- “Donat el caràcter i la finalitat exclusivament docent i eminentment il·lustrativa de les explicacions a classe d'aquesta presentació, l’autor s’acull a l’article 32 de la Llei de propietat intel·lectual vigent respecte de l'ús parcial d'obres alienes com ara imatges, gràfics o altre material contingudes en les diferents diapositives”
- “Dado el carácter y la finalidad exclusivamente docente y eminentemente ilustrativa de las explicaciones en clase de esta presentación, el autor se acoge al artículo 32 de la Ley de Propiedad Intelectual vigente respecto al uso parcial de obras ajenas como imágenes, gráficos u otro material contenidos en las diferentes diapositivas”.

1

UNIVERSITAT DE BARCELONA

Complex Networks

Albert Díaz Guilera
http://diaz-guilera.net
@anduviera

complexity lab barcelona

COMPLEXITAT

2

Multiplex & Time Dependent Networks

Albert Díaz Guilera
http://diaz-guilera.net
@anduviera

complexity lab barcelona

COMPLEXITAT

3

Big data => more complex networks

- New paradigms:
  - Network of networks
  - Time-dependent networks
  - Interconnected networks
  - Multilayer networks. Multiplex networks

4

Interconnected networks

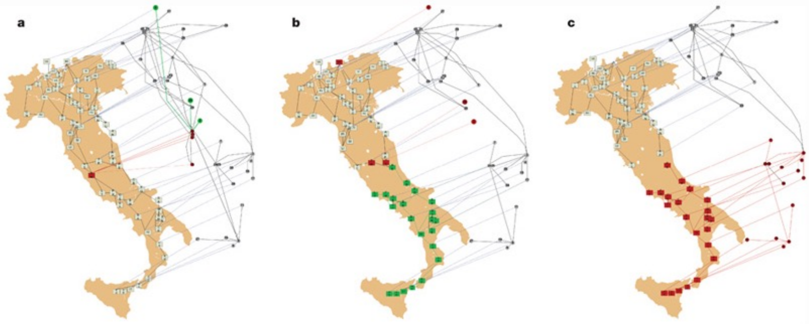
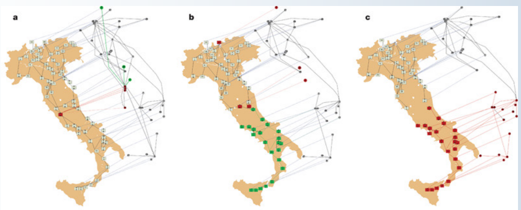


Illustration of an iterative process of a cascade of failures using real-world data from a power network (located on the map of Italy) and an Internet network (shifted above the map) that were implicated in an electrical blackout that occurred in Italy in September 2003[20]. The networks are drawn using the real geographical locations and every Internet server is connected to the geographically nearest power station. **a**, One power station is removed (red node on map) from the power network and as a result the Internet nodes depending on it are removed from the Internet network (red nodes above the map). The nodes that will be disconnected from the giant cluster (a cluster that spans the entire network) at the next step are marked in green. **b**, Additional nodes that were disconnected from the Internet communication network giant component are removed (red nodes above map). As a result the power stations depending on them are removed from the power network (red nodes on map). Again, the nodes that will be disconnected from the giant cluster at the next step are marked in green. **c**, Additional nodes that were disconnected from the giant component of the power network are removed (red nodes on map) as well as the nodes in the Internet network that depend on them (red nodes above map).

5

Big Data <-> Big Networks *Albert Díaz Guilera*

Modelling a blackout in Italy.



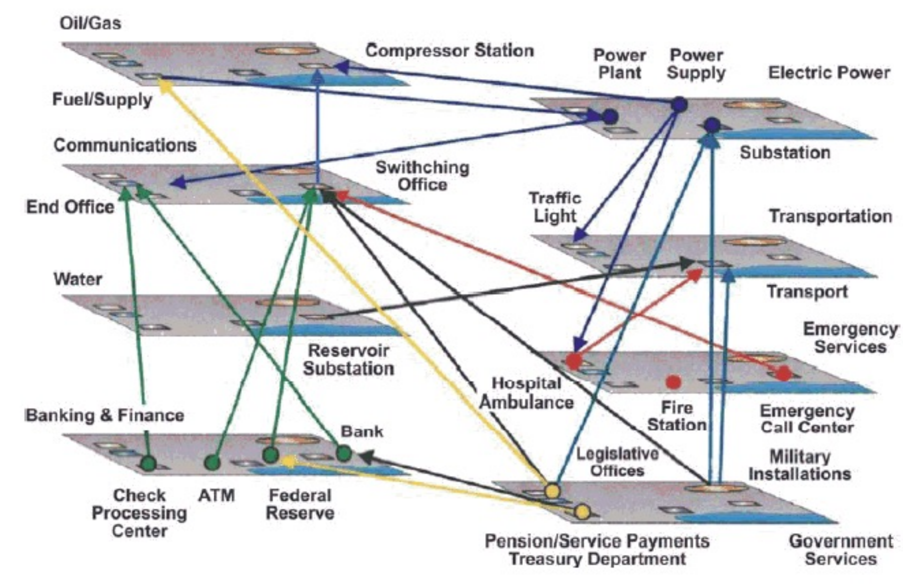
SV Buldyrev *et al.* *Nature* **464**, 1025-1028 (2010) doi:10.1038/nature08932

nature

6

Big Data <-> Big Networks Albert Díaz Guilera

Infrastructures



7

Big Data <-> Big Networks Albert Díaz Guilera

Cascades and failures

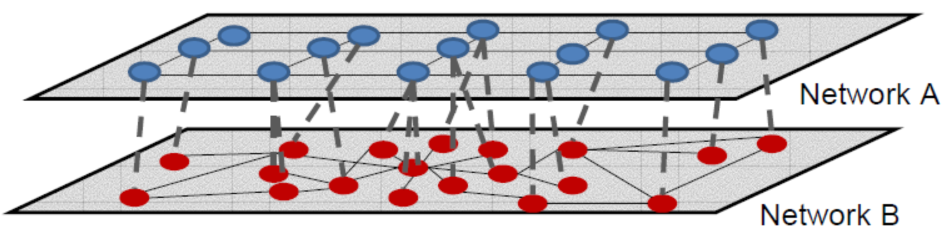


FIG. 1: (Color online) Demonstration of two interdependent networks. Nodes in network B (communications network) are dependent on nodes in network A (power grid) for power; nodes in network A are dependent on network b for control information. General case is represented in which there is not a one-one correspondence between nodes in networks.

8

Networks of networks

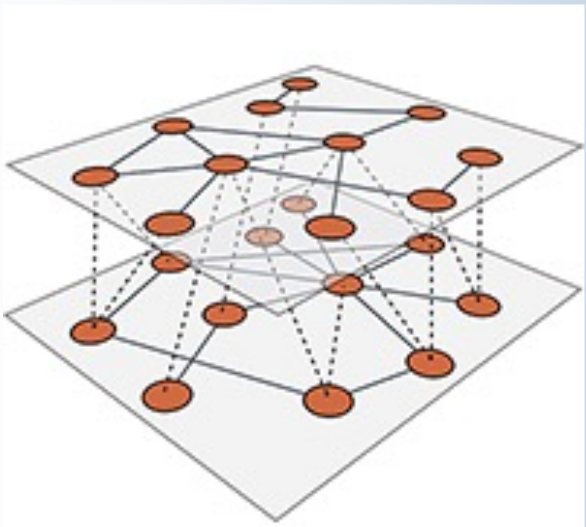
airline transportation network

commuting network

Balcan et al., PNAS 196, 21484 (2009)

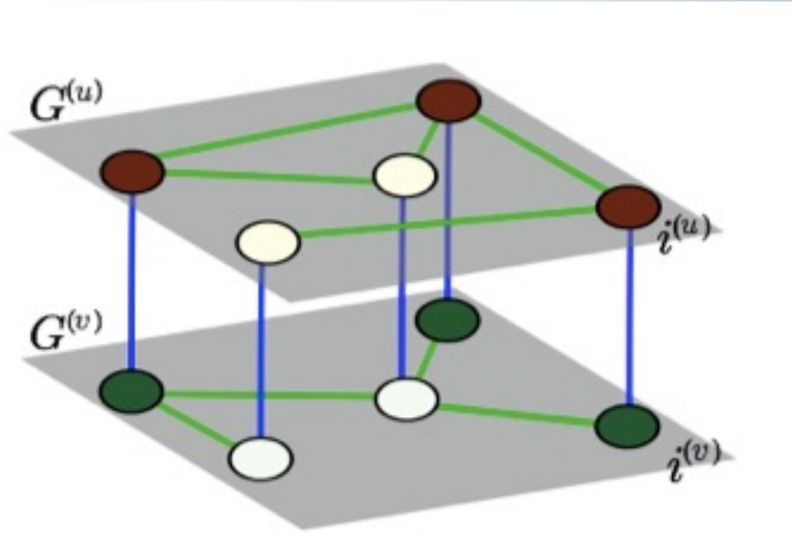
9

Multilayer



10

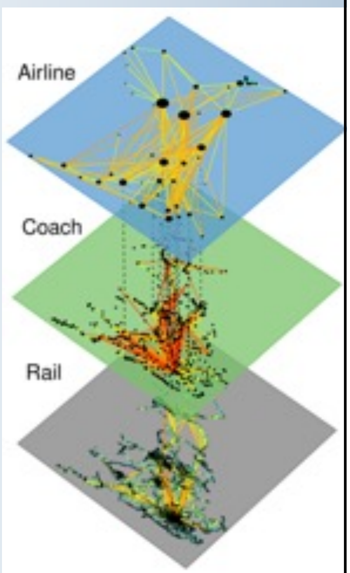
Multiplex



11

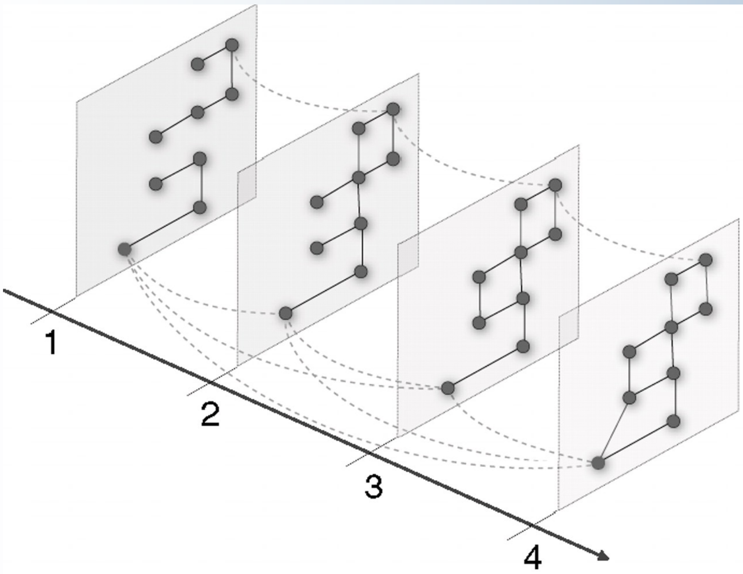
Big Data <-> Big Networks *Albert Díaz Guilera*

Transportation



15

Time dependent networks



16

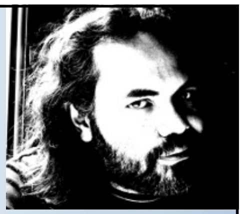
Big Data <-> Big Networks
Albert Díaz Guilera

Data sets

19

Big Data <-> Big Networks

Enzo Nicosia's

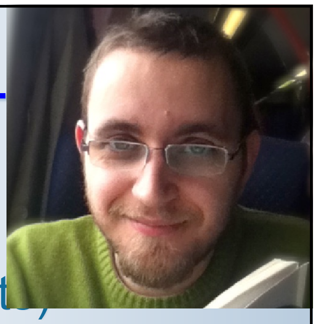


- C. Elegans (synapses and gap junctions)
- BioGrid (Physical and Genetic interactions among all proteins)
- OpenFlights continental airport networks (each continent, layers represent airlines)
- APS scientific collaboration network (layers are PACS 0-9)

20

Big Data <-> Big Networks

Manlio de Domenico



- Twitter (different events)
- Genetic interactions for different organisms
- FAO Multiplex Trade Network (layers are for specific products)

22

Big Data <-> Big Networks
Albert Díaz Guilera

Resources

SOFTWARE FOR MULTIPLEX METRICS AND MODELS (back to top)

MAMMULT (**M**etrics **A**nd **M**odels for **MULT**iplex networks) is a collection of tools for the analysis and modelling of multiplex networks.

MAMMULT can be downloaded from GitHub: **https://github.com/KatolaZ/mammult**

MultiRed is a Python module for the structural reduction of multiplex networks, based on the paper:

M. De Domenico*, V. Nicosia*, A. Arenas, V. Latora, "Structural Reducibility of Multilayer Networks", *Nat. Commun.* 6, 6864 (2015).

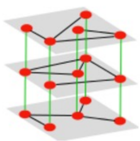
MultiRed can be downloaded from GitHub: **https://github.com/KatolaZ/multired**

23

Big Data <-> Big Networks
Albert Díaz Guilera

MultinetX

https://github.com/nkoub/multinetx



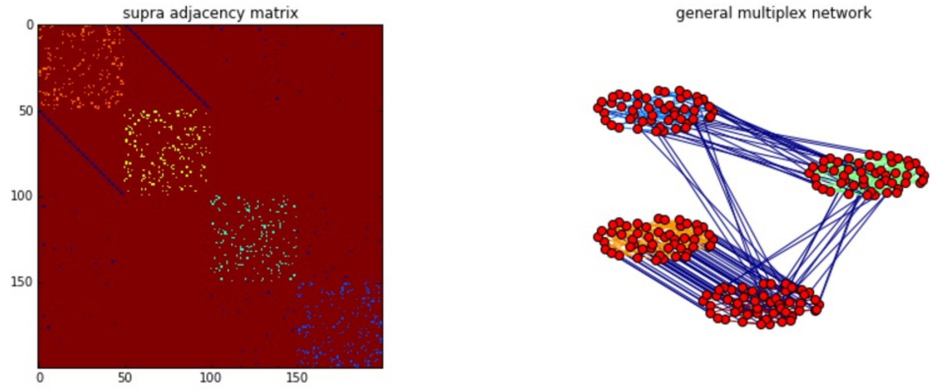
multiNetX v1.0

multiNetX is a python package for the manipulation and visualization of multilayer networks. The core of this package is a MultilayerGraph, a class that inherits all properties from networkx.Graph().

This allows for:

- Creating networks with weighted or unweighted links (only undirected networks are supported in this version)
- Analysing the spectral properties of adjacency or Laplacian matrices
- Visualizing dynamical processes by coloring the nodes and links accordingly

24



25

Multiplex network

$$M_{ij}^{\alpha\beta} = \begin{cases} \alpha = \beta & M_{ij}^{\alpha} \text{ Adjacency matrix of layer} \alpha \\ i = j & M_{ii}^{\alpha\beta} = 0,1 \text{ Possible links between layers} \\ i \neq j \;\&\; \alpha \neq \beta & M_{ij}^{\alpha\beta} = 0 \end{cases}$$

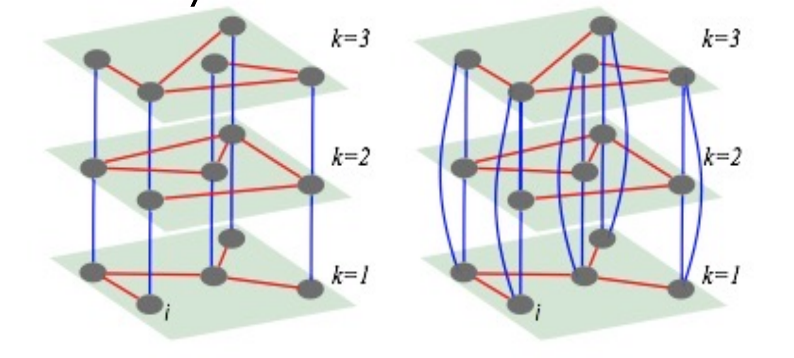
$$\frac{dn_i^\alpha}{dt} = F_i^\alpha \left[n_i^\alpha(t), M_{ij}^\alpha n_j^\alpha(t), M_{ii}^{\alpha\beta} n_i^\beta(t) \right]$$

26

One-to-one

$$M_{ii}^{\alpha\beta} = M^{\alpha\beta} = 0,1 \text{ the same } \forall\, i$$

Matrix of layers



$$\frac{dn_i^\alpha}{dt} = F_i^\alpha\left[n_i^\alpha(t), M_{ij}^\alpha n_j^\alpha(t), M^{\alpha\beta} n_i^\beta(t)\right]$$

27

Big Data <-> Big Networks
Albert Díaz Guilera

CHARACTERIZATION

- Degree
- Clustering
- Centrality
- Communities
- Layer comparison

28

Degree. Activity

$$\boldsymbol{k}_i = \left\{k_i^{[1]}, k_i^{[2]}, \ldots, k_i^{[M]}\right\}$$

We say that node i, with $i = 1, 2, \ldots, N$, is *active* at layer α if $k_i^{[\alpha]} > 0$. We can then associate to each node i a *node-activity vector*

$$\boldsymbol{b}_i = \left\{b_i^{[1]}, b_i^{[2]}, \ldots, b_i^{[M]}\right\} \quad (2)$$

where

$$b_i^{[\alpha]} = 1 - \delta_{0,k_i^{[\alpha]}}$$

i.e., $b_i^{[\alpha]} = 1$ if node i has at least one edge at layer α, and is 0 otherwise. We call *node-activity* B_i of node i the number of layers on which node i is active:

$$B_i = \sum_\alpha b_i^{[\alpha]} \quad (3)$$

29

Degree. Participation ratio

$$P_i = \frac{M}{M-1}\left[1 - \sum_{\alpha=1}^{M}\left(\frac{k_i^{[\alpha]}}{o_i}\right)^2\right]$$

30

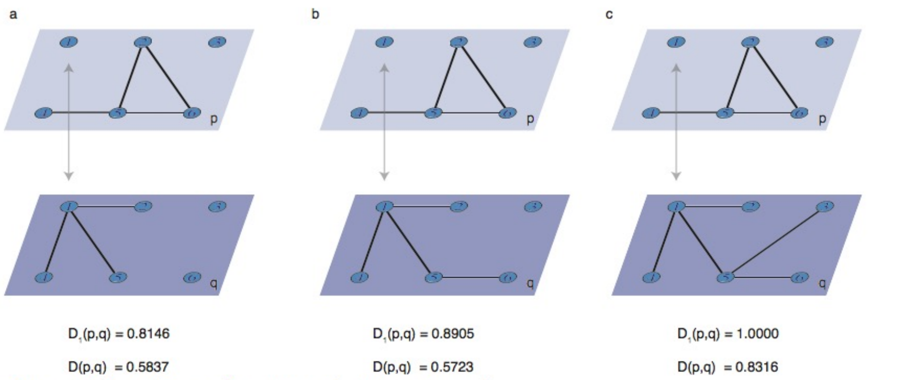
Layers comparison

- Hamming distance

$$H^{[\alpha,\beta]} = \frac{\sum_i b_i^{[\alpha]}(1-b_i^{[\beta]}) + b_i^{[\beta]}(1-b_i^{[\alpha]})}{\min(\mathrm{N}^{[\alpha]}+\mathrm{N}^{[\beta]},\mathrm{N})}.$$

33

Dissimilarity

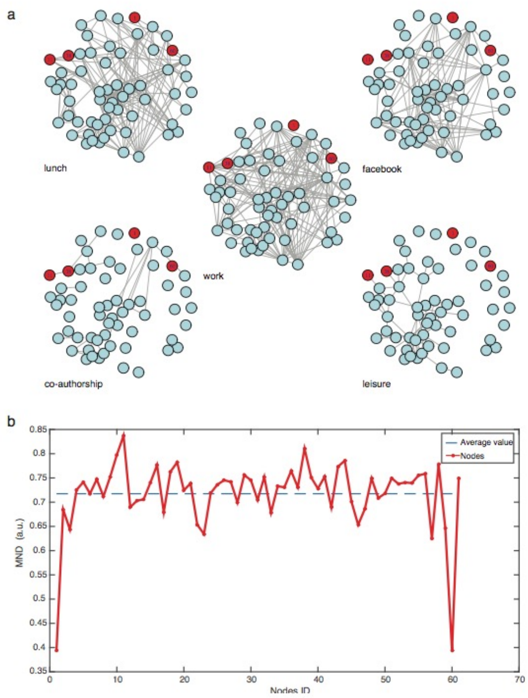


Quantification of network structural dissimilarities

Tiago A. Schieber, Laura Carpi, Albert Díaz-Guilera, Panos M. Pardalos, Cristina Masoller & Martín G. Ravetti

Nature Communications **8**, Article number: 13928 (2017)
doi:10.1038/ncomms13928
Received: 26 May 2016
Accepted: 15 November 2016
Published online: 09 January 2017

34



35

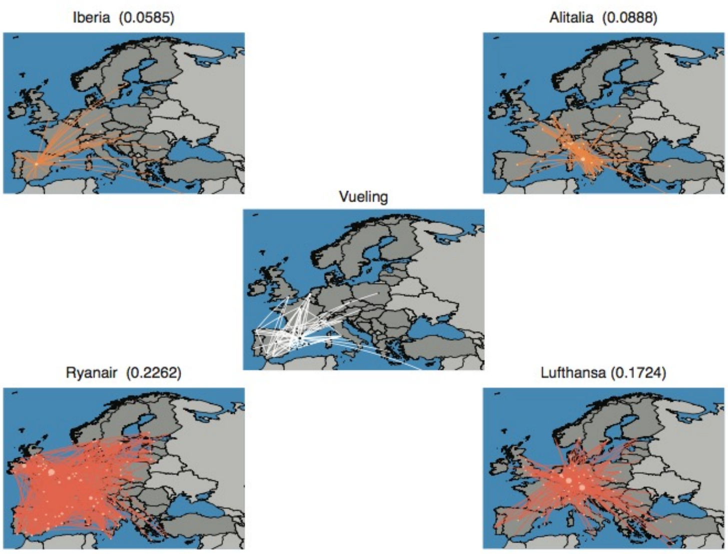
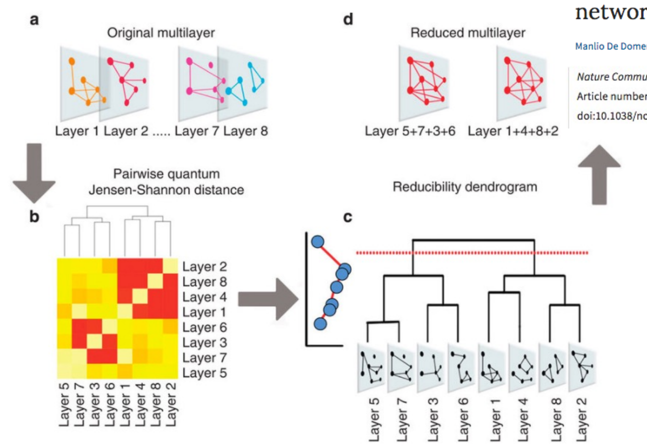


FIG. 6. **Dissimilarity values between Vueling and 4 other companies.** Dissimilarity values between Vueling (center), and the two closest (top) and more distant (bottom) networks, Iberia and Alitalia, and Ryanair and Lufthansa, respectively.

36

Reducibility

From: Structural reducibility of multilayer networks



Given a multilayer network (**a**), we compute the Jensen–Shannon distance $\mathcal{D}_{JS}$ between each pair of its layers (**b**), which is a proxy for layer redundancy. Such resulting distance matrix allows to perform a hierarchical clustering, whose output is a hierarchical diagram (a dendrogram) whose leaves represent the initial layers and internal nodes denote layer merging (**c**). At each step, the two clustered layers (or group of layers) corresponding to the smallest value of $\mathcal{D}_{JS}$ are aggregated and the quality of the new layer configuration in terms of distinguishability from the aggregated graph is quantified by the global quality function $q(\cdot)$, shown by the curve on the left-hand side of **c**. The best partition is the one for which $q(\cdot)$ is maximal (**d**).

Structural reducibility of multilayer networks

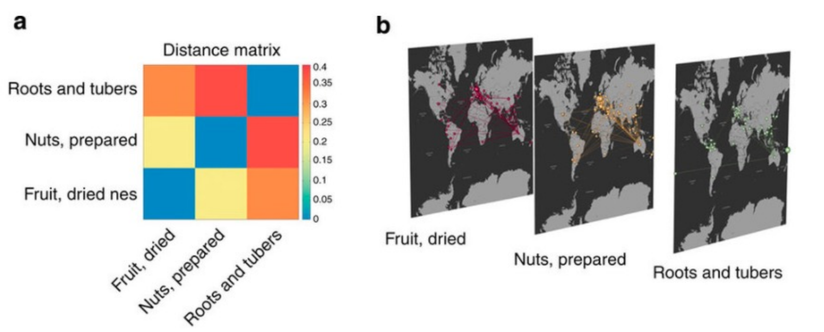
Manlio De Domenico, Vincenzo Nicosia, Alexandre Arenas & Vito Latora

Nature Communications **6**, Article number: 6864 (2015)
doi:10.1038/ncomms7864
Received: 03 July 2014
Accepted: 07 March 2015
Published online: 23 Apr

37

Figure 5: Structural reducibility of the FAO worldwide food import/export network.

From: Structural reducibility of multilayer networks



The distance matrix of three layers of the FAO worldwide food import/export data set, corresponding to three specific products (that is, 'roots and tubers', 'prepared nuts' and 'dried fruit'), is shown in **a**, whereas the topology of the three layers is reported in **b**. The layers corresponding to 'prepared nuts' and 'dried fruits', which are more similar to each other (that is, closer with respect to the Jensen–Shannon divergence), are indeed aggregated by the algorithm in a single cluster, whereas the 'roots and tubers' layer, which is characterized by a remarkably different topology as evident from **b**, is kept separated.

38

Clustering

II. STATISTICS OF CYCLES

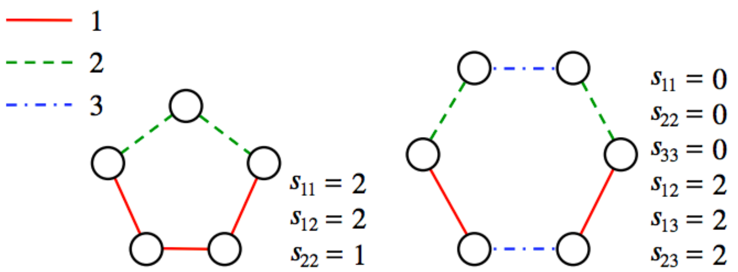


FIG. 1. Notation of multiplex cycles.

We characterize a given cycle in a multiplex network, by the matrix $\mathbf{s} = \{s_{ab}\}_{a,b=1,\ldots,M}$, where each element s_{ab} defines the number of nodes in the cycle which connect an edge of type a with an edge of type b. When

40

Big Data <-> Big Networks Albert Díaz Guilera

Centralities

Ranking in interconnected multilayer networks reveals versatile nodes

Manlio De Domenico, Albert Solé-Ribalta, Elisa Omodei, Sergio Gómez & Alex Arenas

Nature Communications **6**,
Article number: 6868 (2015)
doi:10.1038/ncomms7868

Received: 17 October 2014
Accepted: 07 March 2015
Published online: 23 April 2015

41

Communities

Detection of gene communities in multi-networks reveals cancer drivers

Laura Cantini, Enzo Medico, Santo Fortunato & Michele Caselle

Scientific Reports **5**, Article number: 17386 (2015)
doi:10.1038/srep17386

Received: 08 September 2015
Accepted: 29 October 2015
Published online: 07 December 2015

45

- To address this problem, we combined, in a single multi-network, four different gene networks:
  - (i) Transcription Factor (TF) co-targeting network,
  - (ii) microRNA co-targeting network,
  - (iii) Protein-Protein Interaction (PPI) network
  - (iv) gene co-expression network.

46

Community detection in the multi-network

After filtering, all the layers of the multi-network were sparse enough to perform community detection. The design of community detection algorithms on multi-networks is still an open problem[28]. We propose here a possible solution based on the use of the consensus clustering procedure described in[29]. We used five widely adopted algorithms: Infomap[11], OSLOM[12], Label propagation[13], Louvain[14] and Modularity optimization via simulated annealing[15]. We integrated in our software all five algorithms, leaving the choice of the preferred one to the user. We

47

Big Data <-> Big Networks
Albert Díaz Guilera

MODELS

- Individual layers
- Correlations between layers

48

Time dependent networks

50

Static network

- Topology is fixed
- The state of the nodes evolve according to:
  - Their own rules
  - The interaction with their neighbours

51

Dynamic network

- Links change!!!
- What determines the rules of change of the links?

52

Big Data <-> Big Networks
Albert Díaz Guilera

Classification

- Adaptive systems
- Time dependence of the links

53

Adaptive networks

© 2009

Adaptive Networks

Theory, Models and Applications

Editors: **Gross**, Thilo, **Sayama**, Hiroki (Eds.)

This volume is a state-of-the-art introduction to and survey on adaptive networks, combining aspects of the evolution of networks and dynamics on networks

Dynamics of epidemic diseases on a growing adaptive network

Güven Demirel, Edmund Barter & Thilo Gross

Scientific Reports **7**, Article number: 42352 (2017)
doi:10.1038/srep42352

Received: 12 October 2016
Accepted: 08 January 2017
Published online: 10 February 2017

54

- Nodes look for improving their outcome.
- “Better” neighbours, depending on node states
- Complex systems are usually self-organized, not engineered, not centralized

55

- Social agent: looking for more connected
- Economic agent: looking for more wealthy
- Access to limited information

56

Big Data <-> Big Networks
Albert Díaz Guilera

Dynamics of the system

- Importance of the two time scales in the final outcome of the complex system

58

JOURNAL OF THE ROYAL SOCIETY INTERFACE

Home Content Information for About us Sign up Submit

Adaptive coevolutionary networks: a review

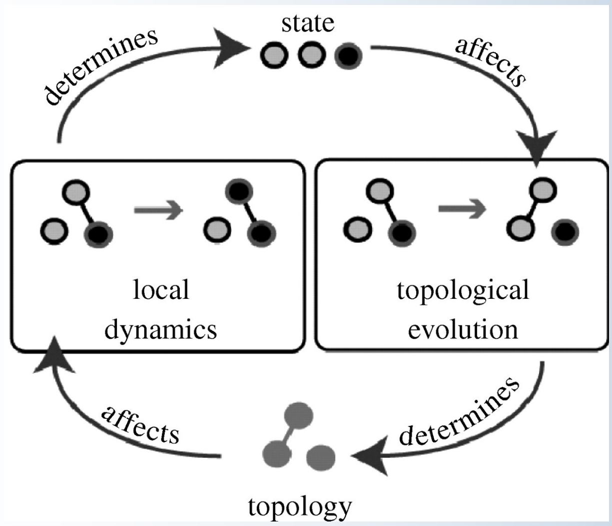
Thilo Gross, Bernd Blasius

Published 6 March 2008. DOI: 10.1098/rsif.2007.1229

59

Big Data <-> Big Networks Albert Díaz Guilera

In an adaptive network, the evolution of the topology depends on the dynamics of the nodes.



Thilo Gross, and Bernd Blasius J. R. Soc. Interface 2008;5:259-271

Interface

60

Time-dependent

- Links change because some external (not coupled to states) rules
  - Random
  - Due to mobility → contact networks

65

Physics Reports
Volume 519, Issue 3, October 2012, Pages 97–125

Temporal networks

Petter Holme[a, b, c], Jari Saramäki[d]
Show more

https://doi-org.sire.ub.edu/10.1016/j.physrep.2012.03.001
Get rights and content

66

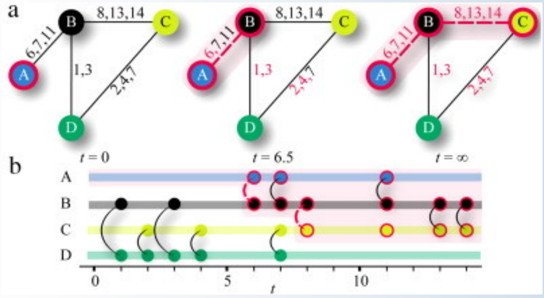


Illustration of the reachability issue and the intransitivity of temporal networks (more specifically a contact sequence). In (a), the times of the contacts between vertices A–D are indicated on the edges. Assume that, for example, a disease starts spreading at vertex A and spreads further as soon as a contact occurs. The dashed lines and vertices show this spreading process for four different times. The spreading will not continue further than what is indicated in the t=∞ picture, i.e. D cannot get infected. However, if the spreading started at vertex D, the entire set of vertices would eventually be infected. Aggregating the edges into one static graph cannot capture this effect that arises from the time ordering of contacts. Panel (b) visualizes the same situation by showing the temporal dimension explicitly. The colors of the lines in (b) match the vertex colors in (a).

Temporal networks
Petter Holme, Jari Saramäki

Physics Reports, Volume 519, Issue 3, 2012, 97–125

http://dx.doi.org/10.1016/j.physrep.2012.03.001

67

68

Albert Díaz Guilera

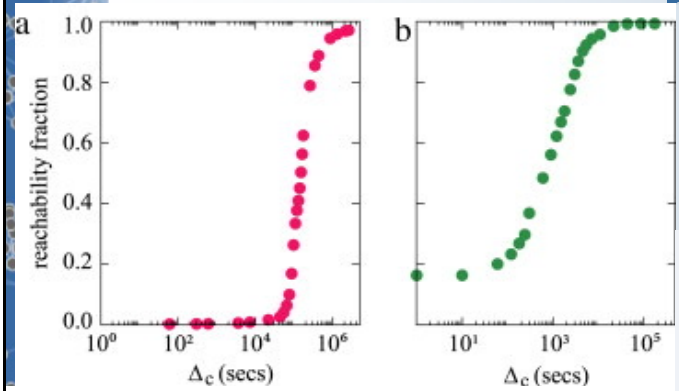
New measures
Sequence is important

•

69

Albert Díaz Guilera

Reachability



Petter Holme, Jari Saramäki
Temporal networks
Physics Reports, Volume 519, Issue 3, 2012, 97–125
http://dx.doi.org/10.1016/j.physrep.2012.03.001

The reachability ratio as a function of the maximum allowed delay of a time-respecting path at each vertex. Panel (a) shows data for a mobile telephone call network, (b) comes from the connections of an airline system.

70

Reachability graph

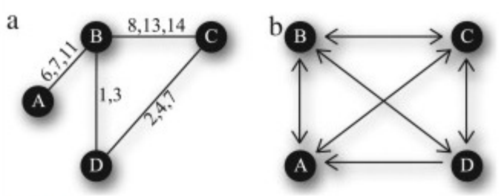


Fig. 10.
Reachability graphs. Panel (a) shows a contact sequence (same as in Fig. 1) and (b) shows its reachability graph.

71

Big Data <-> Big Networks Albert Díaz Guilera

Generalizations

- Distances
- Components
- Centrality measures
- Communities

72

Physics Reports

Volume 486, Issues 3–5, February 2010, Pages 75–174

Community detection in graphs

Santo Fortunato

Show more

https://doi-org.sire.ub.edu/10.1016/j.physrep.2009.11.002

73

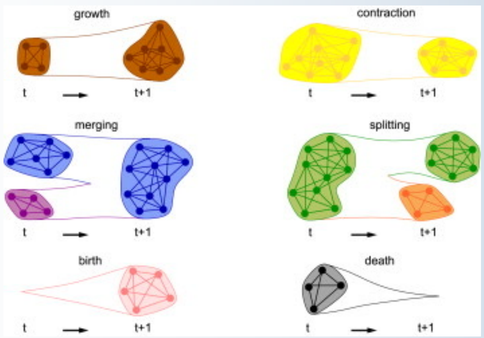


Fig. 27. Possible scenarios in the evolution of communities. Reprinted figure with permission from Ref. [380].

Santo Fortunato

Community detection in graphs

Physics Reports, Volume 486, Issues 3–5, 2010, 75–174

http://dx.doi.org/10.1016/j.physrep.2009.11.002

74

Contact networks

- Agents (humans) move
- Creating a time-dependent network of connectivities

75

SocioPatterns

http://www.sociopatterns.org/

SocioPatterns

ABOUT | GALLERY | PUBLICATIONS

WELCOME

SocioPatterns is an interdisciplinary research collaboration formed in 2008 that adopts a data-driven methodology to study social dynamics and human activity.

Since 2008, we have collected longitudinal data on the physical proximity and face-to-face contacts of individuals in numerous real-world environments, covering widely varying contexts across several countries: schools, museums, hospitals, etc. We use the data to study human behaviour and to develop agent-based models for the transmission of infectious diseases.

We make most of the collected data freely available to the scientific community.

76

Albert Díaz Guilera

Data

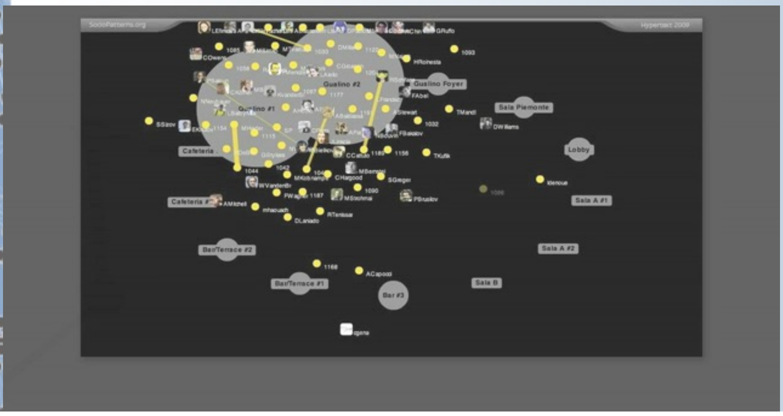
DATASETS

This page provides a collection of datasets obtained through the SocioPatterns sensing platform.

- Household
- Schools
- Hospitals
- Conferences

77

Conference

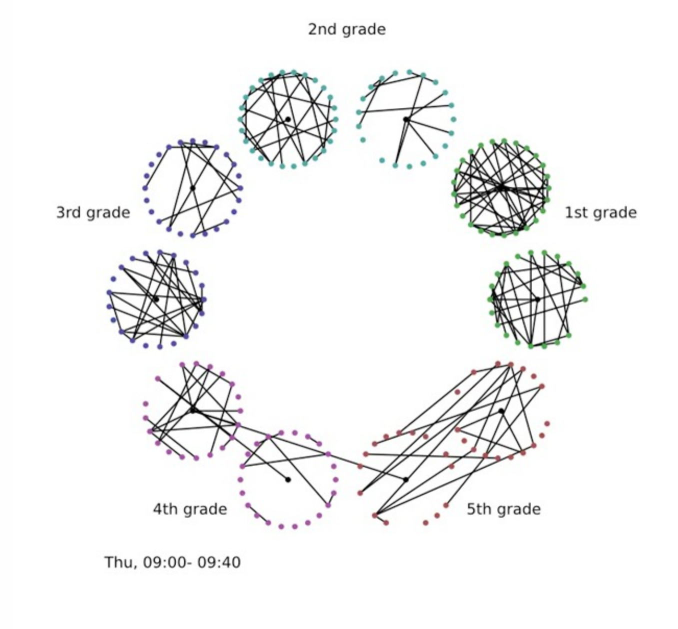


What's in a crowd? Analysis of face-to-face behavioral networks. L. Isella et al.

J. Theor. Bio. 271 (2011) 166

78

Bi *Albert Díaz Guilera*



79

Modeling Human Dynamics of Face-to-Face Interaction Networks

Michele Starnini, Andrea Baronchelli, and Romualdo Pastor-Satorras
Phys. Rev. Lett. **110**, 168701 – Published 15 April 2013

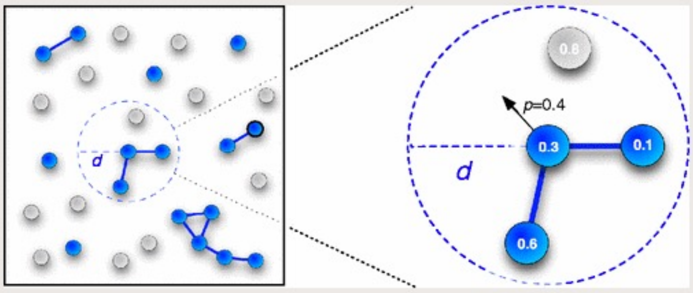


FIG. 1

+Open in new window

Left: Blue (dark) colored agents are active, gray (light) agents do not move nor interact. Interacting agents, within a distance d, are connected by a link. Right: Each individual is characterized by a number representing her attractiveness. The probability for the central individual to move is $p = 1.0 - 0.6 = 0.4$, since the attractiveness of the inactive agent is not taken into account.

80