Introduction to data analytics for networks – a historical perspective and major advances

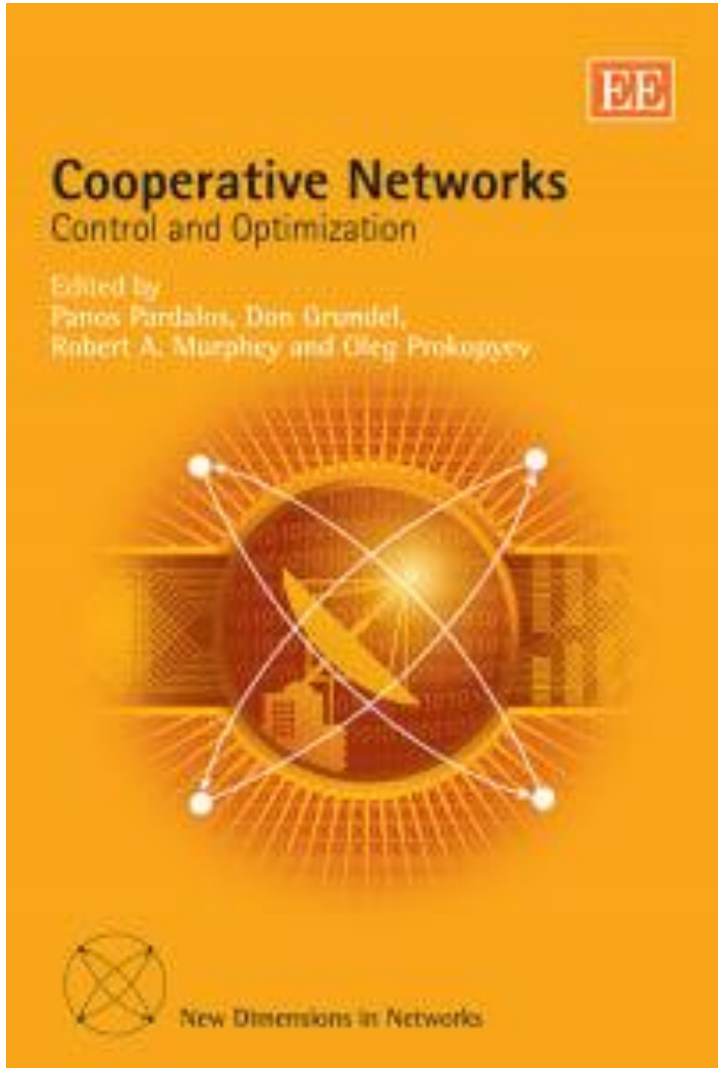
Panos M. Pardalos

Distinguished Emeritus Professor

University of Florida

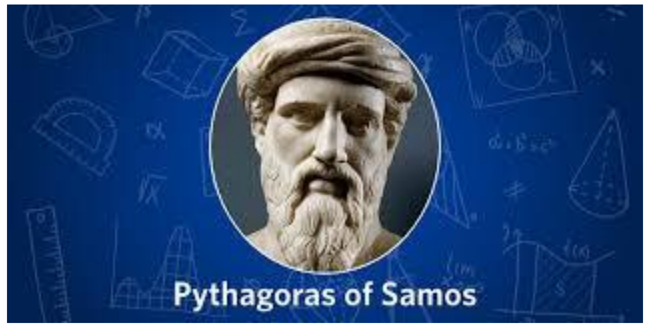
http://www.ise.ufl.edu/pardalos/

Human Cooperative Networks



Animal Cooperative Networks

- The oldest, shortest words— "yes" and "no"— are those which require the most thought



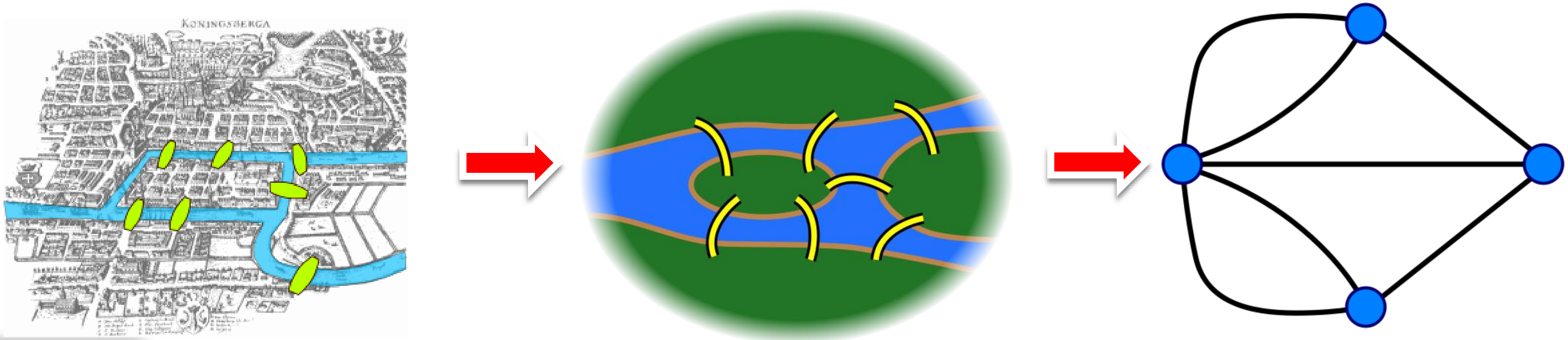
- Pythagoras of Samos (Πυθαγόρας; c. 570 BC – c. 496 BC)

- Pythagoras was the first person to call the universe a Cosmos. The Greek term, which is the root of the word cosmetic, refers to an equal presence of order and beauty

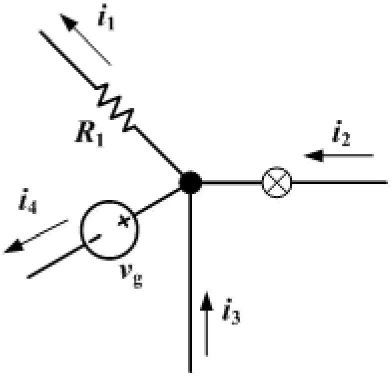
From a single network to network of networks

- In the 18th century Euler solved the famous Königsberg bridge problem.

- Euler's solution is considered to be the first theorem of network analysis and graph theory.



- In the 19th century Gustav Kirchhoff initiated the theory of electrical networks.
- Kirchhoff was the first person who defined the flow conservation equations, one of the milestones of network flow theory.
- After the invention of the telephone by Alexander Graham Bell in the 19th century the resulting applications gave the network analysis a great stimulus



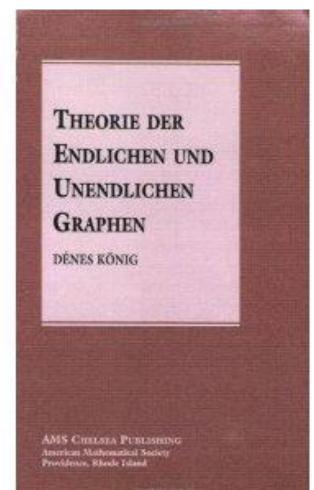
- In the nineteen century Gustav Kirchhoff initiated the theory of electrical networks.
- Kirchhoff was the first person who define the flow conservation equations, one of the milestones of network flow theory.

- After the invention of the telephone in the 19th century by Alexander Graham Bell, the resulting applications gave the network analysis a great stimulus.
- Nicola Tesla and the modern smartphone (1926) https://bigthink.com/words-of-wisdom/nikola-tesla-2/

The field evolved dramatically after the 19th century.

- The first graph theory book was written by Dénes König in 1936.

- As in many other fields, WWII played a crucial role in the development of the field.
- Many algorithms and techniques were developed to solve logistic problems from the military.
- After the war, these technological advances were applied successfully in other fields.



The field evolved dramatically after the 19th century.

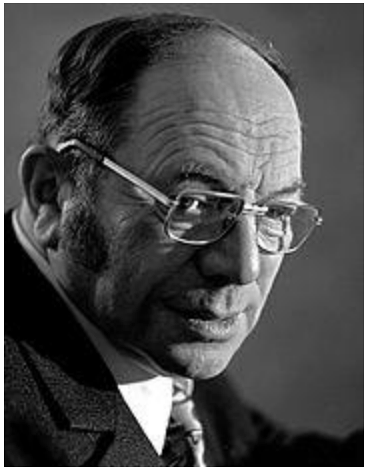
- The first graph theory book was written by Dénes König in 1936.

- As in many other fields, WWII played a crucial role in the development of the field.
- Many algorithms and techniques were developed to solve logistic problems from the military.
- After the war, these technological advances were applied successfully in other fields.

The field evolved dramatically after the 19th century.

- The earliest linear programming model was developed by Leonid Kantorovich in 1939 during World War II, to plan expenditures to reduce the costs of the army.
- In 1940, also during World War II, Tjalling Koopmans formulated also linear optimization models to select shipping routes to send commodities from America, to Specific destinations in England.

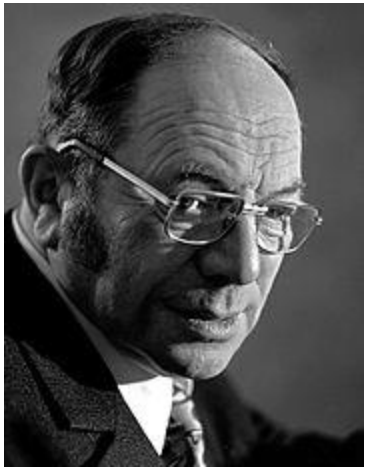
- For their work in the theory of optimum allocation of resources, these two researchers were awarded with the Nobel price in Economics in 1975.

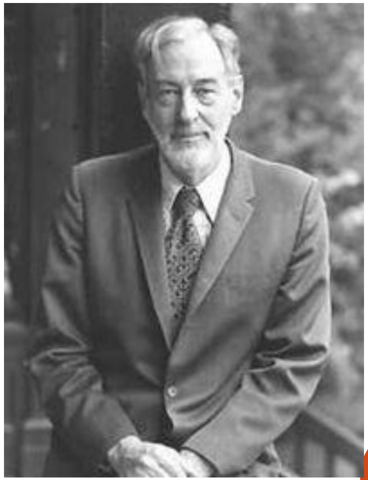


The field evolved dramatically after the 19th century.

- The earliest linear programming model was developed by Leonid Kantorovich in 1939 during World War II, to plan expenditures to reduce the costs of the army.
- In 1940, also during World War II, Tjalling Koopmans formulated also linear optimization models to select shipping routes to send commodities from America, to Specific destinations in England.

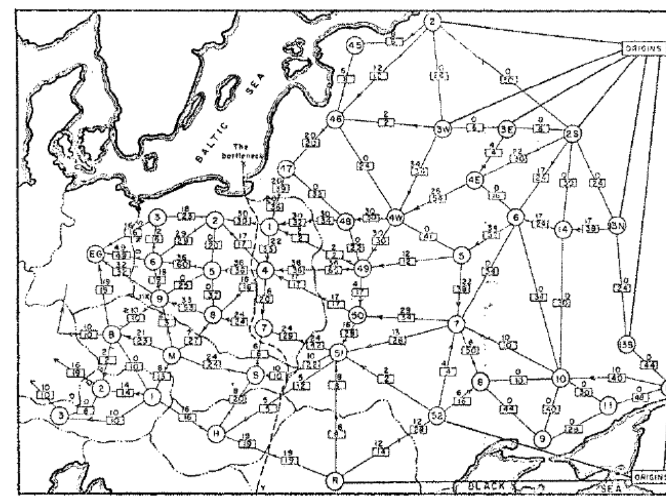
- For their work in the theory of optimum allocation of resources, these two researchers were awarded with the Nobel price in Economics in 1975.





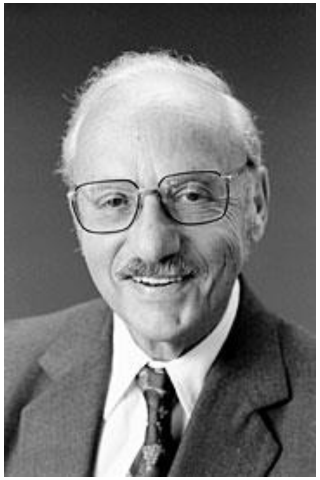
- The first complete algorithm for solving the transportation problem was proposed by Frank Lauren Hitchcock in 1941. Is this true?

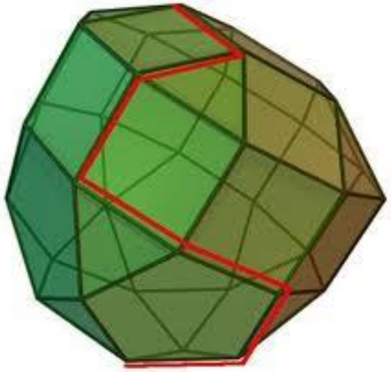
- With the development of the Simplex Method for solving linear programs by George B. Dantzig in 1957, a new framework for solving several network problems became available.
- The network simplex algorithm is still in practice one of the most efficient algorithms for solving network flow problems.



- The first complete algorithm for solving the transportation problem was proposed by Frank Lauren Hitchcock in 1941.

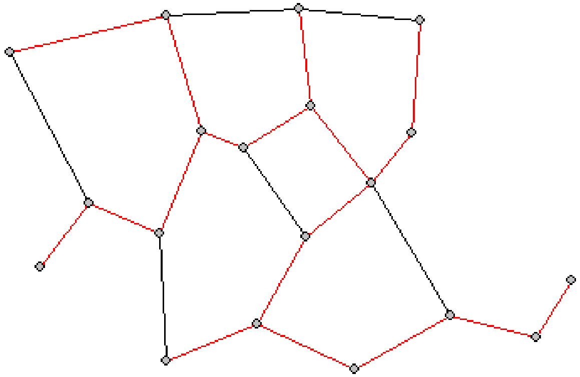
- With the development of the Simplex Method for solving linear programs by George B. Dantzig in 1957, a new framework for solving several network problems became available.
- The network simplex algorithm is still in practice one of the most efficient algorithms for solving network flow problems.



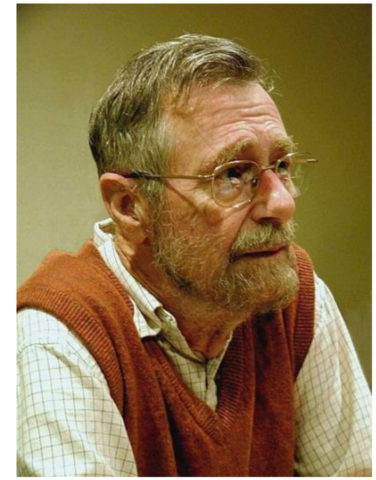


- Many other authors proposed efficient algorithms for solving different network problems.
- Joseph Kruskal in 1956 and Robert C. Prim in 1957 developed algorithms for solving the minimum spanning tree problem.

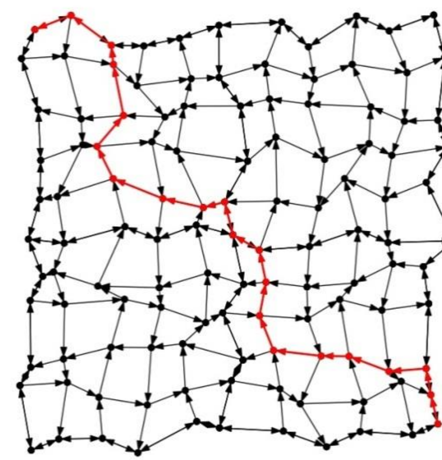
- In 1956 Edsger W. Dijkstra developed his algorithm for solving the shortest path problem, one of the most recognized algorithms in network analysis.



- Many other authors proposed efficient algorithms for solving different network problems.
- Joseph Kruskal in 1956 and Robert C. Prim in 1957 developed algorithms for solving the minimum spanning tree problem.

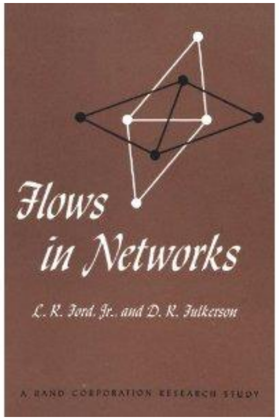


- In 1956 Edsger W. Dijkstra developed his algorithm for solving the shortest path problem, one of the most recognized algorithms in network analysis.



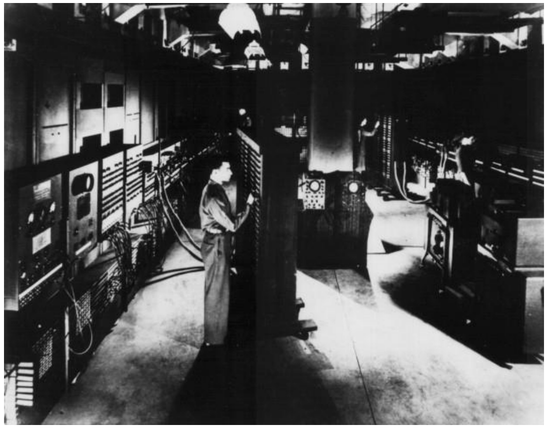
The field evolved dramatically after the 19th century.

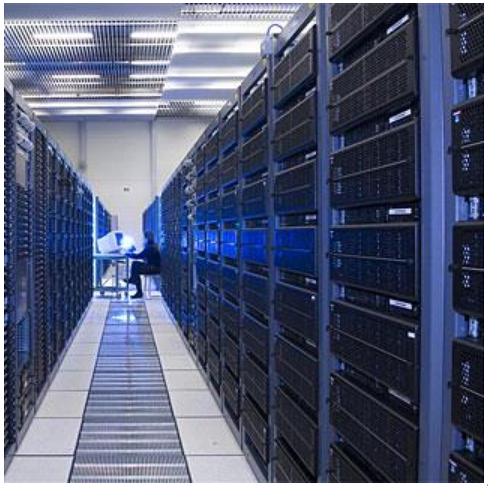
- As it happened in other fields, computer science and networks influenced each other in many aspects.
- In 1963 the book by Lester R. Ford and Delbert R. Fulkerson introduced new developments in data structures techniques and computational complexity into the field of networks.



In recent years the evolution of computers have changed the field. We are now able to analyze and solve large-scale network problems.

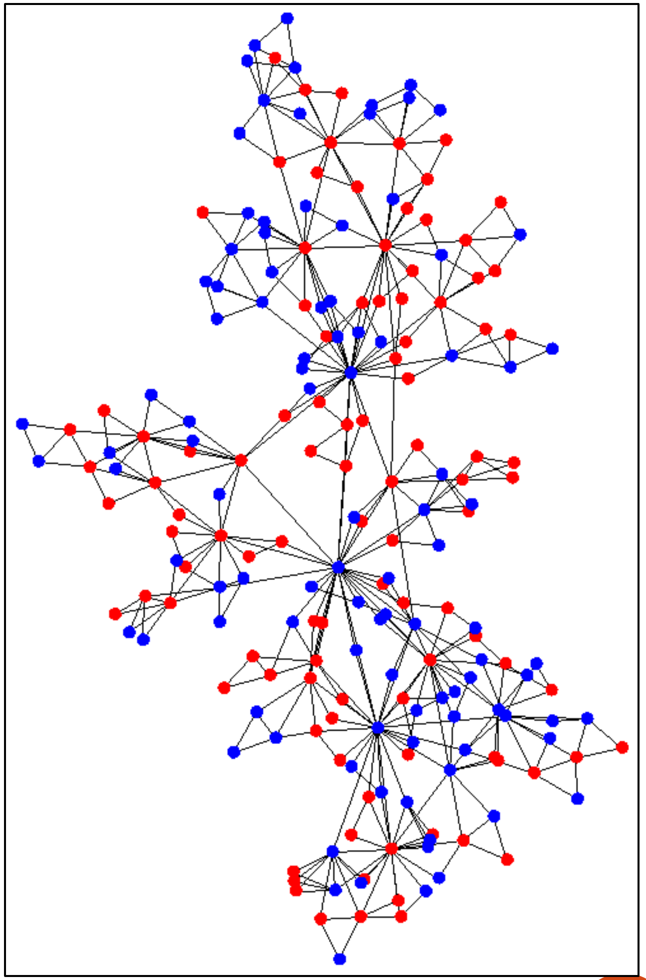
- Parallel computing
- Grid computing
- Cloud/Skype computing
- GPU (Graphics Processing Unit)
- Quantum computing

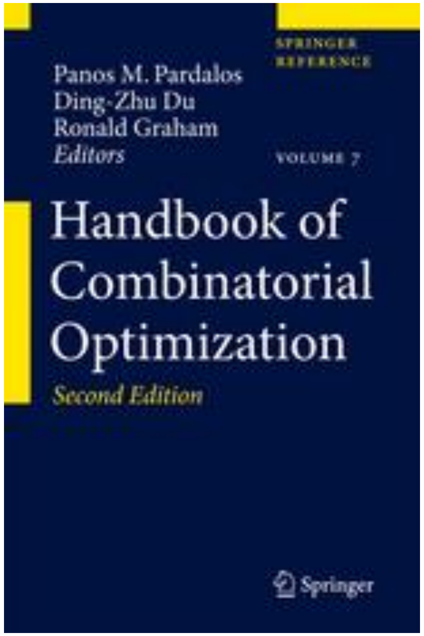




- **Network Analysis has become a major research topic over the last years.**

- **The broad range of applications that can be described and analyzed by means of a network is bringing together researches from numerous fields:**
    - Operations Research
    - Computer Science
    - Transportation
    - Biomedicine
    - Energy
    - Social Sciences
    - Computational Neuroscience
    - Others.



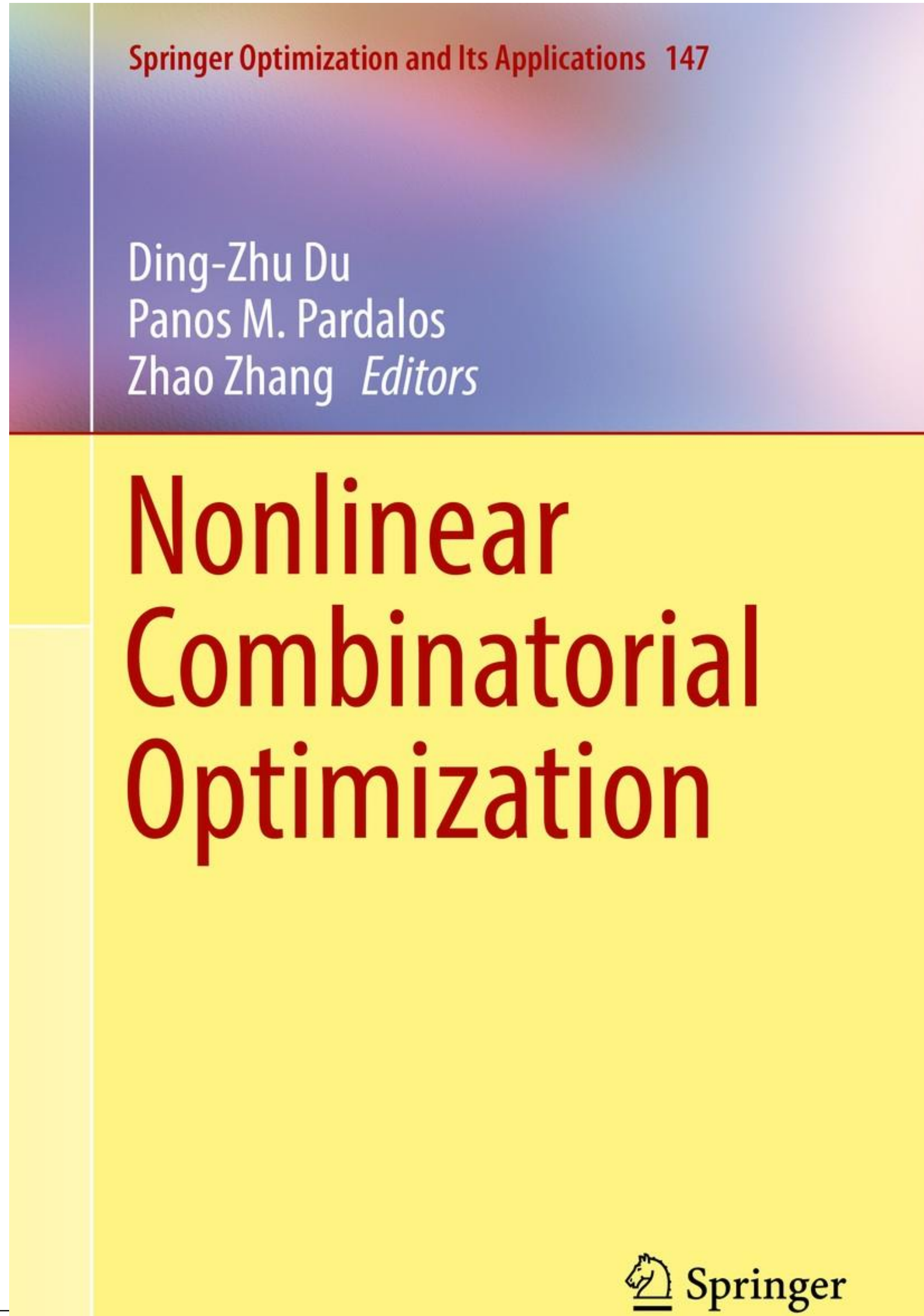


- **Handbook of Combinatorial Optimization (2nd Edition) , Springer Nature 2013, 4930 p. [In 7 volumes] -3rd edition in progress**
  Panos M. Pardalos, Ding-Zhu Du, and Ronald L. Graham (Eds.)
  http://www.springer.com/mathematics/book/978-1-4419-7996-4#

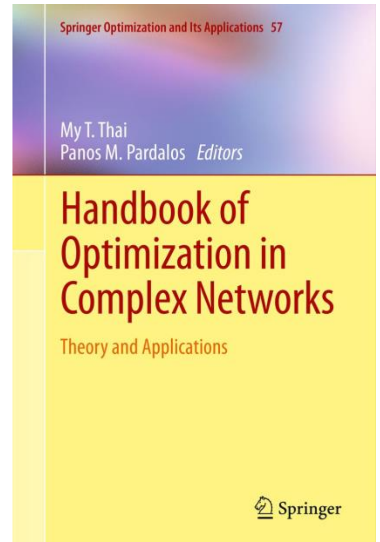
- **Every combinatorial optimization problem can be reduced to an equivalent shortest path problem**.

Ding-Zhu Du , Panos M. Pardalos , Xiaodong Hu ,Weili Wu,
Introduction to Combinatorial Optimization, Springer (2022)

- Max Flow problems in networks, Minimum cost problems in networks, and Shortest path problems in networks

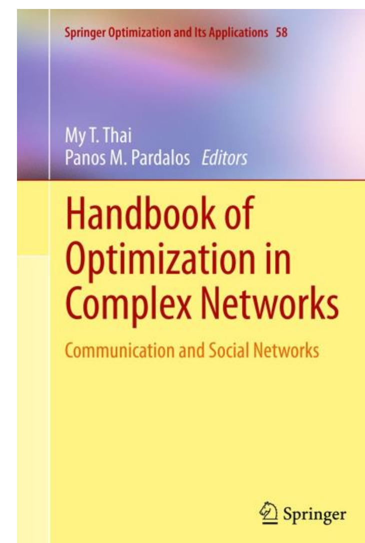

Springer Optimization and Its Applications 147
Ding-Zhu Du
Panos M. Pardalos
Zhao Zhang Editors
Nonlinear Combinatorial Optimization
Springer

- **Handbook of Optimization in Complex Networks**

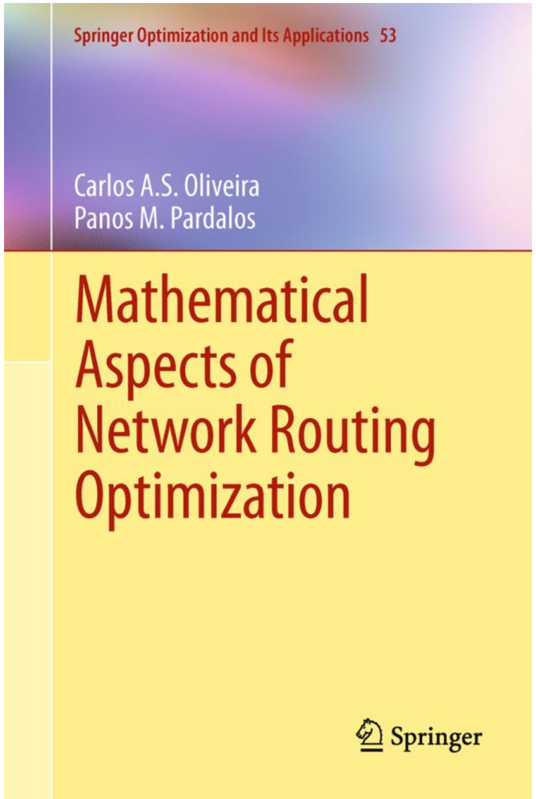


- **Volume 1: Theory and Applications**
  My T. Thai and Panos M. Pardalos (Eds.) Springer. Series: Springer Optimization and Its Applications , 2012. Vol. 57. ISBN 978-1-4614-0753-9

- **Volume 2: Communications and Social Networks** My T. Thai and Panos M. Pardalos (Eds.) Springer. Series: Springer Optimization and Its Applications , 2012. Vol. 58. ISBN 978-1-4614-0856-7

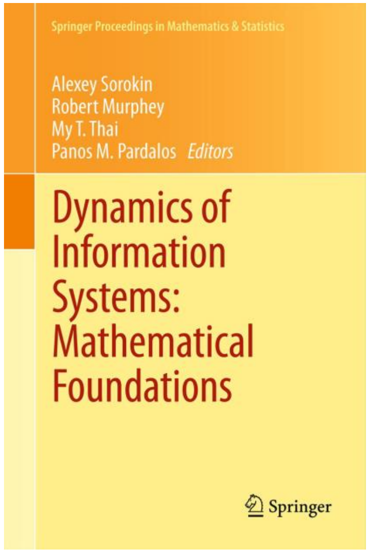


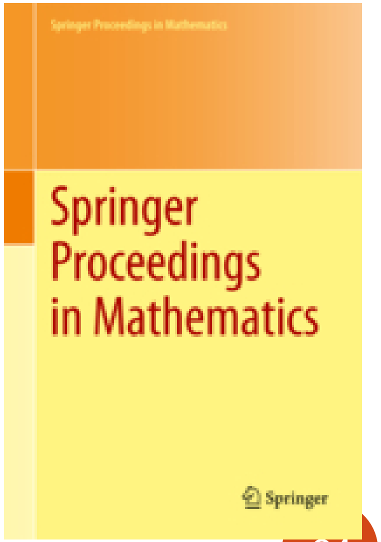
- **Mathematical Aspects of Network Routing Optimization**
  Carlos Oliveira and Panos M. Pardalos.
  Springer. Series: Springer Optimization and Its Applications , 2011. Vol. 53. ISBN 978-1-4614-0310-4



- **Dynamics of Information Systems: Mathematical Foundations**
  Alexey Sorokin, Robert Murphey, My T. Thai, and Panos M. Pardalos (Eds.) Springer. Springer Proceedings in Mathematics & Statistics, 2012. Vol. 20. ISBN 978-1-4614-3905-9

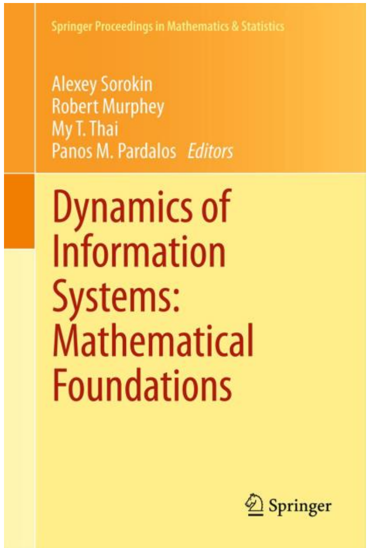
- **Dynamics of Information Systems: Algorithmic Approaches**
  Alexey Sorokin, Robert Murphey, My T. Thai, and Panos M. Pardalos (Eds.) Springer. Springer Proceedings in Mathematics & Statistics. Forthcoming

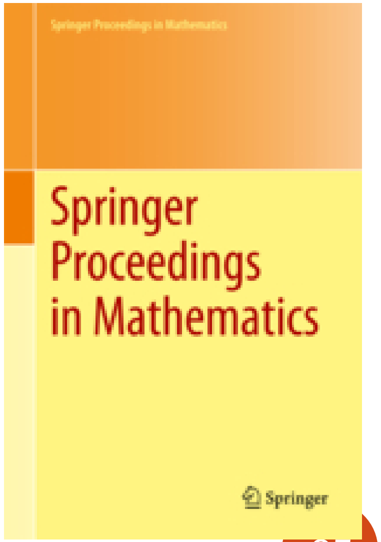




- **Dynamics of Information Systems: Mathematical Foundations**
  Alexey Sorokin, Robert Murphey, My T. Thai, and Panos M. Pardalos (Eds.) Springer. Springer Proceedings in Mathematics & Statistics, 2012. Vol. 20. ISBN 978-1-4614-3905-9

- **Dynamics of Information Systems: Algorithmic Approaches**
  Alexey Sorokin, Robert Murphey, My T. Thai, and Panos M. Pardalos (Eds.) Springer. Springer Proceedings in Mathematics & Statistics. Forthcoming





Recent Graph Theory &Networks

- A major step in graph theory was the development of the concept called "random graphs" by Erdos and Renyi**.

** Erdős, P., Rényi, A., 1959. **On random graphs**. Publ. Math. 6, 290–297.

Recent Graph Theory & Networks

- Another important step in graph theory was the development of an idea called "six degrees of freedom" by Frigyes Karinthy*.
- A study conducted by Watts and Strogatz** presented the transition of a network from one extreme case "regular network" to another extreme case "random network".
- Their most interesting observation was the development of intermediate type of networks called "Small World Networks".

*Milgram, S., 1967. The small-world problem. Psychol. Today 1, 61–67.
**Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of "small-world" networks. Nature 393, 440–442.

Recent Graph Theory & Networks

- The results of the study on three different types of network structures (regular, random and small world) tremendously increased the usage of network theory in solving real world problems.
- One of the prominent usage of the network theory is in analyzing complex systems.

Network Models in Nature

- Networks provide information on which system components interact with each other, putting aside the mechanisms of these interactions.
- The dynamics of the system is tightly connected to the structure of the underlying network.
- Studying the topological properties of system's networks can provide an insight about the system's high level structure and its function.

Networks can be very large

Nature April 2025

- https://www.nature.com/articles/d41586-025-01088-x

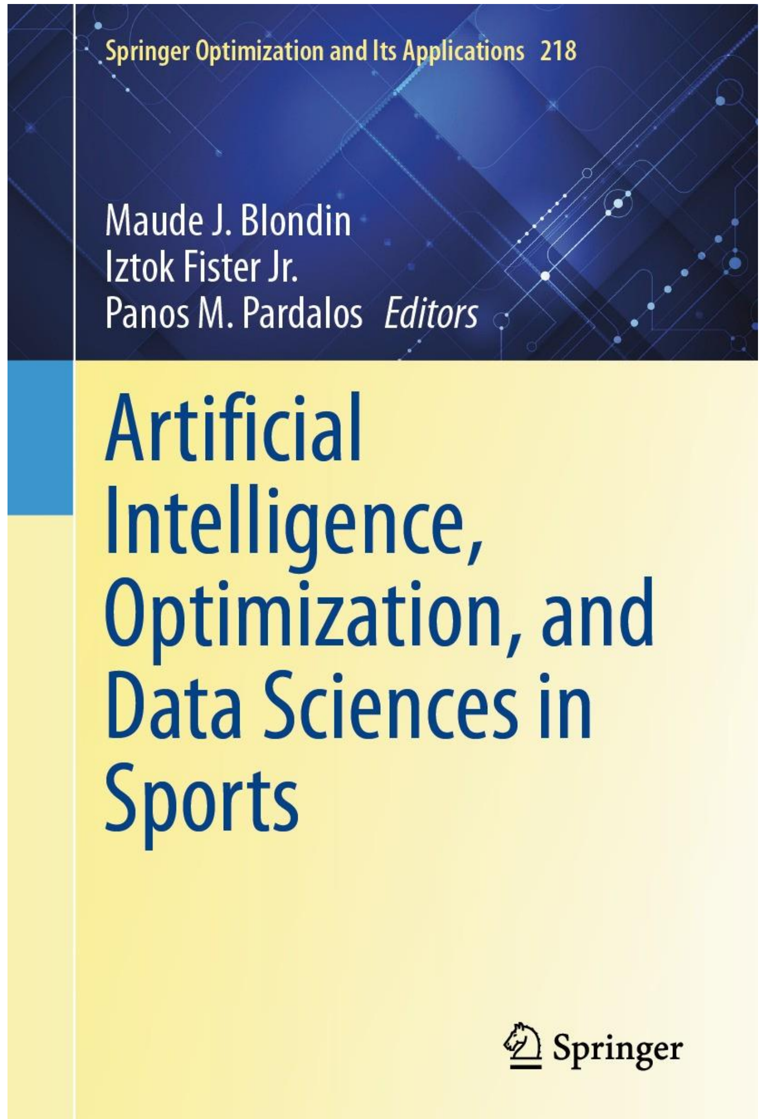
Biggest brain map ever details huge number of neurons and their activity

3D reconstruction is the first to overlay neuronal activity on a large-scale map of brain cells.

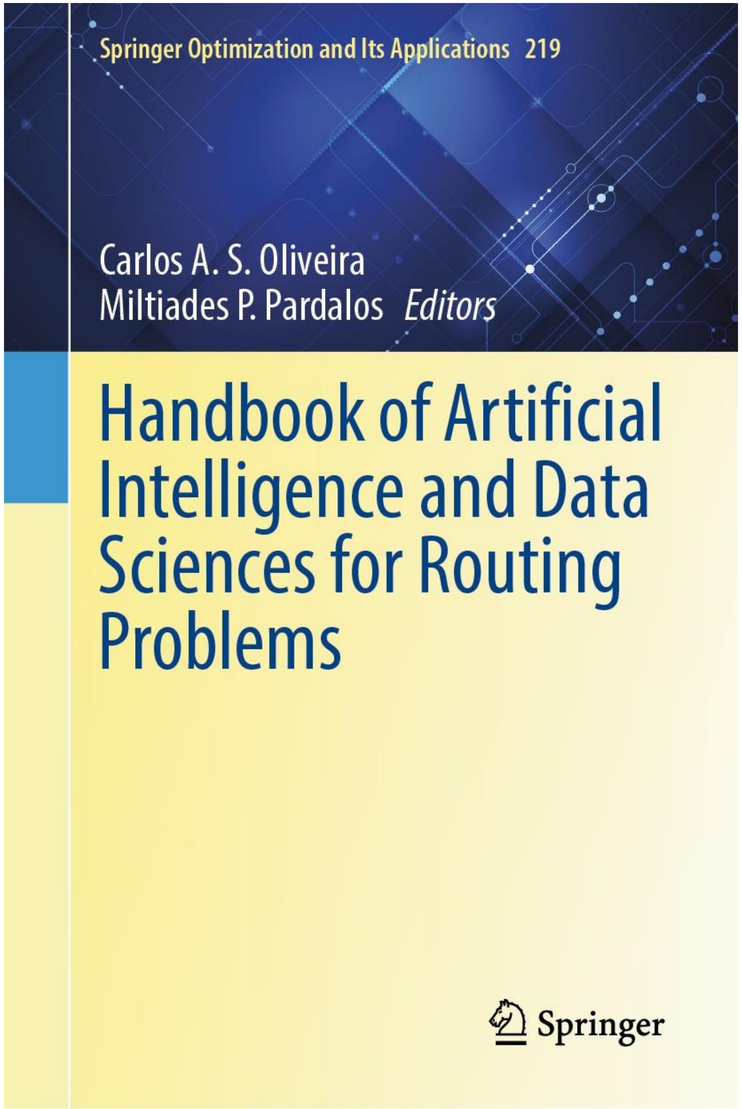
Our Work on Networks Includes:

- Telecommunication networks (call graph)
- Financial networks (market graph)
- Brain networks (Epileptic and Parkinson brain)
- Biological Networks
- Social networks (sports networks)
- Transportation networks (evacuation networks)
- Energy Networks (smart grid)
- Smart Cities
- Massive Networks

https://link.springer.com/book/10.1007/978-3-031-76047-1



https://link.springer.com/book/10.1007/978-3-031-78262-6



Networks of Networks

- IoT (**Internet of Things**)

A worldwide network of interconnected objects that are uniquely addressable via standard communication protocols.

- Cooperative networks
- Multicast networks
- Interdepended networks
- Networks of networks
- Sustainable interdepended networks

Critical Elements and Robustness

- We studied **critical elements** in regards to connectivity

- A new **measure of robustness** has been introduced

- Work:
  - "On New Approaches of Assessing Network Vulnerability: Hardness and Approximation" (T. N. Dinh, Y. Xuan, M. T. Thai, P.M. Pardalos, and T. Znati), **IEEE/ACM Transactions on Networking (ToN)**, Vol. 20, No. 2 (2012), pp. 609-619.

  - "Detecting Critical Vertex Structures on Graphs: A Mathematical Programming Approach" (Walteros, J. L., Veremyev, A., Pardalos, P. M., and E.L. Pasiliao), **Networks**, Vol. 73, No. 1 (2019), pp. 48-88.

  - "Quantification of networks structural dissimilarities (Tiago A. Schieber, Laura Carpi, Albert D´ıaz-Guilera, Panos M. Pardalos, Cristina Masoller and Mart´ın G. Ravetti), **Nature Communications** 8, online, Article number: 13928 (2017). https://www.nature.com/articles/ncomms13928

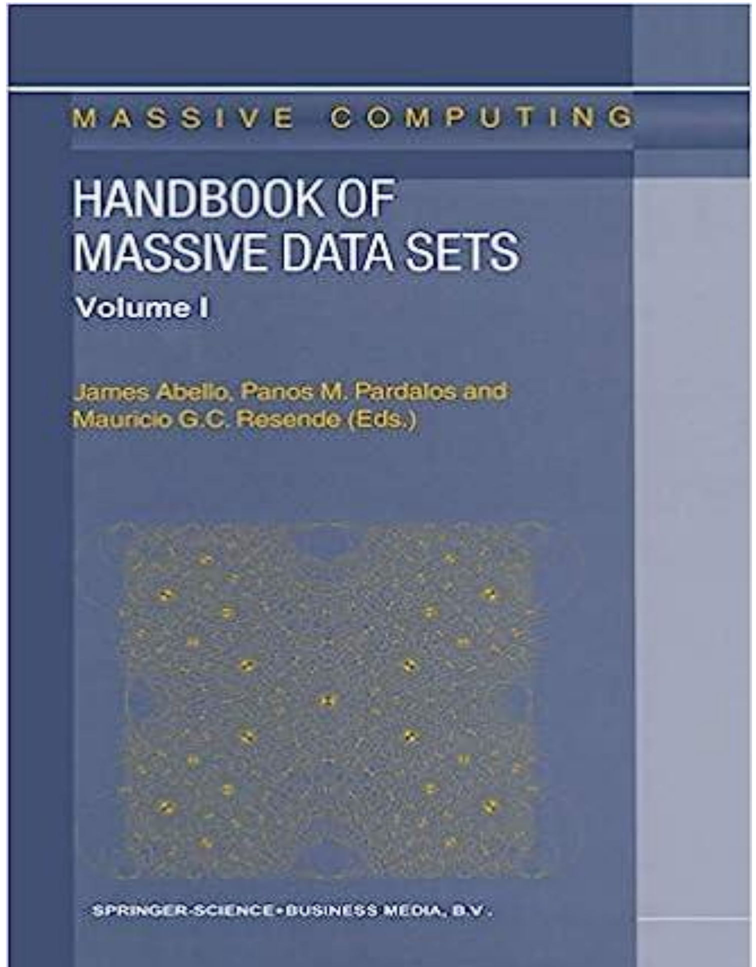
New Challenges

- Large, Massive and Complex (social networks, telecommunication networks, smart grid, biological networks, financial networks, etc)
- **Handbook of Massive Data Sets** (Springer 2013)

(James Abello, Panos M. Pardalos, Mauricio G.C. Resende)

(Semi) External memory algorithms

Network(s) Representation


MASSIVE COMPUTING
HANDBOOK OF
MASSIVE DATA SETS
Volume I
James Abello, Panos M. Pardalos and
Mauricio G.C. Resende (Eds.)
SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

New Challenges

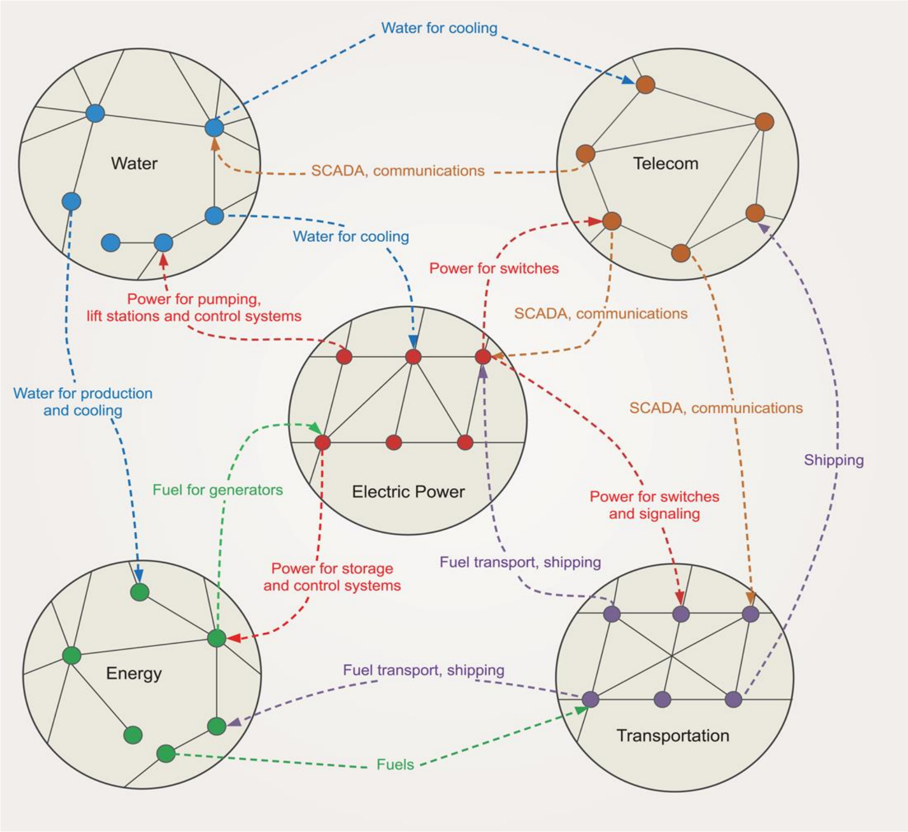
- Schieber, T., Carpi, L., Díaz-Guilera, A., Pardalos, P.M. *et al.* **Quantification of network structural dissimilarities**. *Nat Commun* **8**, 13928 (2017). https://doi.org/10.1038/ncomms13928

- Carpi, L.C., Schieber, T.A., Pardalos, P.M. *et al.* **Assessing diversity in multiplex networks**. *Sci Rep* **9**, 4511 (2019). https://doi.org/10.1038/s41598-019-38869-0

Interdependent and multiplex networks

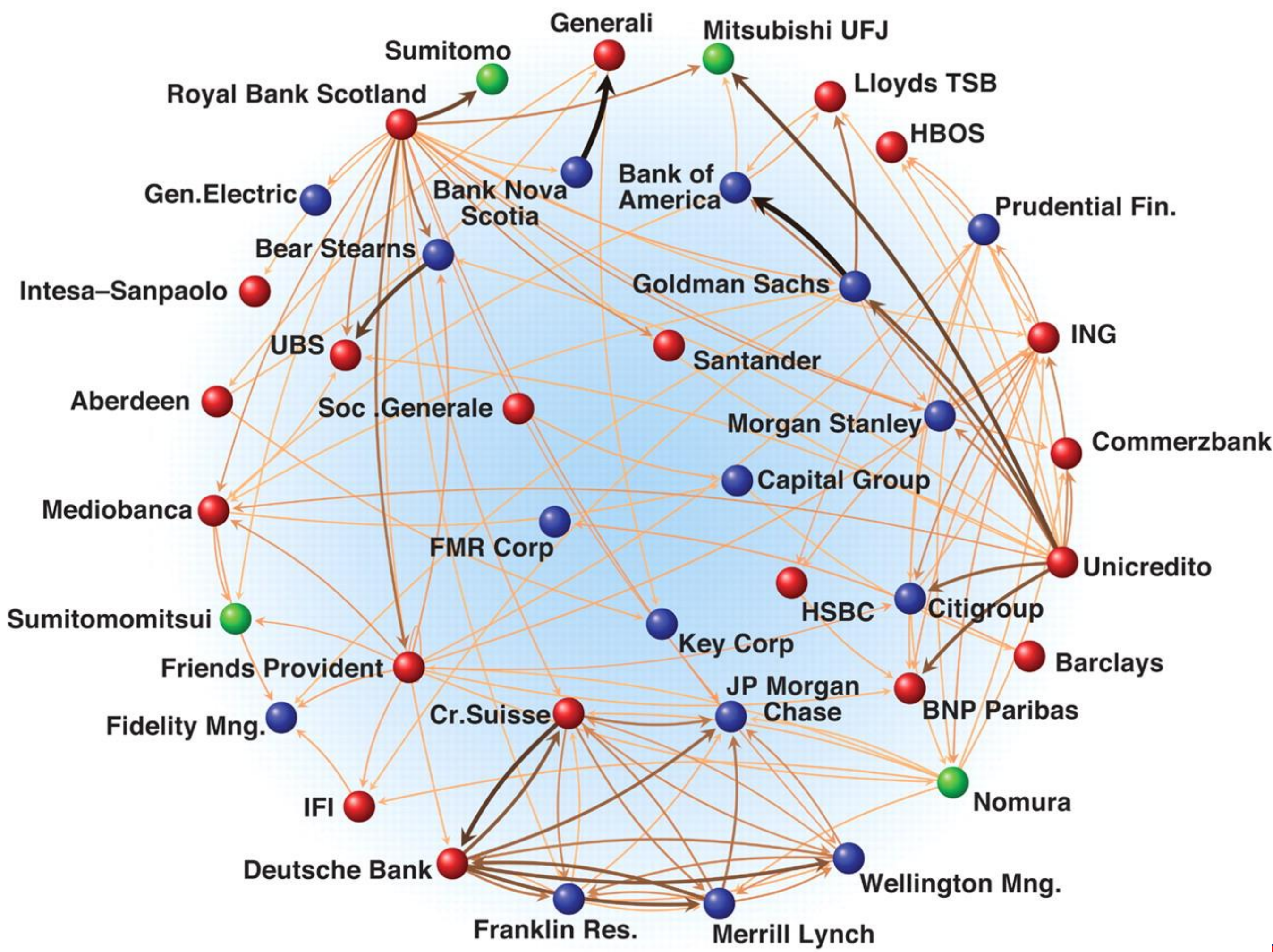
- Many complex systems (in nature or man made) are represented not by single networks but by sets of interdependent networks.
- In the simplest case, interdependent networks are equivalent to the so-called multiplex networks containing vertices of one type but several kinds of edges.
- Connectivity properties of these networks and their robustness are different from ordinary networks.

Figure 1: Illustration of the interdependent relationship among different infrastructures [69].



National Science Review, Volume 1, Issue 3, September 2014, Pages 346–356, https://doi.org/10.1093/nsr/nwu020

Figure 2: A sample of the international financial network, where the nodes represent major financial institutions and the links are both directed and weighted and represent the strongest existing relations among them.

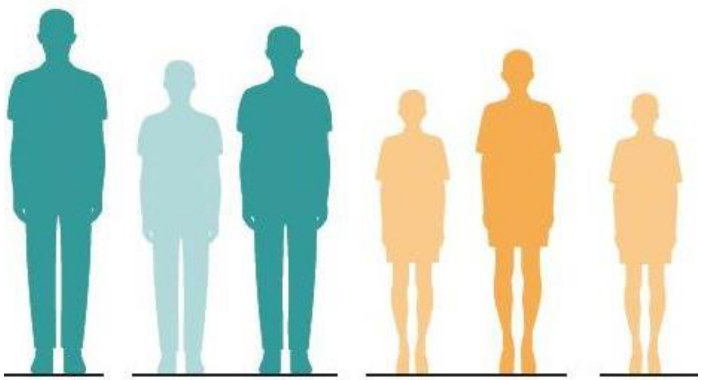


Frank Schweitzer et al. Science 2009;325:422-425

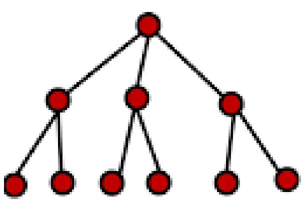
Published by AAAS

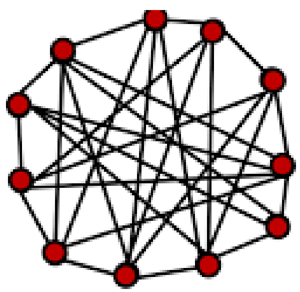
What is diversity?

Take into account three characteristics of a population:



- Diversity of attributes (e.g. atoms with different masses, people with different heights, etc.),
- Diversity of types (e.g. atoms or molecules, males or females, etc.),
- Diversity of configurations (configurations of atoms in molecules, patterns in a network , etc.).



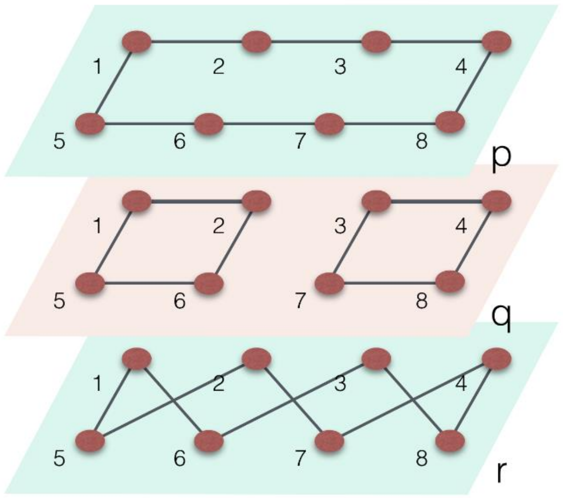


In multiplex networks, **what is diversity?**

A multiplex network with M layers, each one with the same set of N nodes, is represented by a set of adjacency matrices: $M(NxN)$

$$\mathcal{A} = \{A^{[1]}, A^{[2]}, \ldots, A^{[M]}\}$$

Then, diversity refers to the **variety** or **heterogeneity** of connectivity configurations or patterns, a set of elements (nodes layers), possesses.



To quantify diversity we first need to define a **distance** or a **dissimilarity** measure to compare:

- **Local** connectivity patterns of a node in the different layers;
- **Global** connectivity patterns of pairs of layers.

New distances

$$\mathcal{D}_i(\overline{p},\overline{q}) = \frac{\sqrt{\mathcal{J}(\mathcal{N}_i^{\overline{p}},\mathcal{N}_i^{\overline{q}})}+\sqrt{\mathcal{J}(\mathcal{T}_i^{\overline{p}},\mathcal{T}_i^{\overline{q}})}}{2\sqrt{\log(2)}}$$

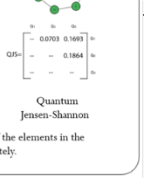
$\mathcal{D}_i$ quantifies the differences of the connectivity paths of a node in layers $\bar{p}$ and $\bar{q}$.

$\mathcal{N}_i^{\overline{p}}$ is the node distance distribution (NDD) of node i in layer $\overline{p}$: $\mathcal{N}_i^{\overline{p}}(d)$ is the fraction of nodes that are at distance d (shortest path) from node i in layer $\bar{p}$.

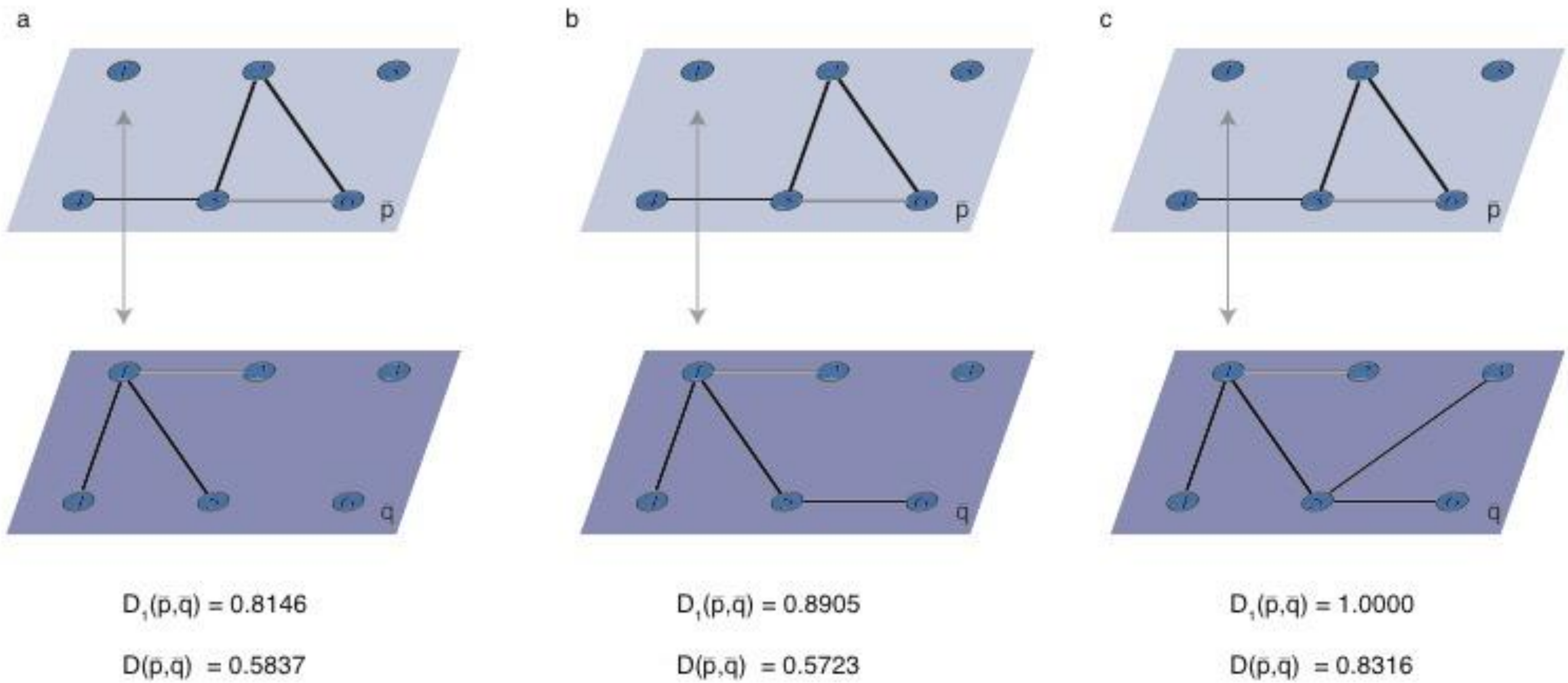
$\mathcal{T}^{\overline{p}}$ is the transition matrix of layer $\overline{p}$: $\mathcal{T}_i^{\overline{p}}(j)$ is the probability that node j in layer $\bar{p}$ to be reached in one step, by a random walker located at node i in $\bar{p}$.

$$\mathcal{D}(\overline{p},\overline{q}) = \langle \mathcal{D}_i(\overline{p},\overline{q}) \rangle_i$$

$\mathcal{D}(\overline{p},\overline{q})$ quantifies differences between layers $\bar{p}$ and $\bar{q}$.



Quantification of nodes and layers differences

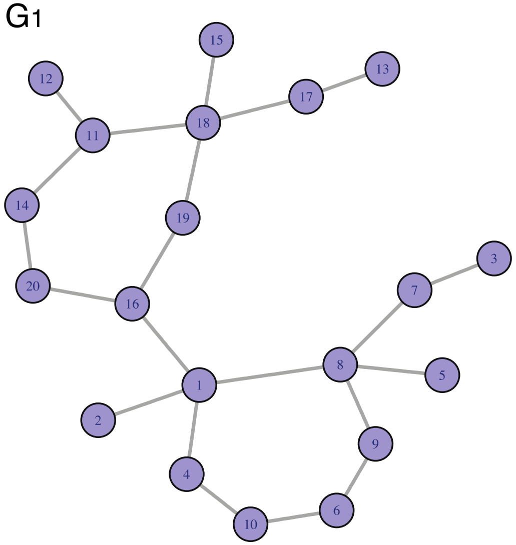


$\mathcal{D}_i = 0$ indicates that node i has identical connectivity paths in layers $\bar{p}$ and $\bar{q}$.

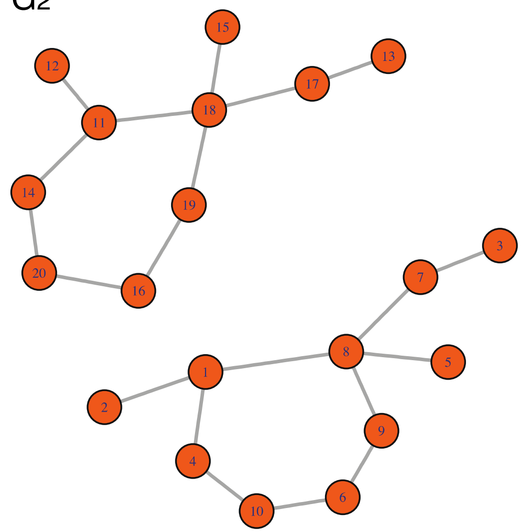
$\mathcal{D}_i = 1$ indicates that node i is not connected in one layer, while there are paths connecting it to all nodes in the other layer.

Comparing distances …

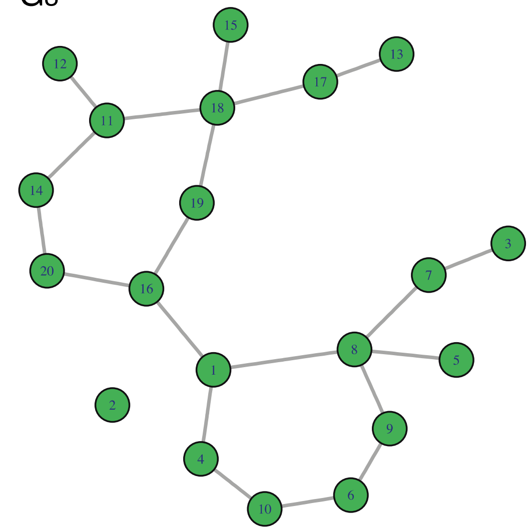
G_1



G_2



G_3



D=

	G_1	G_2	G_3
G_1	--	0.3597	0.1406
G_2	--	--	0.3447
G_3	--	--	--

Our D

GED=

	G_1	G_2	G_3
G_1	--	1	1
G_2	--	--	2
G_3	--	--	--

Graph edit distance

QJS=

	G_1	G_2	G_3
G_1	--	0.0703	0.1693
G_2	--	--	0.1864
G_3	--	--	--

Quantum Jensen-Shannon

Only our new distance captures the topological role of the elements in the network, quantifying their presence or absence adequately.

Diversity Measure

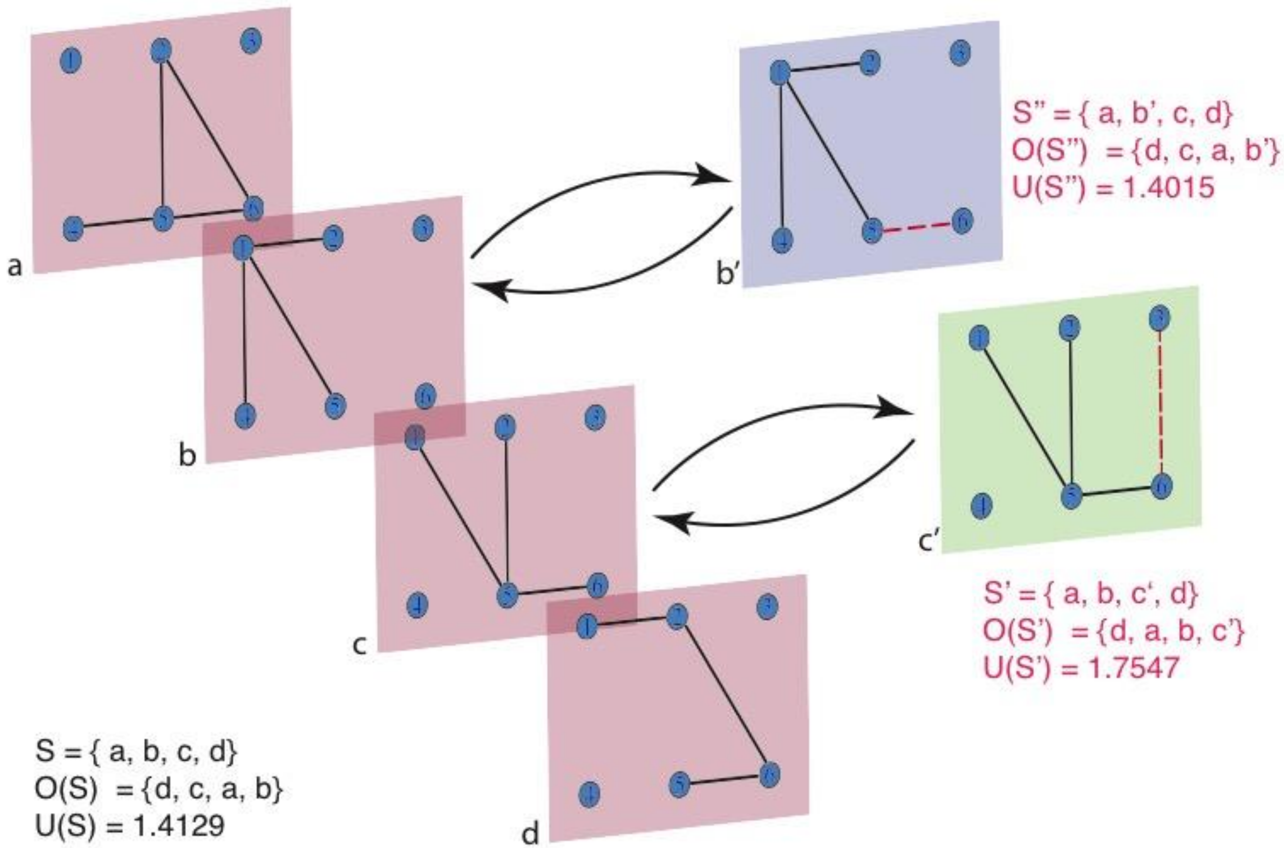
- The diversity of a system is defined by the distances between its elements: the larger the distances, the more diverse the system is.
- The distance between the element $\overline{g} \notin S$ and the set S, $\mathcal{D}(\overline{g}, S)$ is the smallest distance between $\overline{g}$ and any of the elements in S:

$$\mathcal{D}(\overline{g}, S) = \min_{\overline{s_i} \in S} \mathcal{D}(\overline{g}, \overline{s_i}).$$

- Diversity $U : \tilde{S} \to \mathbb{R}_+$ is defined recursively:

$$U(S) = max_{\overline{s_i} \in S}\{U(S \setminus \overline{s_i}) + \mathcal{D}(\overline{s_i}, S \setminus \overline{s_i})\}$$

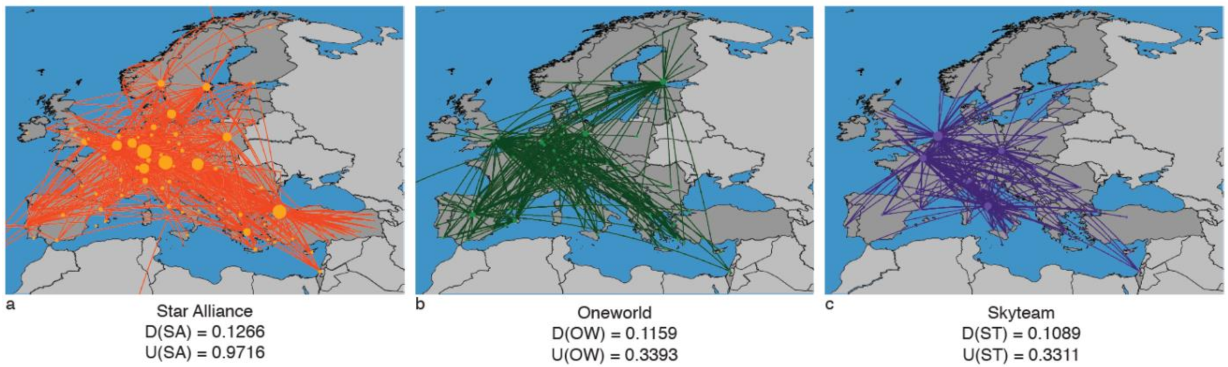
Variation of the global diversity value when adding different links.



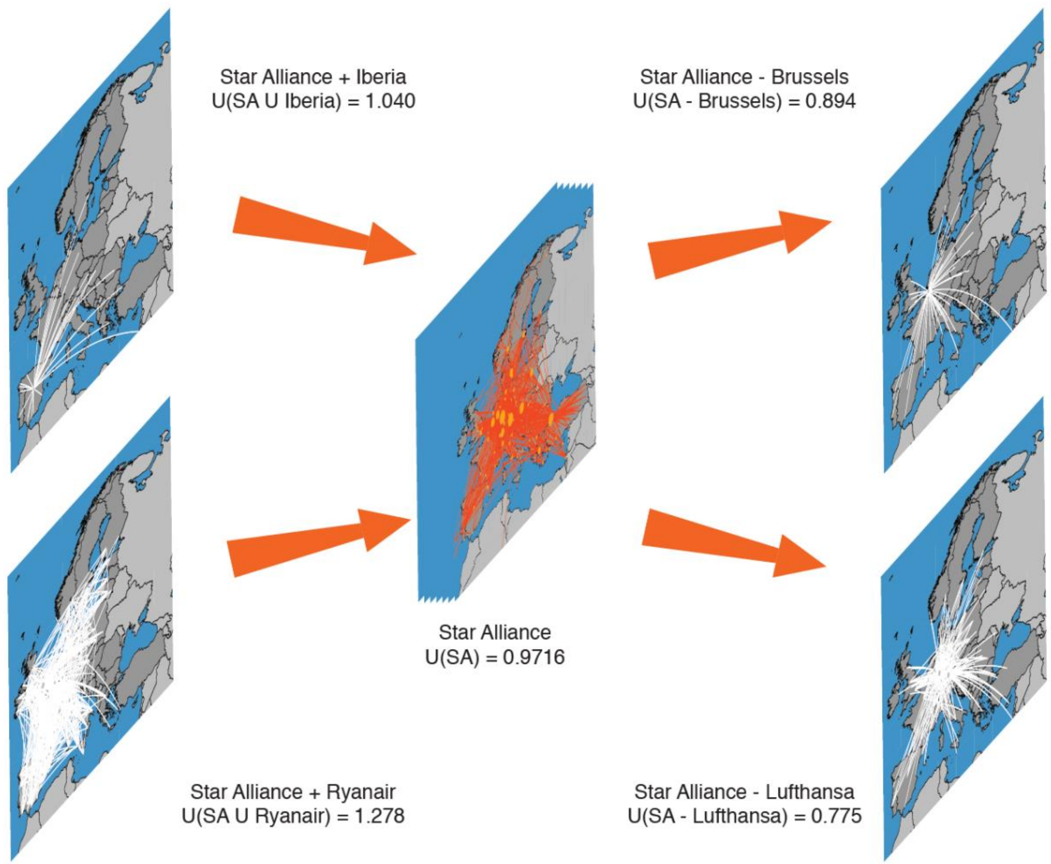
Diversity ordering O(S): order of layers depending of their contribution to global diversity(from less to more contribution).

Global diversity of the European Air Transportation Network (ATN).

a.



b.



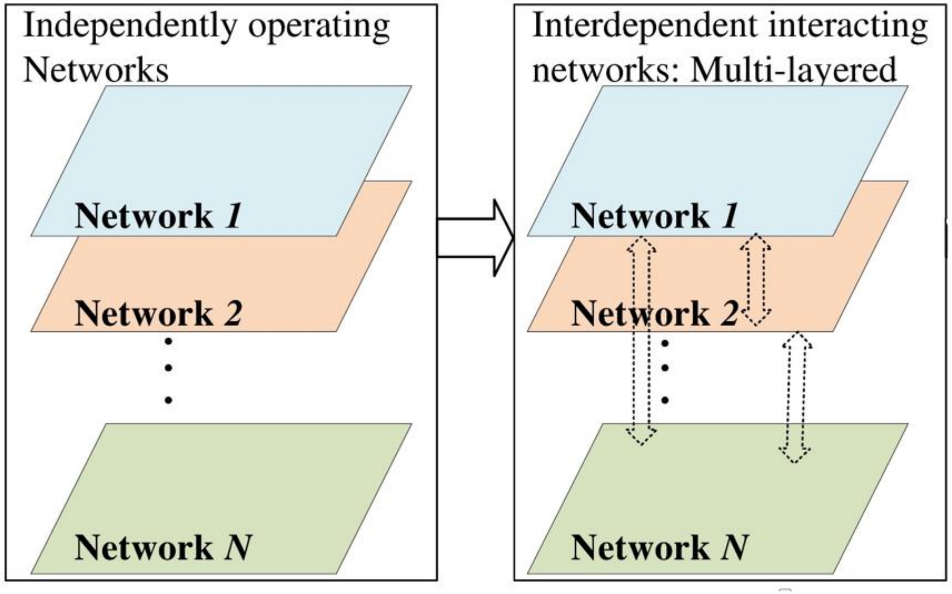
O(Star Alliance):
Brussels Airlines (BEL),
Swiss Air (SWR),
Polish Airlines (LOT),
Air Portugal (TAP),
Aegean Airlines (AEE),
Austrian Airlines (AUA),
Scandinavian Airlines (SAS),
Turkish Airlines (THY),
Lufthansa (DLH).

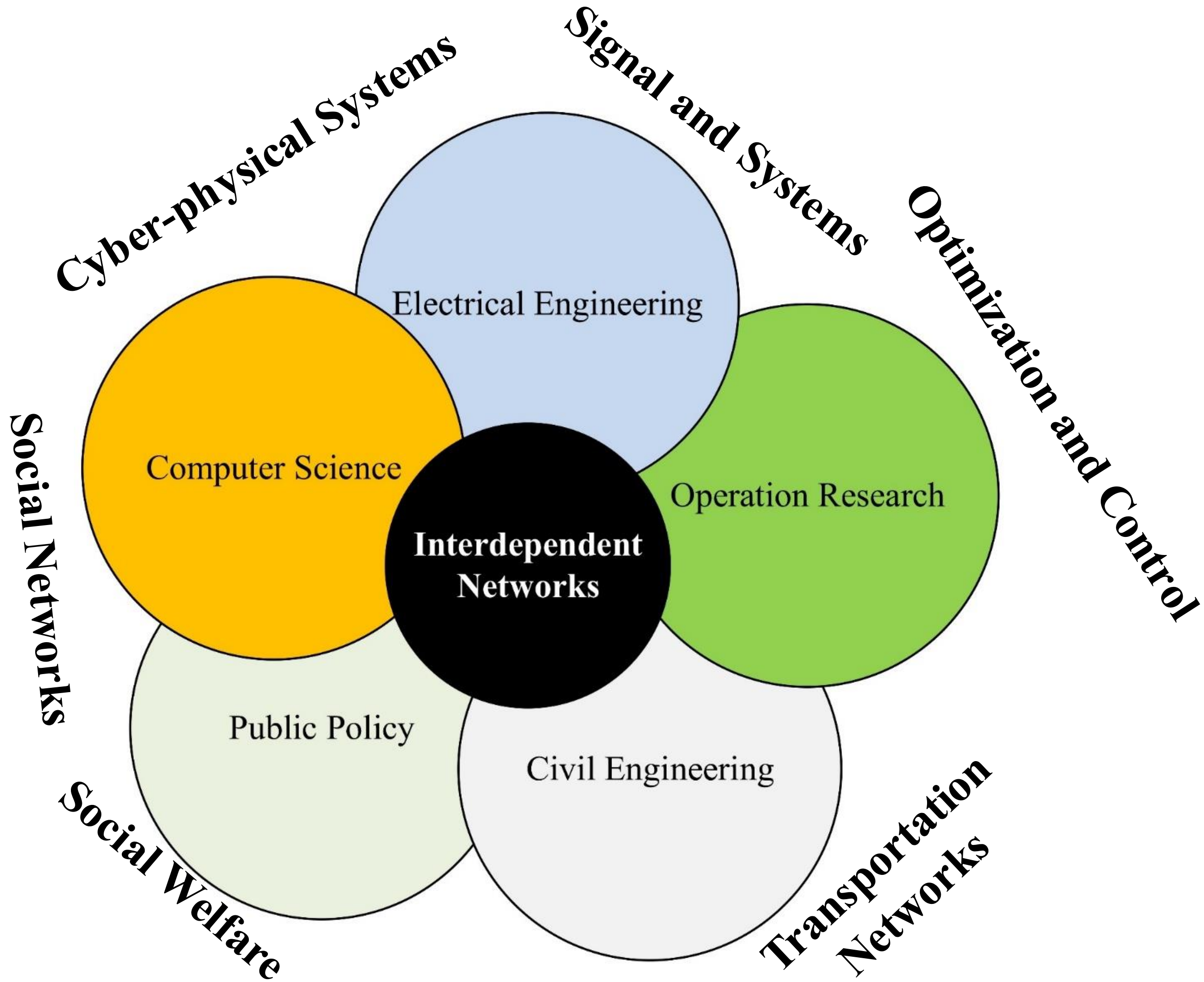
Results

- New distance between unweighted labelled graphs.
- Used to define the diversity of the connectivity paths of a node in the different layers, and the diversity of the connectivity paths of the whole set of layers.
- Analysis of air alliances reveals which airlines, when joining an alliance, optimally increase the diversity, bringing new routes while minimizing overlapping ones; and which ones, when leaving the alliance, less compromise the diversity of the routes offered by the alliance.
- Other applications: Carpi et al, **Nature Sci. Reports** 9, 4511 (2019) https://www.nature.com/articles/s41598-019-38869-0

Future Vision

From Independent Operation to Multi-layer Interdependent Networks




Cyber-physical Systems
Signal and Systems
Optimization and Control
Social Networks
Electrical Engineering
Computer Science
Interdependent Networks
Operation Research
Public Policy
Civil Engineering
Social Welfare
Transportation Networks

Sustainable Interdependent Networks Book Series

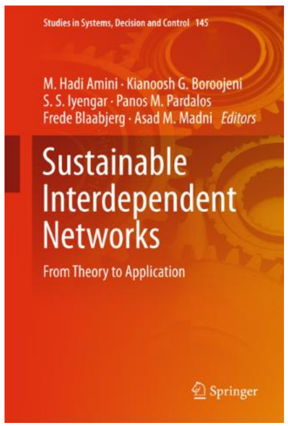
Features:

- ✓ **Interdisciplinary Nature**
- ✓ **Covering Both Theoretical and Application Aspects**
- ✓ **Introducing Novel Ideas for Researchers**
- ✓ **Engaging Outstanding Experts in Each Area**
- ✓ **Endorsed by World-renowned Scientists**
- ✓ **Bridges the Gap Between Various Disciplines**

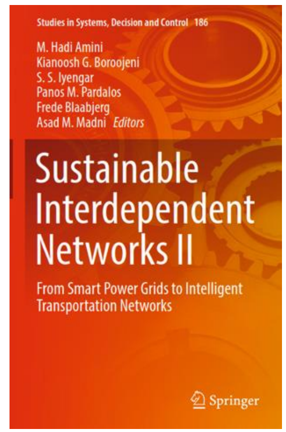
More information: http://interdependentnetworks.com/

Future Works: Studies in Systems, Decision, and Control

2016-2018



2018-2019

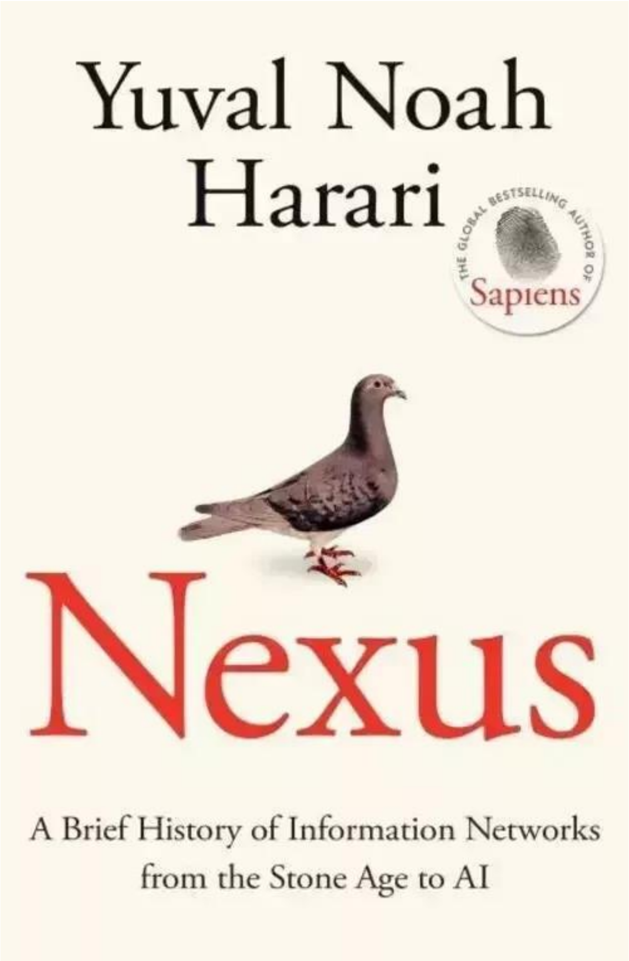


2019-

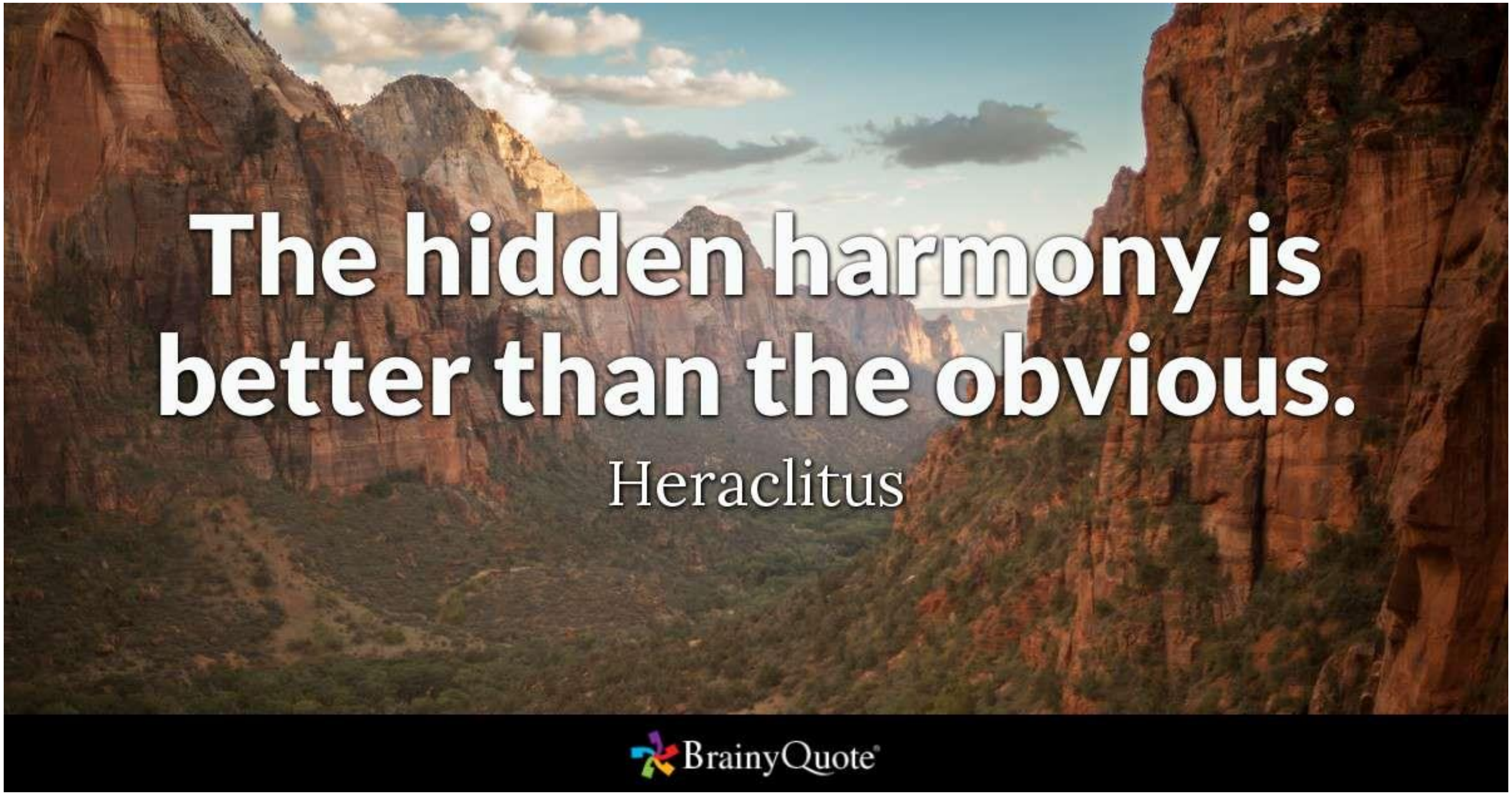
Sustainable Interdependent Networks III

- ✓ IoT-based Smart Cities
- ✓ Engaging Researchers and Industry Experts from Computer Science, Civil Engineering, Social Science, Distributed Systems, and Network Optimization

https://www.nytimes.com/2024/09/10/books/review/nexus-yuval-noah-harari.html



Questions?


The hidden harmony is better than the obvious.
Heraclitus
BrainyQuote