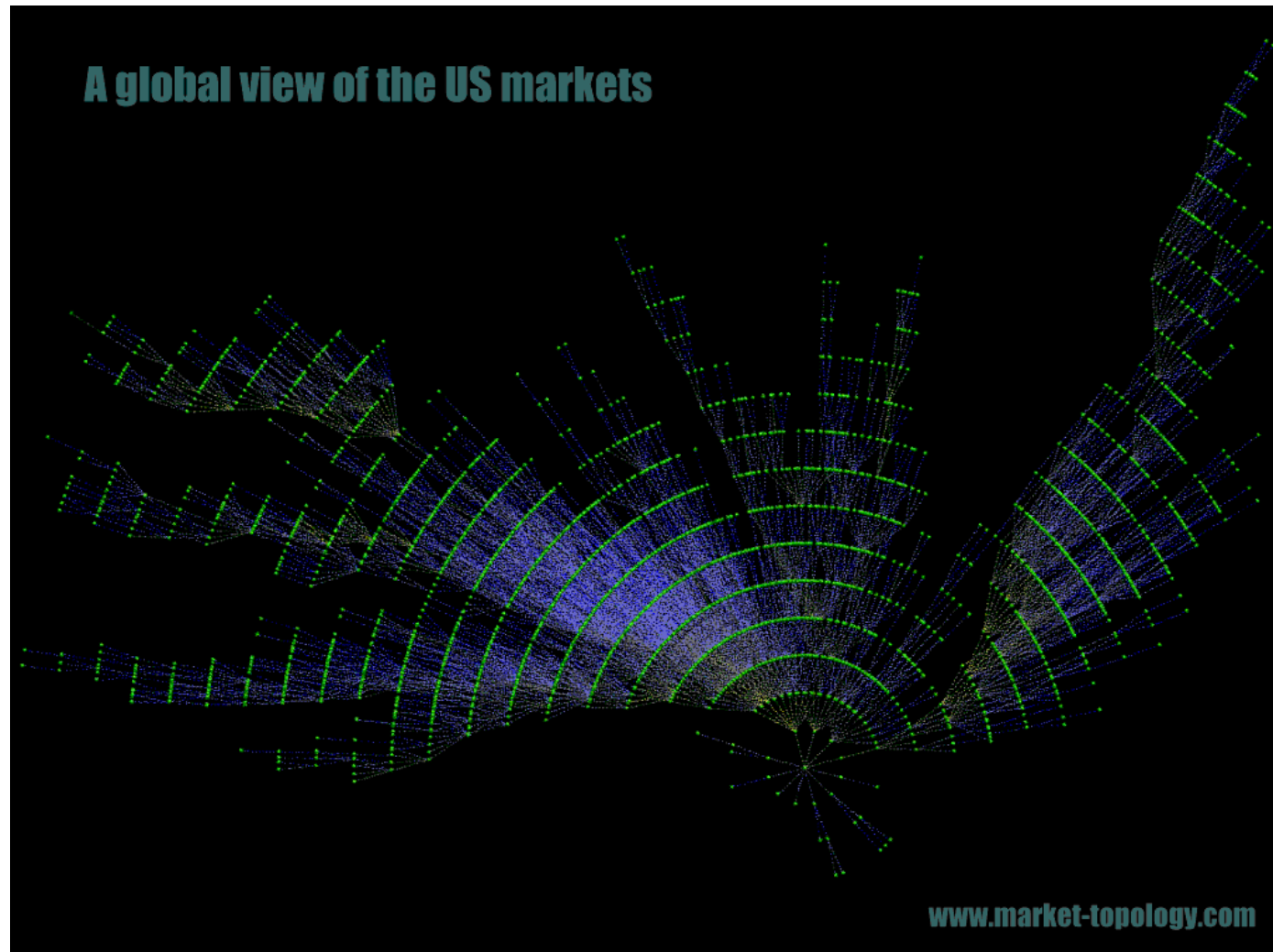
A stock is represented by a vertex and two vertices are connected by an edge if the cross-correlation of price fluctuations of the corresponding pair of stocks (calculated for a certain period of time) is above a threshold θ :

$$-1 \leq \theta \leq 1,$$

$$C_{ij} = \frac{\langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle}{\sqrt{\langle R_i^2 - \langle R_i \rangle^2 \rangle \langle R_j^2 - \langle R_j \rangle^2 \rangle}},$$

where $R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)}$ defines the return of the stock i for day t . $P_i(t)$ denotes the price of the stock i on day t .

Market Graph





Evolution of the Market Graph

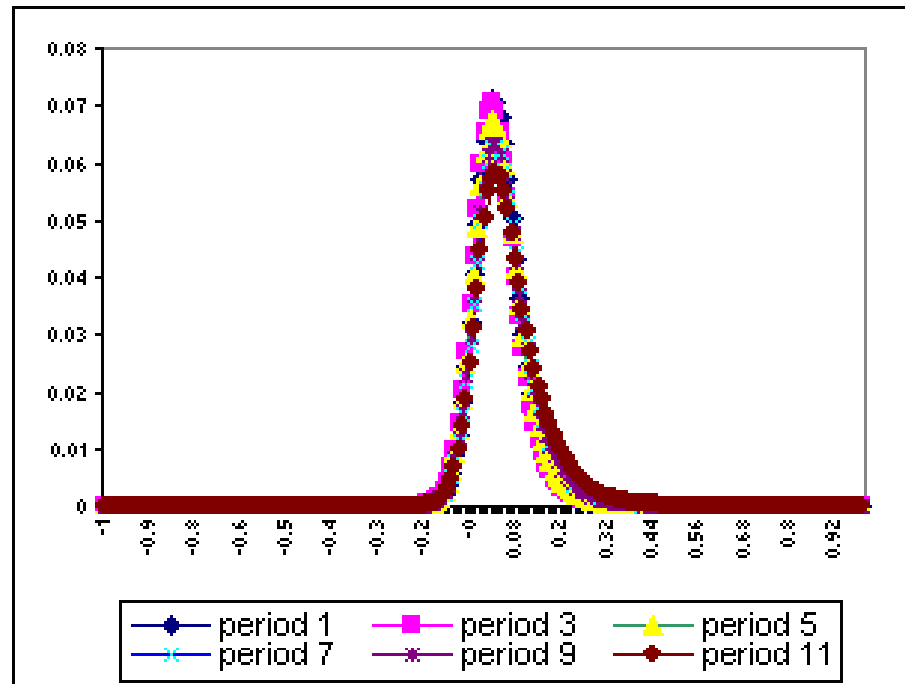
Period #	Starting date	Ending date
1	09/24/1998	09/15/2000
2	12/04/1998	11/27/2000
3	02/18/1999	02/08/2001
4	04/30/1999	04/23/2001
5	07/13/1999	07/03/2001
6	09/22/1999	09/19/2001
7	12/02/1999	11/29/2001
8	02/14/2000	02/12/2002
9	04/26/2000	04/25/2002
10	07/07/2000	07/08/2002
11	09/18/2000	09/17/2002

Evolution of the Market Graph

days\shifts	1	2	3	4	5	6	7	8	9	10	11
50	■										
100	■	■									
150	■	■	■								
200	■	■	■	■							
250	■	■	■	■	■						
300	■	■	■	■	■	■					
350	■	■	■	■	■	■	■				
400	■	■	■	■	■	■	■	■			
450	■	■	■	■	■	■	■	■	■		
500	■	■	■	■	■	■	■	■	■	■	
550	■	■	■	■	■	■	■	■	■	■	■
600		■	■	■	■	■	■	■	■	■	■
650			■	■	■	■	■	■	■	■	■
700				■	■	■	■	■	■	■	■
750					■	■	■	■	■	■	■
800						■	■	■	■	■	■
850							■	■	■	■	■
900								■	■	■	■
950									■	■	■
1000										■	■

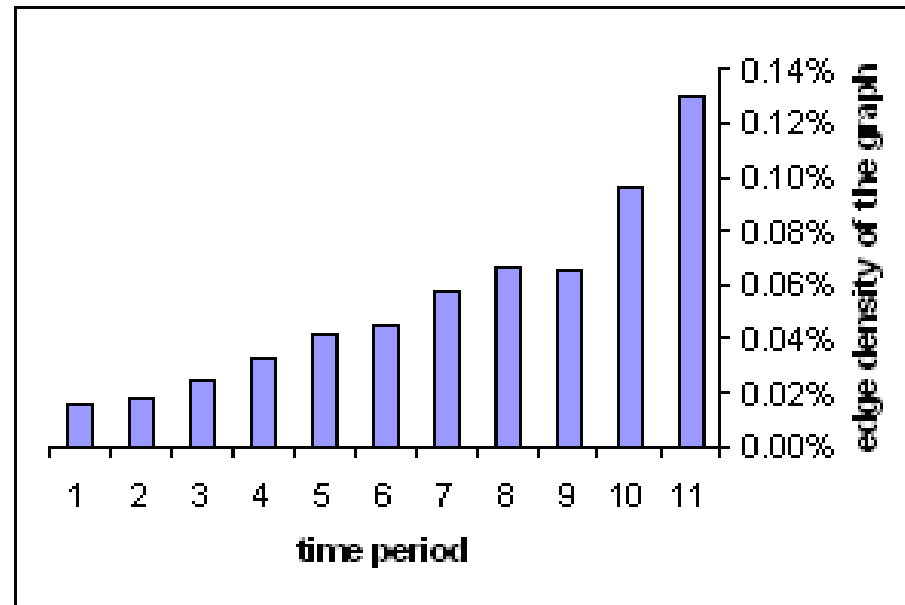
Time shifts used for studying the evolution of the market graph structure

Evolution of the Market Graph



Distribution of the correlation coefficients between all considered pairs of stocks in the market, for odd-numbered time shifts

Evolution of the Market Graph



Growth dynamics of the edge density of the market graph over time
($\theta = 0.5$)



Market Graph

- The edge density of the market graph corresponding to the latest period was more than 8 times higher than for the first period.
- The dramatic jump of the edge density suggests that there is a trend to the “**globalization**” of the modern stock market, which means that nowadays more and more stocks significantly affect the behavior of the others, and the structure of the market becomes not purely random.

Market Graph

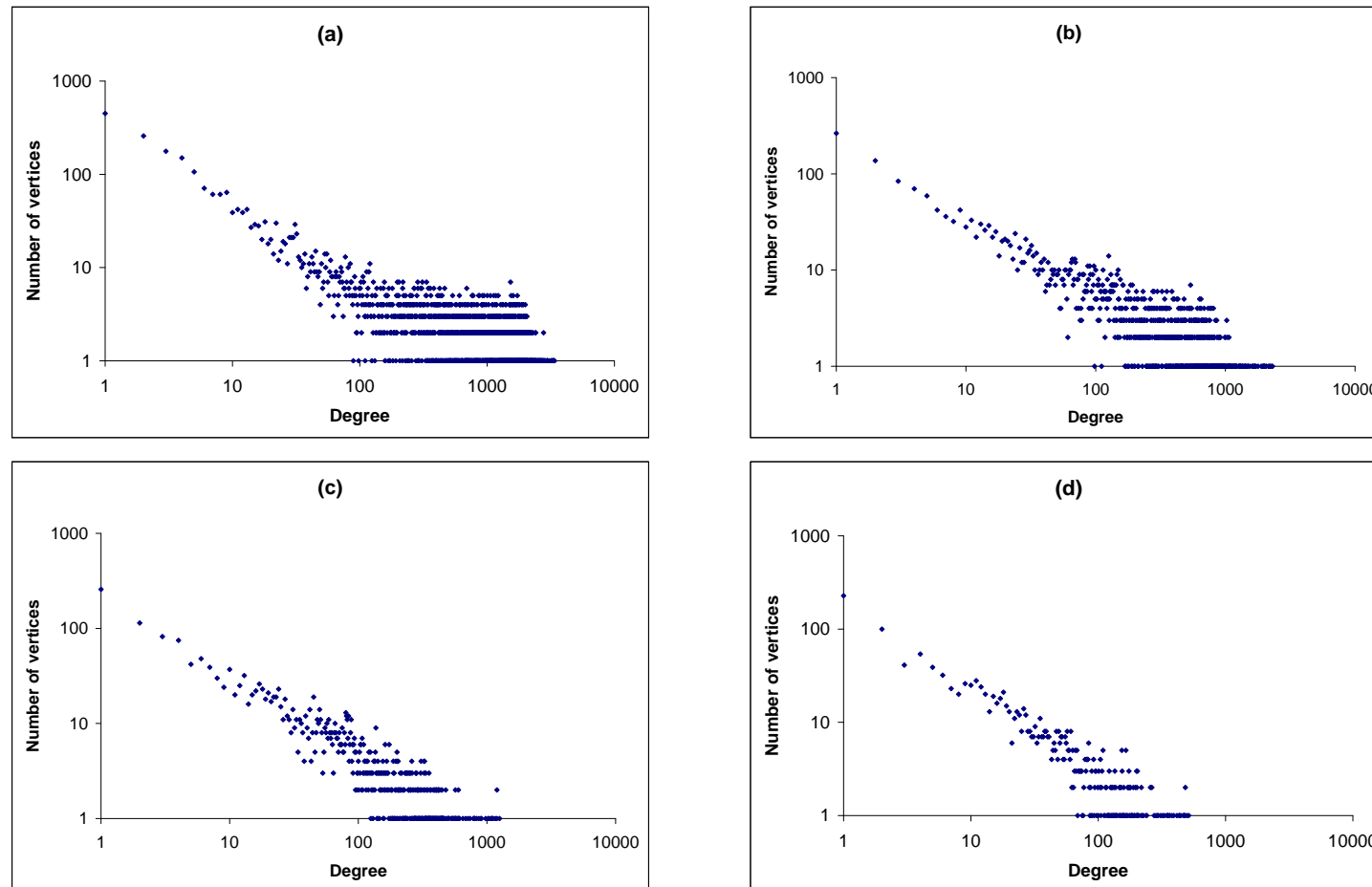


Figure 1: Degree distribution of the market graph for (a) $\theta = 0.2$; (b) $\theta = 0.3$; (c) $\theta = 0.4$; (d) $\theta = 0.5$

Market Graph

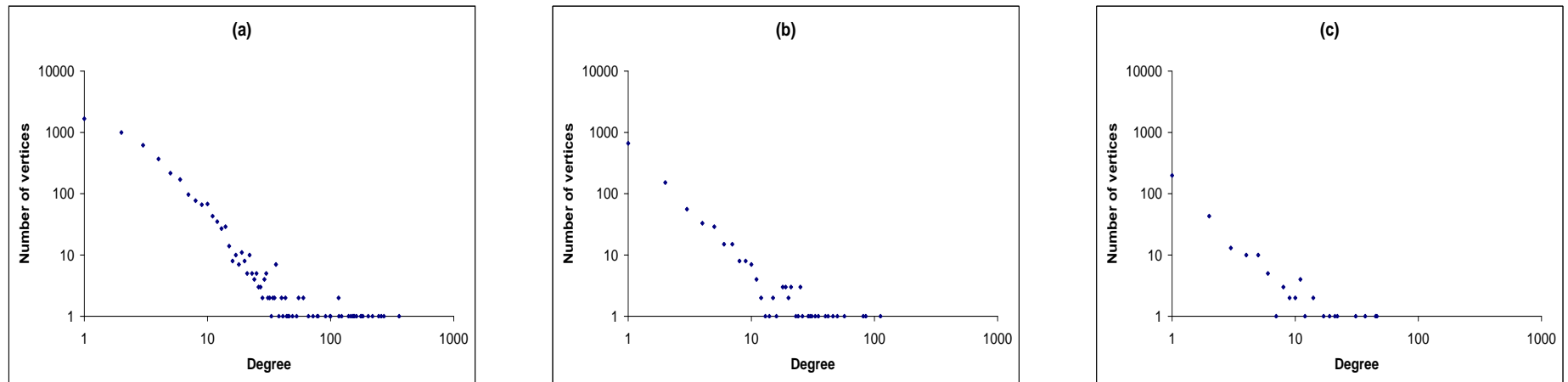
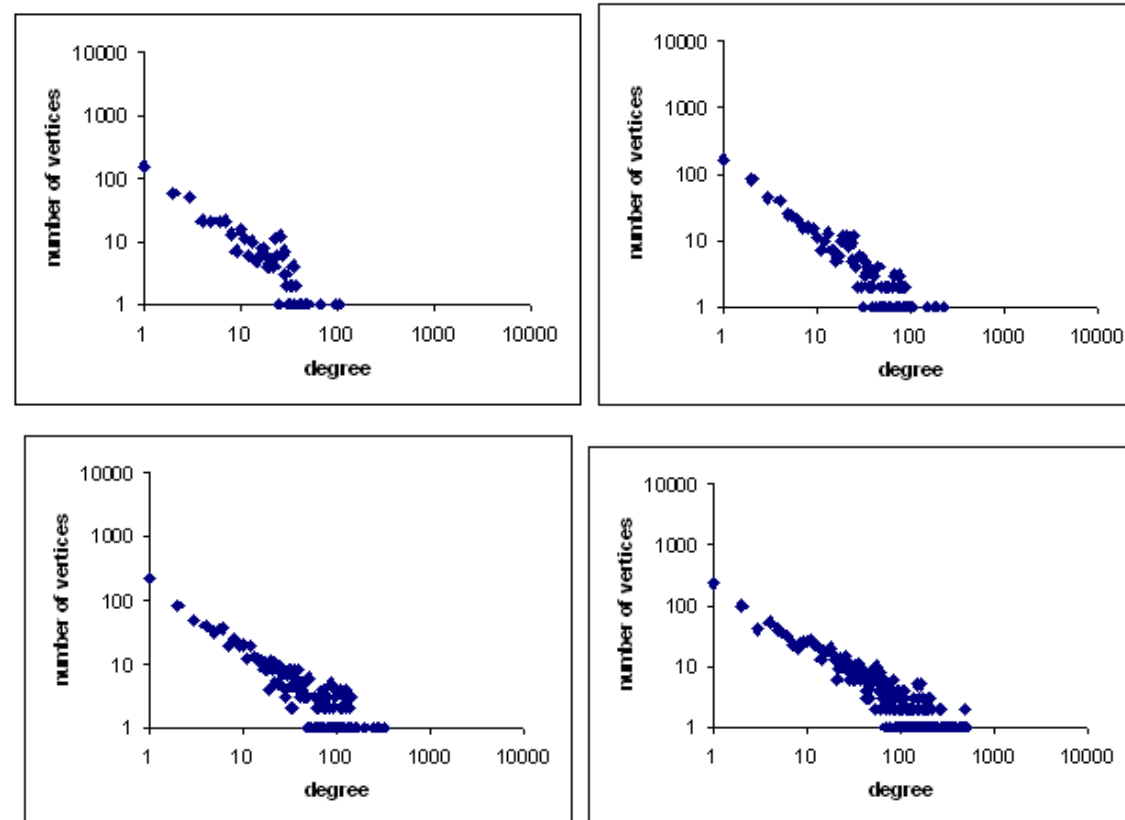


Figure 2: Degree distribution of the complementary market graph for (a) $\theta = -0.15$; (b) $\theta = -0.2$; (c) $\theta = -0.25$

Evolution of the Market Graph



Degree distribution of the market graph for periods 1, 4, 7, and 11 (logarithmic scale, $\theta = 0.5$).



Market Graph

- The degree distribution represents the *global pattern of connections* in the market graph.
- It turns out that the degree distribution of the market graph has a highly specific (*power-law*) structure. This structure remains stable over the considered time intervals in 2000-2002.
- An interesting fact is that besides the market graph, many other graphs representing real-life datasets arising in diverse application areas also have a well-defined power-law structure. This fact served as a motivation to introduce a concept of “*self-organized networks*”, and it turns out that this phenomenon also takes place in finance.



Power-Law Model

The fraction $P(k)$ of vertices of degree k in the graph is proportional to some power of k , *i.e.*,

$$P(k) \propto k^{-\gamma}. \quad (1)$$

Equivalently, one can represent it as

$$\log P \propto -\gamma \log k, \quad (2)$$

which demonstrates that this distribution would form a **straight line in the logarithmic scale**, and the slope of this line would be equal to the value of the parameter γ .



Examples of Power-Law Graphs

- **Call Graph**

The vertices are telephone numbers, and two vertices are connected by an edge if a call was made from one number to another.

- **Web Graph**

The vertices of this graph are documents and the edges are links (URLs) that point from one document to other.

- **Internet Graph**

The vertices are *routers* navigating packets of data or groups of routers (*domains*). The edges in this graph represent wires or cables in the physical network.

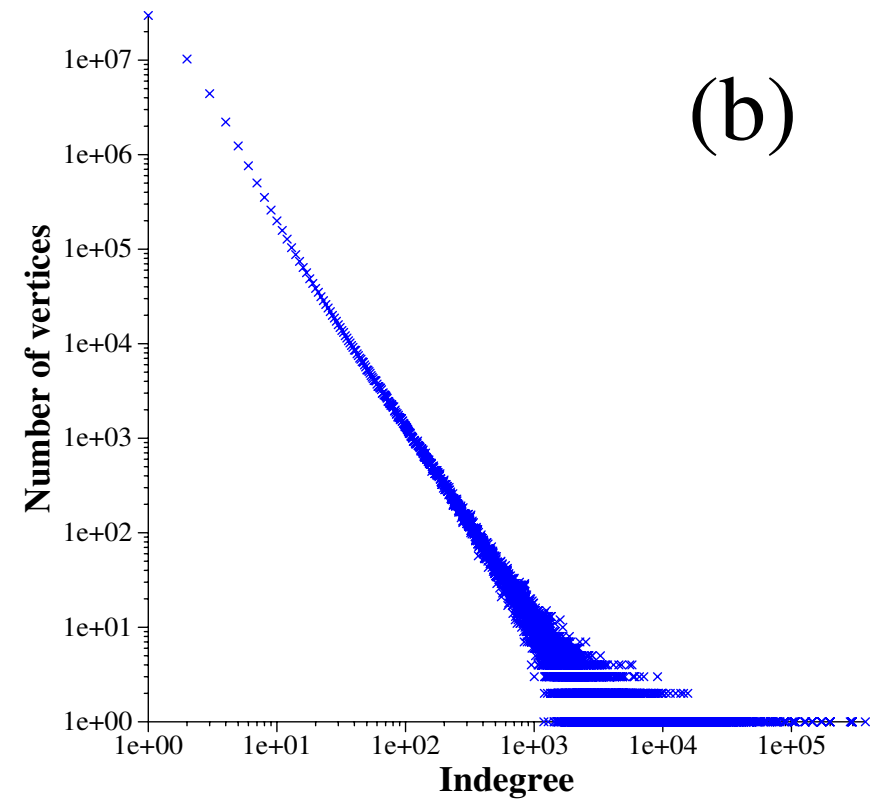
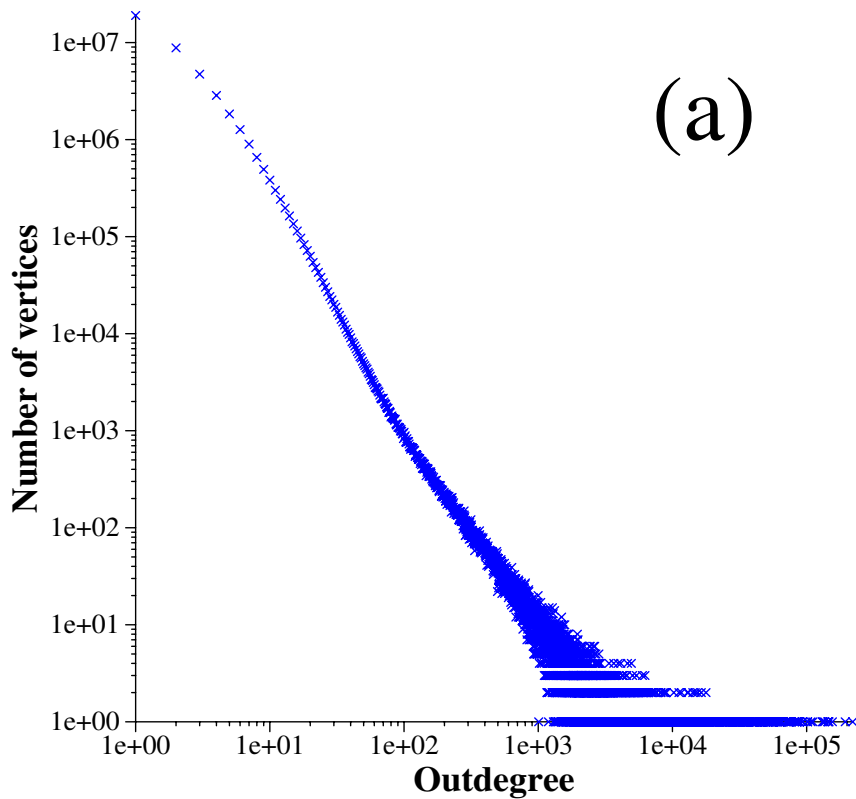


Call Graph

Abello, Pardalos, Resende (1999): AT&T billing records

- A graph based on one 20-day period had 290 million vertices and 4 billion edges.
- The analyzed one-day call graph had **53,767,087 vertices and over 170 millions of edges.**
- 3,667,448 connected components, most of them tiny; only 302,468 (or 8%) components had more than 3 vertices. A giant connected component with 44,989,297 vertices was computed.

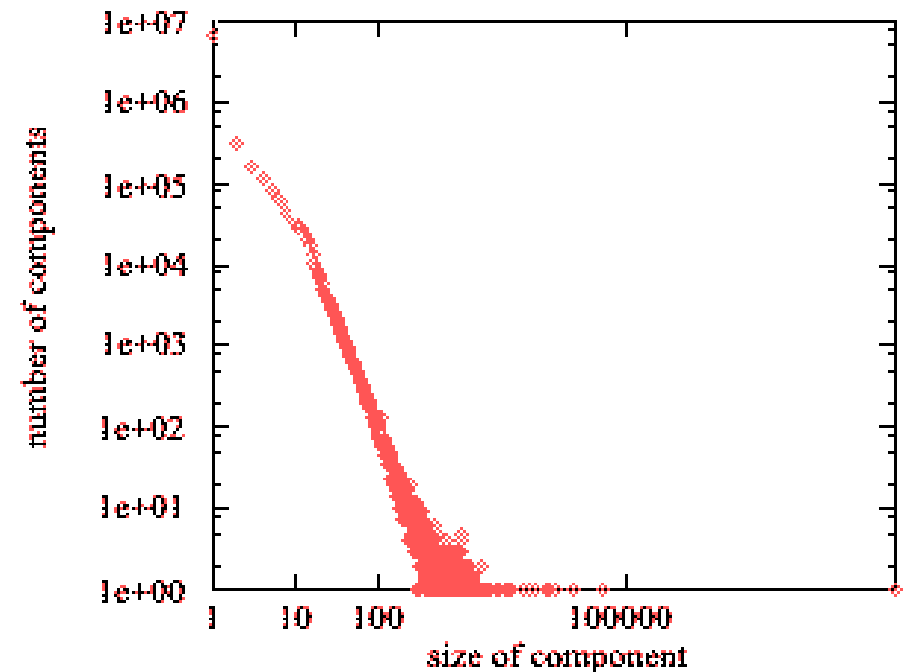
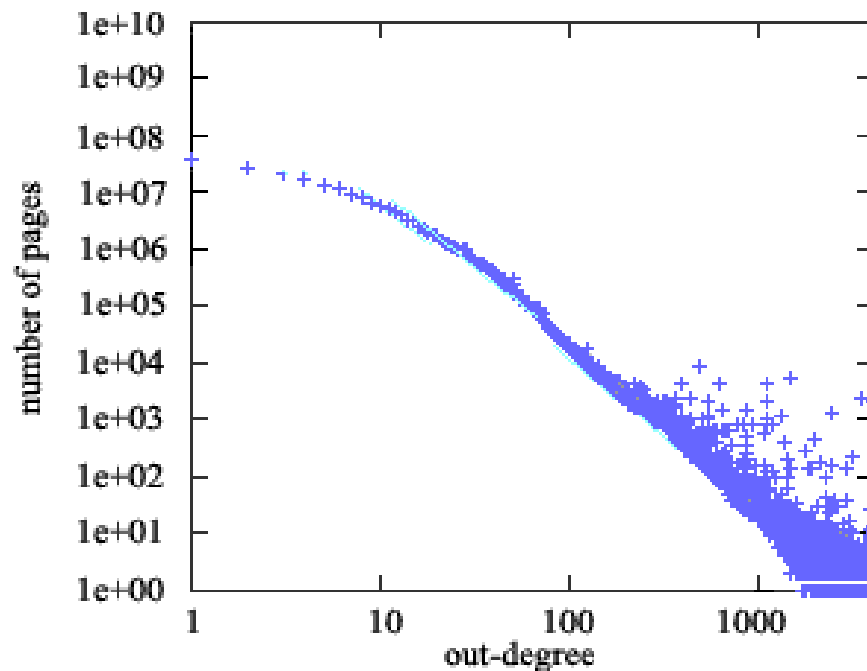
Call Graph



Distributions of outdegrees (a), indegrees (b)

Web Graph

Broder et al. used two Altavista crawls, each with about 200 million pages and 1.5 billion links in their experiments.



Number of vertices with various out-degrees (left) and distribution of sizes of strongly connected components (right) in Web graph.



Power-Law Graphs

- Barabási et al. (1999): $\gamma \approx 2.45$ for outdegrees, $\gamma \approx 2.1$ for indegrees of Web graph.
- Aiello et al. (1999): $\gamma \approx 2.1$ for Call graph.
- **Market graph:** $\gamma \approx 0.68(\theta = 0.3); 0.83(\theta = 0.4)$.
- **Complementary market graph:**
 $\gamma \approx 1.41(\theta \in [-0.2, -0.15])$.



Market Graph

- **Cliques** in the market graph represent classes of stocks whose prices **exhibit similar behavior** over time, *i.e.*, a stock in a clique is correlated with **all** other stocks in this clique.
- **Independent sets** in the market graph represent **diversified portfolios**, where every stock is negatively correlated with all other stocks in the independent set.
- Finding cliques and independent sets provides a new tool of classification of stocks and diversified portfolio selection.



Market Graph

To find a large clique, we apply the following greedy algorithm. Let us denote by $N(i) = \{j | (i, j) \in E\}$ the set of neighbors of i in $G = (V, E)$.

$$C = \emptyset, G_0 = G;$$

do

$$G_0 = \bigcap_{i \in C} N(i) \setminus C;$$

$C = C \cup j$, where j is a vertex of largest degree in G_0 ;

until $G_0 = \emptyset$.



Market Graph

$G'(V', E')$ is the graph induced by vertices remaining after recursively removing from the graph all of the vertices which are not in C and whose degree is less than $|C|$.

θ	edge density	clique size	$ V' $	edge dens. in G'
0.35	0.0090	168	535	0.6494
0.4	0.0047	104	405	0.6142
0.45	0.0024	109	213	0.8162
0.5	0.0013	84	146	0.8436
0.55	0.0007	61	102	0.8701
0.6	0.0004	45	70	0.8758
0.65	0.0002	23	80	0.5231
0.7	0.0001	21	33	0.7557



Market Graph

Edge formulation of the Maximum Clique Problem:

$$\begin{aligned}\omega(G) = \max \quad & \sum_{i=1}^n x_i \\ \text{s.t.} \quad & x_i + x_j \leq 1, \forall (i, j) \notin E, \\ & x_i \in \{0, 1\}, i = 1, \dots, n.\end{aligned}$$

This integer program can be solved using modern commercial software packages.



Market Graph

θ	edge density	$\omega(G)$
0.35	0.0090	193
0.4	0.0047	144
0.45	0.0024	109
0.5	0.0013	85
0.55	0.0007	63
0.6	0.0004	45
0.7	0.0001	22

Clique number of the market graph for different values of the correlation threshold

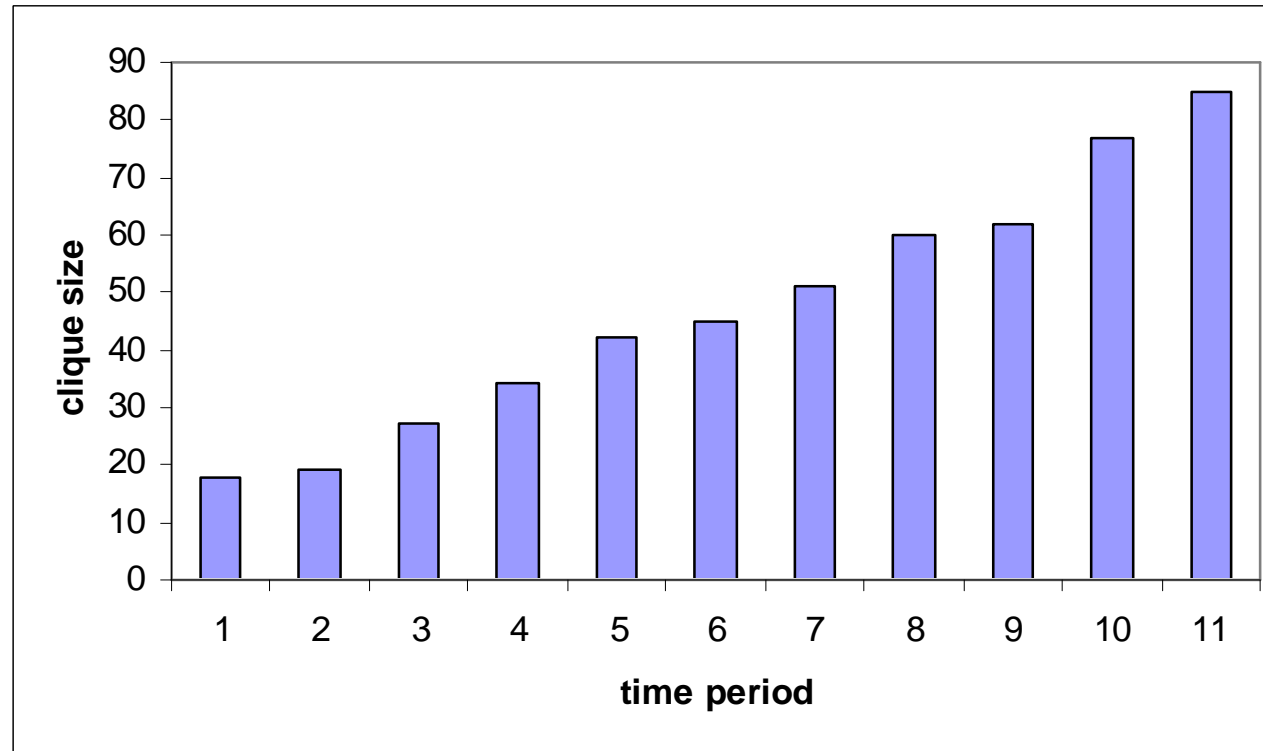


Market Graph

θ	edge density	clique size
0.05	0.5206	1369
0.1	0.2837	943
0.15	0.1469	651
0.2	0.0741	461
0.25	0.0364	341
0.3	0.0178	257

Size of a clique found in the market graph for different values of the correlation threshold

Market Graph



Evolution of the maximum clique size in the market graph
($\theta = 0.5$)



Market Graph

θ	edge density	indep. set size
0.05	0.4794	36
0.0	0.2001	11
-0.05	0.0431	5
-0.1	0.005	3
-0.15	0.0005	2

Size of an independent set found in the market graph and density of the complementary graph for different values of the correlation threshold



Market Graph

Clustering coefficients of the market graph (* - complementary graph)

θ	edge density	clustering coef.
-0.15*	0.0005	2.64×10^{-5}
-0.1*	0.0050	0.0012
0.3	0.0178	0.4885
0.4	0.0047	0.4458
0.5	0.0013	0.4522
0.6	0.0004	0.4872
0.7	0.0001	0.4886



Market Graph

- It turns out that the sizes of independent sets detected in the market graph are significantly smaller than clique sizes, which indicates that one would not expect to find a large diversified portfolio in the modern stock market, but there are a lot of stocks exhibiting similar behavior.
- This fact provides another argument in support of the idea of the globalization of the stock market.



Market Graph

symbol	degree	symbol	degree
QQQ	216	IVW	156
IWF	193	IAH	155
IWO	193	IYY	154
IYW	193	IWB	153
XLK	181	IYV	150
IVV	175	BDH	144
MDY	171	MKH	143
SPY	162	IWM	142
IJH	159	IJR	134
IWV	158	SMH	130

20 instruments with highest degrees ($\theta = 0.6$).



Market Graph

- Stocks with high degrees in the market graph are correlated with many other stocks, which implies that high-degree stocks **most accurately reflect the behavior of the market**.
- All high-degree vertices correspond to **Exchange Traded Funds** (stocks corresponding to *indices*).



Concluding Remarks

- The evolution of the market graph over time confirms the idea of globalization of the stock market
- The degree distribution of the market graph remains stable and has a highly specific structure. In this sense, the stock market can be considered as a “self-organized” system
- Independent sets and sparse subgraphs (quasi-independent sets) – *diversified portfolios*
- Clique partitioning – a natural approach to the *classification* of financial instruments
- Other problems?



References

- V. Boginski, S. Butenko and P. M. Pardalos. On Structural Properties of the Market Graph, in: *Innovations in Financial and Economic Networks* (A. Nagurney, editor), Edward Elgar Publishers, 28–45, 2003.
- V. Boginski, S. Butenko and P. M. Pardalos. Statistical Analysis of Financial Networks. *Computational Statistics and Data Analysis*, to appear in 2004. Available online at www.sciencedirect.com
- V. Boginski, S. Butenko, and P.M. Pardalos. Mining Market Data: A Network Approach, submitted to *Computers and Operations Research*, 2004.



References

- V. Boginski, S. Butenko and P. Pardalos, “Modeling and Optimization in Massive Graphs”, *Novel Approaches to Hard Discrete Optimization* (P.M. Pardalos and H. Wolkowicz, editors) Fields Institute Communications Series, American Math. Society, pp. 17-39, 2003.
- J. Abello, P.M. Pardalos and M.G.C. Resende, **Handbook of Massive Data Sets**, Kluwer Academic Publishers, 2002.
- J. Abello, P.M. Pardalos and M.G.C. Resende, “*On maximum clique problems in very large graphs*”, In **DIMACS Series 50**: 119-130, 1999.