Networks In Finance and Economics

Panos M. Pardalos

Center for Applied Optimization (CAO)
University of Florida, USA
www.ise.ufl.edu/pardalos

Laboratory of Algorithms and Technologies for Networks Analysis (LATNA)
National Research University Higher School of Economics, Russia
http://nnov.hse.ru/en/latna/

Section 1

Introduction

Financial Markets and Networks

Financial markets, banks, currency exchanges and other institutions can be modeled and analyzed as network structures where

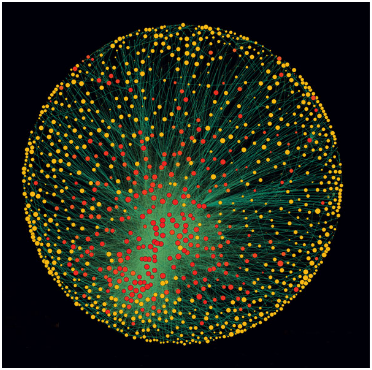
- Nodes are any agents such as companies, shareholders, currencies, countries, etc.
- Edges (can be weighted, oriented, etc.) represent any type of relations between agents, for example, ownership, friendship, collaboration, influence, dependence, and correlation.

We are going to discuss network and data sciences techniques to study the dynamics of financial markets and other problems in economics.

Motivation

- What are the scientific insights into the causes and effects (benefits and drawbacks) of globalization/deglobalization?
- Economic data is massive (financial markets, banks, industries etc). Can we use this data to study the dynamics of global economic interactions?
- Although a-posteriori analysis is more tractable, can we develop tools that predict the dynamics, or identify control parameters, or help design new policies?

"The capitalist network that runs the world"

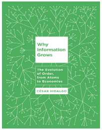


The 1318 transnational corporations that form the core of the economy. Superconnected companies are red, very connected companies are yellow. The size of the dot represents revenue (Image: PLoS One)

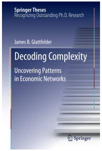
S. Vitali, J. B. Glattfelder, S. Battiston, The Network of Global Corporate Control, PloS ONE, vol. 6, is. 10, e25995, 2011.

Books

- Hidalgo, C. **Why information grows. The evolution of Order, from Atoms to Economies**, New York: Basic Books (2015)



- Glattfelder J.B. **Decoding complexity: uncovering patterns in economic networks,**, Heidelberg: Springer; (2013)



Section 2

Basic definitions

Similarity measure

- Use closure prices of stocks traded on a market for a certain period of time
- $P_i(t)$ is the price of the stock i by the end of the day t
- $R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)}$ is the stock return at the day t
- $\rho_{ij} = \frac{E[(R_i - E[R_i])(R_j - E[R_j])]}{\sqrt{V(R_i)V(R_j)}}$ is the correlation coefficient between the stocks i and j
- $\rho = [\rho]_{n \times n}$ is the correlation matrix which we use as a similarity matrix

Market Graph

V. Boginski, S. Butenko, and P. M. Pardalos, Modeling and optimization in massive graphs. In: Novel Approaches to Hard Discrete Optimization, American Mathematical Society, 17-39, 2003.

A market graph can be constructed as follows (for a stock market):

- each stock is represented by a vertex
- two vertices are connected if the correlation coefficient of the returns of the corresponding pair of stocks is above a predefined threshold $\theta \in [0;1]$

Power Law

Many real-life graphs possess the so-called Power Law property.

The graph $G = (V, E)$ is the Power Law graph if

$$y = e^{\alpha}/x^{\beta} \tag{1}$$

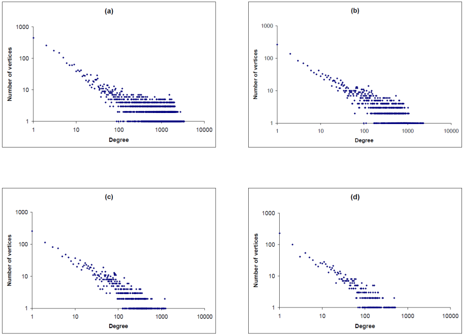
or, equivalently,

$$\log y = \alpha - \beta \log x \tag{2}$$

where y is the number of nodes with degree x.

Power Law in a Market Graph

V. Boginski, S. Butenko, and P. M. Pardalos, On structural properties of the market graph. In: Innovation in Financial and Economic Networks, Edward Elgar Publishers, London, 29-45, 2003.



Degree distribution of the USA market graph with 6546 nodes for (a) $\theta = 0.2$; (b) $\theta = 0.3$; (c) $\theta = 0.4$; and (d) $\theta = 0.5$.

Minimum/Maximum Spanning Tree

- One of the first network structures suggested for analysis of financial markets

R. N. Mantegna, Hierarchical structure in financial markets, Eur. Phys.I J. B, Volume 11, 1999, pages 193-197.

- Can be applied for studying hierarchical structures or the most influential companies on a market.
- Comes as a solution of the well-known Minimum Spanning Tree (MST) problem: For a given weighted undirected connected graph $G = (V, E)$, w_{ij} is the weight of $(i, j) \in E$, find a subgraph such that:
  - It is a tree (connected acyclic graph)
  - Has the smallest possible total sum of the edges weights

Maximum Planar Graph

The logical extension of the MST approach

- Comes as a solution of the Maximum Planar Graph problem: For a given weighted undirected graph $G = (V, E)$, w_{ij} is the weight of $(i, j) \in E$, find a subgraph such that:
  - It is a planar (can be drawn in 2D space without intersection of edges)
  - Has the largest possible sum of edges.
- The most important advantage is that it take into account much more information about a market
- The biggest disadvantage is that it is NP-hard problem

Clique

For a given undirected graph $G = (V, E)$ a clique is any subgraph of G such that it has all possible edges, i.e. all nodes are pairwise connected.

Clique-related problems

All problems are NP-hard.

- Maximum Clique Problem: find the largest clique for a given undirected graph $G = (V, E)$
- Clique Cover Problem: given an undirected graph $G = (V, E)$ find the partitioning of its nodes into the minimum possible number of subsets such that the subgraphs induced by these sets are cliques.
- Clique Listing Problem: find all maximal (cannot add a vertex add still have a clique) cliques of a graph.

Independent Set

- An independent set or stable set is a set of vertices in a graph, no two of which are adjacent.
- Maximum Independent Set Problem: for a given graph find an independent set with the largest possible size.
- Has applications in the Portfolio Selection Problem

Section 3

Dynamics

Cluster structure

- We say that the set X of n agents (stocks) was clustered (partitioned) into p clusters X_i $i = 1, \dots, p$ if
  - $X_i \neq \emptyset \ \forall i$
  - $X_i \cap X_j = \emptyset \ \forall i \neq j$
  - $\cup_{i=1}^{p} X_i = X$
- Among all possible partitionings we look for that one having such a property that stocks inside clusters have the strongest similarity with medians (i.g. centers) of them.
- Such a clustering we call a cluster structure.

Similarity measure

- We use closure prices of stocks traded on a market for a certain period of time
- $P_i(t)$ is the price of the stock i by the end of the day t
- $R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)}$ is the stock return at the day t
- $\rho_{ij} = \frac{E[(R_i - E[R_i])(R_j - E[R_j])]}{\sqrt{V(R_i)V(R_j)}}$ is the correlation coefficient between the stocks i and j
- $\rho = [\rho]_{n \times n}$ is the correlation matrix which we use as a similarity matrix

Clustering by the p-Median Approach

We are ready to find the best clustering (in the sense defined earlier):

- Let S be a subset of X such that $|S| = p$
- We formulate the optimization problem

$$\max_{S \subset X, |S|=p} \left(\sum_{i=1}^{n} \max_{j \in S} (\rho_{ij}) \right) \tag{3}$$

- After the transformation $d_{ij} = 1 - \rho_{ij}$ we get the equivalent P-Median problem

$$\min_{S \subset X, |S|=p} \left(\sum_{i=1}^{n} \min_{j \in S} (d_{ij}) \right) \tag{4}$$

which is known to be NP-hard

Clustering by the *p*-Median Approach. Example

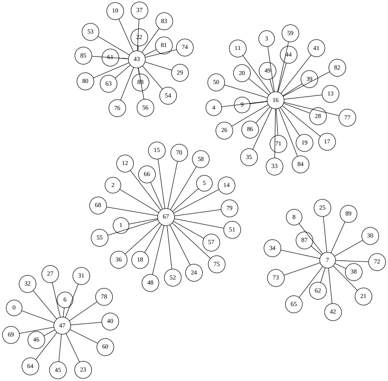


Figure: The 5-cluster structure for the S&P100 Index in 2001.

Studying the dynamics

- Let $G_{p,t}$ be a p-cluster (having p clusters) structure obtained for a time period $t = 1, \ldots, T$ where all observations are divided into T non-overlapping consecutive time periods
- To evaluate the dynamics of the cluster structures we compare all time-neighboring p-cluster structures $G_{p,t}$ and $G_{p,t+1}$
- Thus we have n sequences $\{(G_{p,1}, G_{p,2}), ..., (G_{p,T-1}, G_{p,T})\}$ $\forall p = 1, ..., n$
- For each pair $(G_{p,t}, G_{p,t+1})$ we assign a number which reflects a similarity of these two structures

Similarity between cluster structures

The first similarity measure $\alpha_{p,t}$ is defined as:

$$\alpha_{p,t} = \sum_{1 \leq i < j \leq n} \frac{x_{i,j}^{p,t}}{n(n-1)/2} \tag{5}$$

where $x_{i,j}^{p,t} = 1$ if the stocks i and j were in the same cluster (different clusters) of $G_{p,t}$ and now are the same cluster (different clusters) of $G_{p,t+1}$. This variable equals 0, otherwise.

In other words, we count a proportion of pairs of stocks which either are together in one cluster or remain separate.

Similarity between cluster structures

The second characteristic $\beta_{p,t}$ reflects the behavior of groups of stocks.

- We assume that over the time some stocks in clusters can roam but others stay together and form cores inside each cluster
- We count stocks in the cores and then divide this number by the total number of stocks
- We use the Jonker-Volgenant algorithm in order to find these cores

Results for the USA market

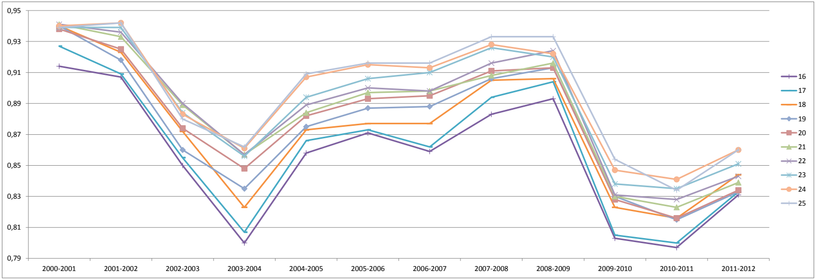


Figure: Dynamics of the cluster structures for $\alpha_{p,t}$. **S&P100 Index**. 10 trends for $p = 16, 17, ..., 25$ number of clusters.

Results for the USA market

Figure: Dynamics of the cluster structures for $\beta_{p,t}$. **S&P100 Index**. 10 trends for $p = 16, 17, ..., 25$ number of clusters.

Results for the USA market

This approach was able to detect two financial crises on the USA market: the Dot-com bubble crisis of the years 2000-2002 and the global economic crisis of the years 2007-2009.

A. Kocheturov, M. Batsyn, P. M. Pardalos, Dynamics of cluster structures in a financial market network, Physica A: Statistical Mechanics and its Applications, Volume 413, 1 November 2014, Pages 523-533.

Modularity-Based Clustering Approach

- Modularity by Newman, 2004 is a common measure of how "good" the clustering is.
- By "good" we mean how far it is from the random partitioning.
- $Q = \textit{the proportion of edges that fall within clusters} - \textit{the expected such number if the edges were distributed at random}$

Modularity-Based Clustering Approach

Q can be rewritten in such a way:

$$Q = \frac{1}{2sum} \sum_i \sum_j (\rho_{ij} - \frac{\rho_i \rho_j}{2sum}) \delta(i,j), \tag{6}$$

where $2sum = \sum_i \sum_j \rho_{ij}$, $\rho_i = \sum_k \rho_{ik}$ and $\delta(i,j)$ is equal to 1 if the stocks i and j are in the same module (cluster), 0 otherwise.

Modularity-Based Clustering Approach

In order to find the best partitioning we define and solve the following optimization problem:

$$\max \frac{1}{2sum} \sum_i \sum_j (\rho_{ij} - \frac{\rho_i \rho_j}{2sum}) \delta(i,j) \quad (7)$$

$$s.t.\ \delta(i,i) = 1\ \forall i \quad (8)$$

$$\delta(i,j) = \delta(j,i)\ \forall i,j \quad (9)$$

$$\delta(i,j) = 1,\ \delta(j,k) = 1 \implies \delta(i,k) = 1\ \forall i,j,k \quad (10)$$

Modularity. Results for the USA market

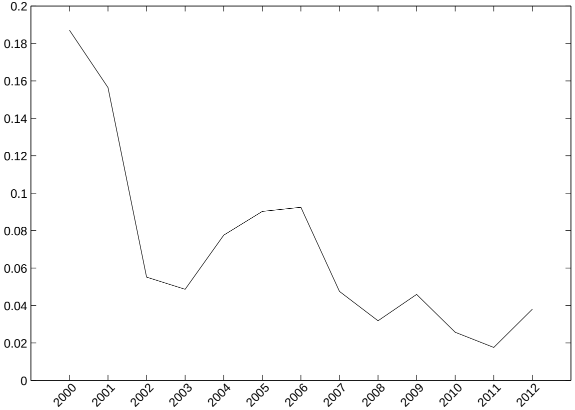


Figure: The modularity Q values for the **S&P100 Index**. From Kocheturov et al., 2014

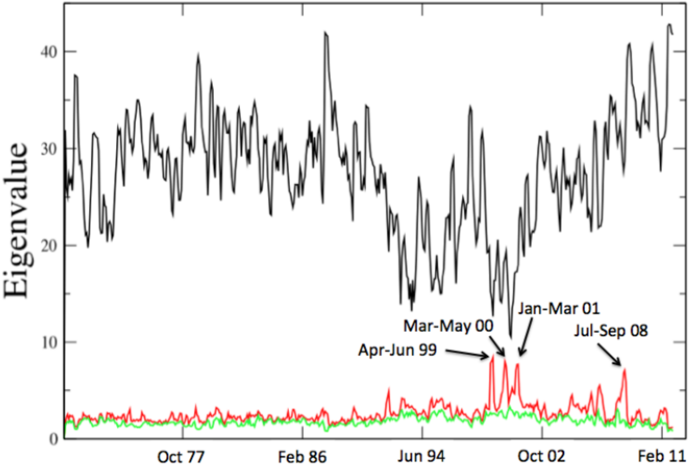
Dynamics of Plenary Maximally Filtered Graphs

G. Buccheri, S. Marmi, and R. N. Mantegna, Evolution of correlation structure of industrial indices of US equity markets, Phys. Rev. E, volume 88 (1), 2013.

- The authors used correlations between industry portfolios of US equity markets
- They studied monthly dynamics of plenary maximally filtered graphs (PMFGs)
- by investigation of the dynamics of the largest eigenvalues and eigenvectors associated with monthly correlation matrices

Dynamics of Plenary Maximally Filtered Graphs

The second eigenvalue (in red) revealed the same crises:



Market Graph Analysis Tools

- Minimum Spanning Tree.
- Planar Maximally Filtered Graph.
- Maximum cliques and clique partitions.
- Maximum independent sets.

Finding Cliques in the Market Graph

- Applying a heuristic algorithm to find a large clique: let $N(i)$ be the set of neighbors of the vertex i:

  $C = \emptyset, G_0 = G;$

  do

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

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

  until $G_0 = \emptyset$.

Finding Cliques in the Market Graph

- Using the IP formulation of the maximum clique problem to find the exact solution:

$$\text{maximize} \quad \sum x_i$$

$$s.t.$$

$$x_i + x_j \leq 1, (i,j) \notin E'$$

$$x_i \in \{0,1\}$$

Maximum Clique size for different correlation thresholds

- Large cliques despite very low edge density - confirms the idea about the "globalization" of the market

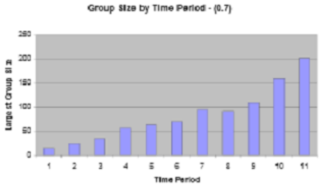
θ	edge density	clique size
0.35	0.0090	193
0.4	0.0047	144
0.45	0.0024	109
0.5	0.0013	85
0.55	0.0007	63
0.6	0.0004	45
0.65	0.0002	27
0.7	0.0001	22

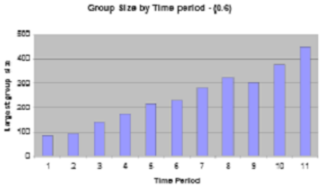
Classification of Stocks Using Clique Partitioning

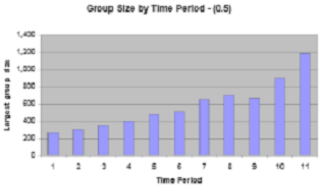
- A clique in the market graph represents a dense cluster of stocks whose prices exhibit a similar behavior over time
- Therefore, dividing the market graph into a set of distinct cliques (clique partitioning) is a natural approach to classifying stocks (dividing the set of stocks into clusters of similar objects an approach to solve the clustering problem)

Connected Components in Market Graph

- Largest Group size by Time Period







Connected Components in Market Graph

Observations

- The increase in the giant component size from oldest to newest time period indicates the globalization tendency, just as in maximum clique size and edge density
- The giant component includes semiconductor industries and the increase in the size of the giant component corroborates the observation that the number of these industries has been increasing with time

Independent Sets in the Market Graph

- Maximum independent set represents the largest "perfectly diversified" portfolio
- Solving the maximum clique problem in the complementary graph
- The preprocessing procedure could not reduce the size of the initial graph, the exact solution could not be found
- Large diversified portfolios are hard to find

Independent set sizes for different correlation thresholds

Relatively small independent sets found by the heuristic algorithm

θ	edge density	indep. set size
0.05	0.4794	36
0.0	0.2001	12
-0.05	0.0431	5
-0.1	0.005	3
-0.15	0.0005	2

Independent Sets in the Market Graph

- Finding a perfectly diversified portfolio containing **any given stock**
- For every vertex in the market graph, an independent set that contains this vertex was detected, and the sizes of these independent sets were almost the same, which means that it is possible to find a diversified portfolio containing any given stock using the market graph methodology.

Analysis of the Russian Stock Market

- The market network constructed using **correlation** as a measure of similarity between stocks.
- We considered **11 shifted periods of 500-day each** from September 1, 2007 to September 16, 2011
- Results are surprising:
  - Russian stock market is dominated by a few **highly correlated** stokes with the biggest value.
  - The nodes of the maximum clique for the threshold are 9 most valuable stocks.
  - The stocks in the clique account for 89% of the total value of the market.
  - The most valuable stocks have the strongest connections between their return.

References

V. Boginski and S. Butenko and P.M. Pardalos,
On structural properties of the market graph, in Innovations in financial and economic networks, Edward Elgar Publishers, 2003

V. Boginski and S. Butenko and P. M. Pardalos,
Statistical analysis of financial networks, Computational Statistics and Data Analysis, 2005, pp. 431443

V. Boginski and S. Butenko and P. M. Pardalos,
Mining market data: A network approach, Computers & Operations Research, 2006, pp. 3171-3184

A. Koldanov and P. Koldanov and V. Kalyagin and P.M. Pardalos,
Statistical procedures for the market graph construction, Computational Statistics & Data Analysis, 2013

G. Bautin and V. Kalyagin and A. Koldanov and P. Koldanov and P.M. Pardalos,
Simple measure of similarity for the market graph construction, Computational Management Science, 2013, pp. 105-124

References (cont.)

- A. Vizgunov and B. Goldengorin and V. Kalyagin and A. Koldanov and P. Koldanov and P.M. Pardalos, **Network approach for the Russian stock market**, Computional Management Science, January 2014, Volume 11, Issue 1-2, pp. 45-55
- V.A. Kalyagin, A.P. Koldanov, P.A. Koldanov, P.M. Pardalos, V.A. Zamaraev, **Measures of uncertainty in market network analysis**, Physica A: Statistical Mechanics and its Applications, Volume 413, 1 November 2014, Pages 59-70
- Anton Kocheturov, Mikhail Batsyn, Panos M. Pardalos, **Dynamics of cluster structures in a financial market network**, Physica A: Statistical Mechanics and its Applications, Volume 413, 1 November 2014, Pages 523-533
- Valery A. Kalygin, Alexander P. Koldanov, Panos M. Pardalos, **A General Approach to Network Analysis of Statistical Data Sets**, In Lecture Notes in Computer Science 8426, Springer 2014, pp. 88-97.

References (cont.)

A. Kocheturov, M. Batsyn and P. Pardalos, **Dynamics of Cluster Structures in Stock Market Networks**, Journal of the New Economic Association, 2015, vol. 28, issue 4, pp. 12-30

Anton Kocheturov, Panos M. Pardalos, **Data Science for Massive Networks**, In: P. Braslavski, I. Markov, P. Pardalos, Y. Volkovich, D. I. Ignatov, S. Koltsov, O. Koltsova (Eds.). Information Retrieval, Volume 573, 2016, Pages 88-100

T.A. Schieber, L. Carpi, A.C. Frery, O.A. Rosso, P.M. Pardalos, and M.G. Ravetti, **Information Theory Perspective on Network Robustness**, Physics Letters A, Volume 380, Issue 3, January 2016, pp. 359-364

V. A Kalyagin, A. P Koldanov, P. M Pardalos, **On multivariate network analysis of statistical data sets with different measures of association**, Annals of Mathematics and Artificial Intelligence, Vol 76, No 1-2 (February 2016), pp. 83-92

References (cont.)

Koldanov P., Komissarova A., **Statistical uncertainty of minimum spanning tree in market network**, In bk.: Models, Algorithms, and Technologies for Network Analysis. NY: Springer International Publishing, 2016. P. 157-164

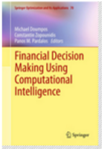
Latyshev A., Koldanov P, **Testing of connections between Pearson and sign correlations in market network**, In bk.: Models, Algorithms, and Technologies for Network Analysis. NY: Springer International Publishing, 2016. P. 175-182.

Kalyagin VA, Koldanov AP, Koldanov PA, Pardalos PM., **Optimal decision for the market graph identification problem in a sign similarity network**,Annals of Operations Research. 2017:1-5.

Selected Books

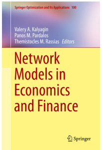
- C. Zopounidis, M. Doumpos and P. Pardalos, **Handbook of Financial Engineering**, Springer Optimization and Its Applications Vol. 18, Springer, 2008

- M. Doumpos, C. Zopounidis and P. Pardalos, **Financial Decision Making Using Computational Intelligence**, Springer Optimization and Its Applications Vol. 70, Springer, 2012

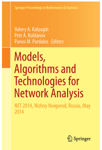


Selected Books

- V.A. Kalyagin, P.M. Pardalos, T.M. Rassias (Eds.) **Network Models in Economics and Finance**, Springer Optimization and Its Applications Vol. 100, Springer, 2014



- V.A. Kalyagin, P.A. Koldanov, P.M. Pardalos, (Eds.) **Models, Algorithms and Technologies for Network Analysis**, Springer Proceedings in Mathematics & Statistics, Springer, 2016



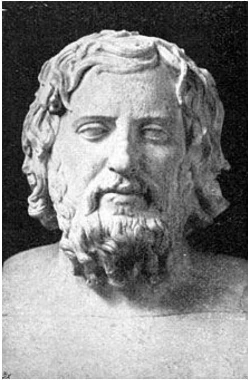
Section 4

Concluding remarks

Concluding remarks

What about a new measure of national wealth that includes quality of life and sustainability of wealth?

Concluding remarks



The Oeconomicus (Greek:) by Xenophon (c. 430-354 BC) is one of the earliest works on economics.

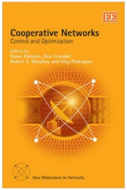
The opening framing dialogue is between Socrates and Critoboulus, the son of Crito. There Socrates discusses the meaning of wealth and identifies it with usefulness and well-being, not merely possessions.

Concluding remarks

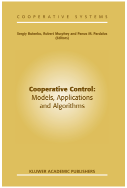
- Network theory can be a complimentary tool to understand the dynamics of financial markets.
- However, we need more sophisticated tools of co-operative networks to understand the complexities of new economies.

Concluding remarks

- P.M. Pardalos, D. Grundel, R.A. Murphey, O. Prokopyev (Eds.)
  Cooperative Networks: Control and Optimization, New Dimensions in Networks, Edward Elgar Pub, 2008



- S. Butenko, R. Murphey, P.M. Pardalos (Eds.), **Cooperative Control: Models, Applications and Algorithms**, Cooperative Systems, Kluwer Academic Publishers, 2003



Concluding remarks

Amini MH, Boroojeni KG, Iyengar SS, Pardalos PM, Blaabjerg F, Madni AM. **Sustainable Interdependent Networks**, Springer (2018)