Overview of Complex Networks

Last updated: 2024/11/11, 20:23:18 EST

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number, 2024–2025

Prof. Peter Sheridan Dodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

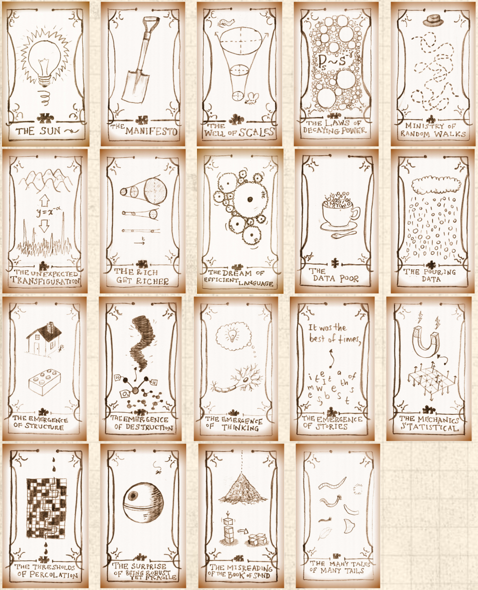
These slides are brought to you by:

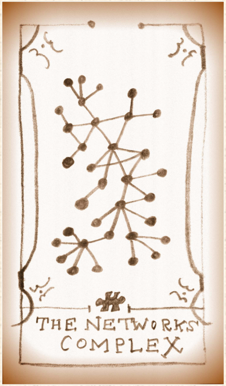
These slides are also brought to you by:

Special Guest Executive Producer

On Instagram at pratchett_the_cat

Outline


THE SUN
THE MANIFESTO
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
MINISTRY OF RANDOM WALKS
THE UNEXPECTED TRANSFIGURATION
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE
THE DATA POOR
THE POURING DATA
THE EMERGENCE OF STRUCTURE
THE EMERGENCE OF DESTRUCTION
THE EMERGENCE OF THINKING
It was the best of times,
THE EMERGENCE OF STORIES
THE MECHANICS STATISTICAL
THE THRESHOLDS OF PERCOLATION
THE SURPRISE OF BEING ROBUST YET FRAGILE
THE MISREADING OF THE BOOK OF SAND
THE MANY TALES OF MANY TAILS


THE NETWORKS COMPLEX

net•work |ˈnetˌwərk|

noun

1 an arrangement of intersecting horizontal and vertical lines.

- a complex system of roads, railroads, or other transportation routes *: a network of railroads.*

2 a group or system of interconnected people or things *: a trade network.*

- a group of people who exchange information, contacts, and experience for professional or social purposes *: a support network.*
- a group of broadcasting stations that connect for the simultaneous broadcast of a program *: the introduction of a second TV network* | [as adj.] *network television.*
- a number of interconnected computers, machines, or operations *: specialized computers that manage multiple outside connections to a network | a local cellular phone network.*
- a system of connected electrical conductors.

verb [trans.]

connect as or operate with a network *: the stock exchanges have proven to be resourceful in networking these deals.*

- link (machines, esp. computers) to operate interactively : [as adj.] (**networked**) *networked workstations.*
- [intrans.] [often as n.] (**networking**) interact with other people to exchange information and develop contacts, esp. to further one's career *: the skills of networking, bargaining, and negotiation.*

Thesaurus deliciousness:

network

noun

1 *a network of arteries* WEB, lattice, net, matrix, mesh, crisscross, grid, reticulum, reticulation; Anatomy plexus.

2 *a network of lanes* MAZE, labyrinth, warren, tangle.

3 *a network of friends* SYSTEM, complex, nexus, web, webwork.

Ancestry:

From Keith Briggs's etymological investigation:

- Opus reticulatum:
- A Latin origin?

[http://serialconsign.com/2007/11/we-put-net-network]

Ancestry:

First known use: Geneva Bible, 1560

'And thou shalt make unto it a grate like networke of brass (Exodus xxvii 4).'

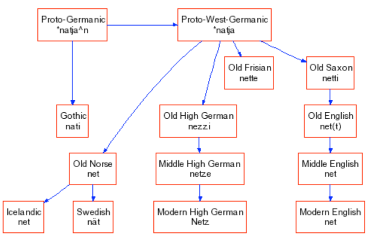
From the OED via Briggs:

- 1658–: reticulate structures in animals
- 1839–: rivers and canals
- 1869–: railways
- 1883–: distribution network of electrical cables
- 1914–: wireless broadcasting networks

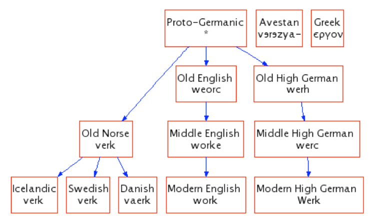
Ancestry:

Net and Work are venerable old words:

- 'Net' first used to mean spider web (King Ælfréd, 888).
- 'Work' appear to have long meant purposeful action.



The network of Germanic 'net' words

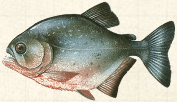


The network of 'work' words

- 'Network' = something built based on the idea of natural, flexible lattice or web.
- c.f., ironwork, stonework, fretwork.

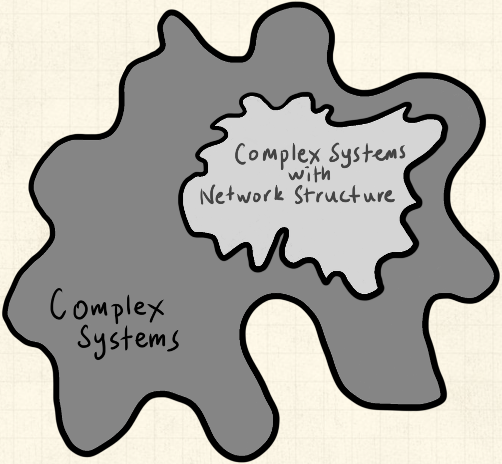
Key Observation:

- Many complex systems
  can be viewed as complex networks
  of physical or abstract interactions.
- Opens door to mathematical and numerical analysis.
- Dominant approach of the first decade was of a theoretical-physics/stat-mechish flavor.
- Mindboggling amount of work published on complex networks since 1998...
- ... largely due to your typical theoretical physicist:



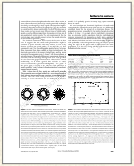
- *Piranha physicus*
- Hunt in packs.
- Feast on new and interesting ideas (see chaos, cellular automata, ...)
- See also: https://xkcd.com/793/

Complex Systems is the Big Story:



- Only a bit networky: Fluids-at-large (the atmosphere, oceans, ...), organism cells, ...

Popularity (according to Google Scholar)

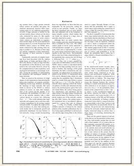


"Collective dynamics of 'small-world' networks"

Watts and Strogatz,

Nature, **393**, 440–442, 1998. [20]

Times cited: $\sim$ 27,184 (as of October 8, 2015)



"Emergence of scaling in random networks"

Barabási and Albert,

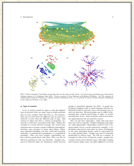
Science, **286**, 509–511, 1999. [2]

Times cited: $\sim$ 23,532 (as of October 8, 2015)

Review articles:

"Complex Networks: Structure and Dynamics"
Boccaletti et al.,
Physics Reports, **424**, 175–308, 2006. [3]

Times cited: ~ 5,791 (as of October 8, 2015)



"The structure and function of complex networks"
M. E. J. Newman,
SIAM Rev., **45**, 167–256, 2003. [15]

Times cited: ~ 13,156 (as of October 8, 2015)

"Statistical mechanics of complex networks"
Albert and Barabási,
Rev. Mod. Phys., **74**, 47–97, 2002. [1]

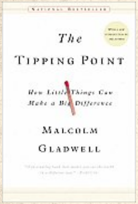
Times cited: ~ 26,636 (as of May 9, 2023)

Popularity according to textbooks:

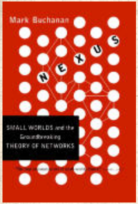
Textbooks:

- Mark Newman (Physics, Michigan)
  "Networks: An Introduction"
- David Easley and Jon Kleinberg (Economics and Computer Science, Cornell)
  "Networks, Crowds, and Markets: Reasoning About a Highly Connected World"

Popularity according to popular books:

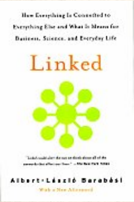


The Tipping Point: How Little Things can make a Big Difference—Malcolm Gladwell [9]

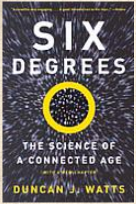


Nexus: Small Worlds and the Groundbreaking Science of Networks—Mark Buchanan

Popularity according to popular books:



Linked: How Everything Is Connected to Everything Else and What It Means—Albert-Laszlo Barabási



Six Degrees: The Science of a Connected Age—Duncan Watts [19]

Numerous others ...

- Complex Social Networks—F. Vega-Redondo [18]
- Fractal River Basins: Chance and Self-Organization—I. Rodríguez-Iturbe and A. Rinaldo [16]
- Random Graph Dynamics—R. Durette
- Scale-Free Networks—Guido Caldarelli
- Evolution and Structure of the Internet: A Statistical Physics Approach—Romu Pastor-Satorras and Alessandro Vespignani
- Complex Graphs and Networks—Fan Chung
- Social Network Analysis—Stanley Wasserman and Kathleen Faust
- Handbook of Graphs and Networks—Eds: Stefan Bornholdt and H. G. Schuster [5]
- Evolution of Networks—S. N. Dorogovtsev and J. F. F. Mendes [8]

More observations

- But surely networks aren’t new ...
- Graph theory was well established ...
- Study of social networks started in the 1930’s ...
- So why all this ‘new’ research on networks?
- Answer: Oodles of Easily Accessible Data.
- We can now inform (alas) our theories with a much more measurable reality.*
- Graph theory missed “becoming”: Stories = Characters + Time
- A worthy goal: establish mechanistic explanations.

**If this is upsetting, maybe string theory is for you ...*

More observations

- Internet-scale data sets can be overly exciting.

Witness:

- The End of Theory: The Data Deluge Makes the Scientific Theory Obsolete (Anderson, Wired)
- "The Unreasonable Effectiveness of Data," Halevy et al. [10].
- c.f. Wigner's "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" [21]

But:

- For scientists, description is only part of the battle.
- We still need to understand.

Super Basic definitions

Nodes = A collection of entities which have properties that are somehow related to each other

- e.g., people, forks in rivers, proteins, webpages, organisms, ...

Links = Connections between nodes

- Links may be directed or undirected.
- Links may be binary or weighted.

Other spiffing words: vertices and edges.

Super Basic definitions

Node degree = Number of links per node

- Notation: Node i's degree = k_i.
- k_i = 0,1,2,....
- Notation: the average degree of a network = $\langle k \rangle$ (and sometimes z)
- Connection between number of edges m and average degree:

$$\langle k \rangle = \frac{2m}{N}.$$

- Defn: N_i = the set of i's k_i neighbors

Super Basic definitions

Adjacency matrix:

- We can represent a network by a matrix A with link weight a_{ij} for nodes i and j in entry (i,j).
- e.g.,

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

- For numerical work, we always use sparse matrices.
- For many real networks, A is a function of time.

Examples

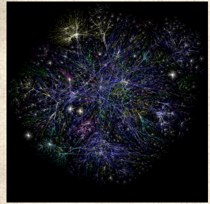
So what passes for a complex network?

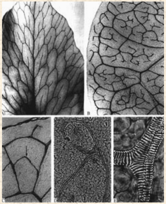
- Complex networks are large (in node number)
- Complex networks are sparse (low edge to node ratio)
- Complex networks are usually dynamic and evolving
- Complex networks can be social, economic, natural, informational, abstract, ...

Examples

Physical networks

- River networks
- Neural networks
- Trees and leaves
- Blood networks
- The Internet (pipes)
- Road networks
- Power grids





- Distribution (branching) versus redistribution (cyclical)

Examples

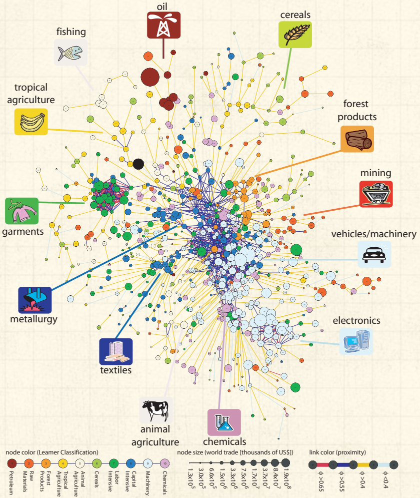
Interaction networks

- The Blogosphere (RIP)
- Biochemical networks
- Gene-protein networks
- Food webs: who eats whom
- Airline networks
- Call networks (AT&T)
- The Media
- The internet (World Wide Web)

datamining.typepad.com

topics:

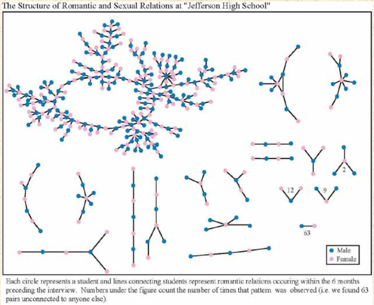
- Hidalgo et al.'s "The Product Space Conditions the Development of Nations" [11]
- How do products depend on each other, and how does this network evolve?
- How do countries depend on each other for water, energy, people (immigration), investments?



Examples

Interaction networks: social networks

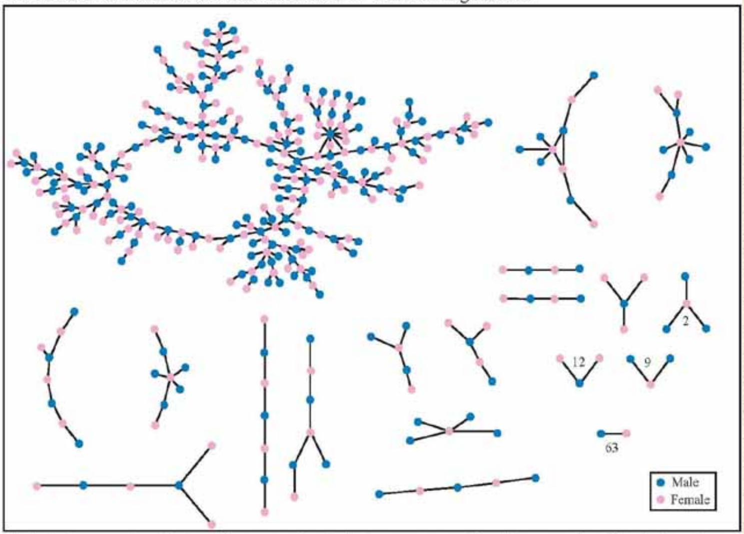
- Snogging
- Friendships
- Acquaintances
- Boards and directors
- Organizations
- facebook twitter,



(Bearman *et al.*, 2004)

- 'Remotely sensed' by: email activity, instant messaging, phone logs (*cough*).

Examples



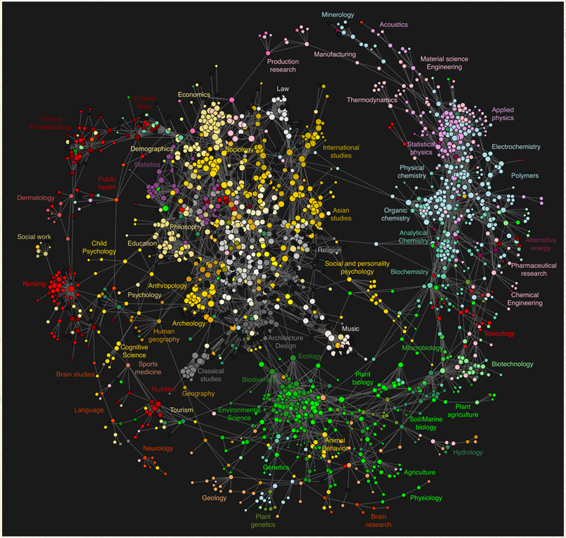
Each circle represents a student and lines connecting students represent romantic relations occuring within the 6 months preceding the interview. Numbers under the figure count the number of times that pattern was observed (i.e. we found 63 pairs unconnected to anyone else).

Examples

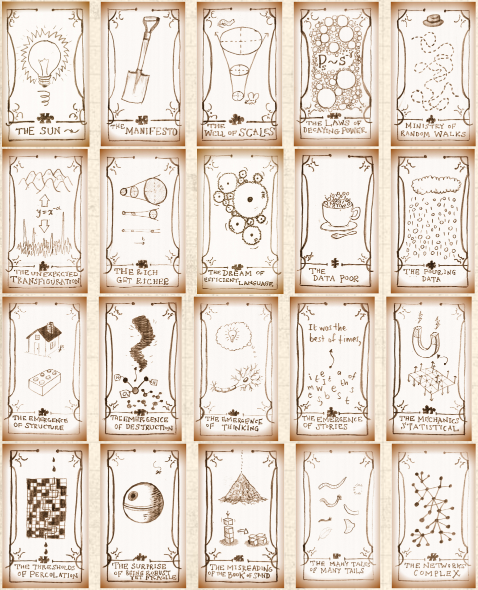
Relational networks

- Consumer purchases (Walmart, Target, Amazon, ...)
- Thesauri: Networks of words generated by meanings
- Knowledge/Databases/Ideas
- Metadata—Tagging, Keywords bit.ly flickr
- Large Language Models

Clickworthy Science:



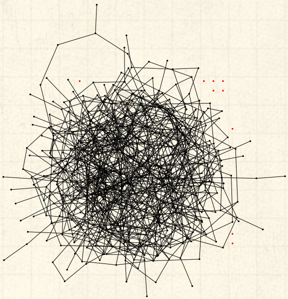
"Clickstream Data Yields High-Resolution Maps of Science", Bollen et al. [4], 2009.


THE SUN
THE MANIFESTO
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
MINISTRY OF RANDOM WALKS
THE UNEXPECTED TRANSFIGURATION
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE
THE DATA POOR
THE POURING DATA
THE EMERGENCE OF STRUCTURE
THE EMERGENCE OF DESTRUCTION
THE EMERGENCE OF THINKING
It was the best of times,
THE EMERGENCE OF STORIES
THE MECHANICS STATISTICAL
THE THRESHOLDS OF PERCOLATION
THE SURPRISE OF BEING ROBUST YET FRAGILE
THE MISREADING OF THE BOOK OF SAND
THE MANY TALES OF MANY TAILS
THE NETWORKS COMPLEX


THE ELEMENTS OF WEBS

A notable feature of large-scale networks:

- Graphical renderings are often just a big mess.



⇐ Typical hairball

- number of nodes $N = 500$
- number of edges $m = 1000$
- average degree $\langle k \rangle = 4$

- And even when renderings somehow look good:
  "That is a very graphic analogy which aids understanding wonderfully while being, strictly speaking, wrong in every possible way"
  said Ponder [Stibbons] —*Making Money*, T. Pratchett.
- We need to extract digestible, meaningful aspects.

Some key aspects of real complex networks:

- degree distribution*
- assortativity
- homophily
- clustering
- motifs
- modularity
- concurrency
- hierarchical scaling
- network distances
- centrality
- efficiency
- interconnectedness
- robustness

- Plus coevolution of network structure and processes on networks.

* Degree distribution is the elephant in the room that we are now all very aware of ...

Properties

1. degree distribution P_k

- P_k is the probability that a randomly selected node has degree k.
- k = node degree = number of connections.
- ex 1: Pure (Erdős-Rényi) random networks have Poisson degree distributions:
  Insert assignment question

$$P_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

- ex 2: "Scale-free" networks: $P_k \propto k^{-\gamma} \Rightarrow$ 'hubs'.
- link cost controls skew.
- hubs may facilitate or impede contagion.

Properties

Note:

- Pure (Erdős-Rényi) random networks are a *mathematical construct*.
- 'Scale-free' networks are growing networks that form according to a plausible mechanism.
- Randomness is out there, just not to the degree of a completely random network.
- "Becoming": Stories = Characters + Time

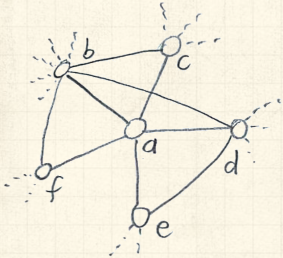
Properties

2. Assortativity/3. Homophily:

- Social networks: Homophily = birds of a feather
- e.g., degree is standard property for sorting: measure degree-degree correlations.
- Assortative network: [14] similar degree nodes connecting to each other.
  Often social: company directors, coauthors, actors.
- Disassortative network: high degree nodes connecting to low degree nodes.
  Often techological or biological: Internet, WWW, protein interactions, neural networks, food webs.

Local socialness:

4. Clustering:



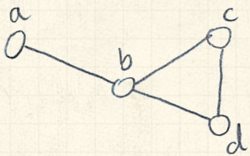
- Your friends tend to know each other.
- Two measures (explained on following slides):
  1. Watts & Strogatz [20]

$$C_1 = \left\langle \frac{\sum_{j_1 j_2 \in N_i} a_{j_1 j_2}}{k_i(k_i-1)/2} \right\rangle_i$$

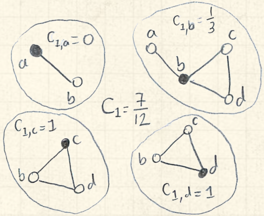
  2. Newman [15]

$$C_2 = \frac{3 \times \#\text{triangles}}{\#\text{triples}}$$

Example network:



Calculation of C_1:



- C_1 is the average fraction of pairs of neighbors who are connected.
- Fraction of pairs of neighbors who are connected is

$$\frac{\sum_{j_1 j_2 \in N_i} a_{j_1 j_2}}{k_i(k_i - 1)/2}$$

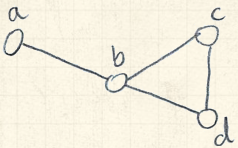
where k_i is node i's degree, and N_i is the set of i's neighbors.

- Averaging over all nodes, we have:

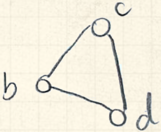
$$C_1 = \frac{1}{n}\sum_{i=1}^{n} \frac{\sum_{j_1 j_2 \in N_i} a_{j_1 j_2}}{k_i(k_i-1)/2} = \left\langle \frac{\sum_{j_1 j_2 \in N_i} a_{j_1 j_2}}{k_i(k_i-1)/2} \right\rangle_i$$

Triples and triangles

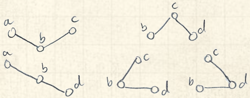
Example network:



Triangles:



Triples:



- Nodes i_1, i_2, and i_3 form a triple around i_1 if i_1 is connected to i_2 and i_3.
- Nodes i_1, i_2, and i_3 form a triangle if each pair of nodes is connected
- The definition $C_2 = \frac{3 \times \#\text{triangles}}{\#\text{triples}}$ measures the fraction of closed triples
- The '3' appears because for each triangle, we have 3 closed triples.
- Social Network Analysis (SNA): fraction of transitive triples.

Clustering:

Sneaky counting for undirected, unweighted networks:

- If the path i–j–ℓ exists then $a_{ij}a_{j\ell} = 1$.
- Otherwise, $a_{ij}a_{j\ell} = 0$.
- We want $i \neq \ell$ for good triples.
- In general, a path of n edges between nodes i_1 and i_n travelling through nodes i_2, i_3, ...i_{n-1} exists $\iff$ $a_{i_1 i_2}a_{i_2 i_3}a_{i_3 i_4} \cdots a_{i_{n-2} i_{n-1}}a_{i_{n-1} i_n} = 1$.
- $$\#\text{triples} = \frac{1}{2}\left(\sum_{i=1}^{N}\sum_{\ell=1}^{N}\left[A^2\right]_{i\ell} - \text{Tr}A^2\right)$$
- $$\#\text{triangles} = \frac{1}{6}\text{Tr}A^3$$

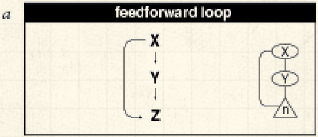
Properties

- For sparse networks, C_1 tends to discount highly connected nodes.
- C_2 is a useful and often preferred variant
- In general, $C_1 \neq C_2$.
- C_1 is a global average of a local ratio.
- C_2 is a ratio of two global quantities.

Properties

5. motifs:

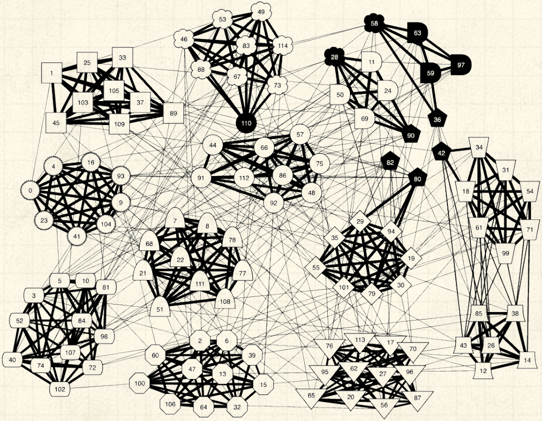
- small, recurring functional subnetworks
- e.g., Feed Forward Loop:



Shen-Orr, Uri Alon, *et al.* [17]

Properties

6. modularity and structure/community detection:



Clauset *et al.*, 2006 [7]: NCAA football

Properties

7. concurrency:

- transmission of a contagious element only occurs during contact
- rather obvious but easily missed in a simple model
- dynamic property—static networks are not enough
- knowledge of previous contacts crucial
- beware cumulated network data
- Kretzschmar and Morris, 1996 [13]
- "Temporal networks" become a concrete area of study for Piranha Physicus in 2013.

Properties

8. Horton-Strahler ratios:

- Metrics for branching networks:
  - Method for ordering streams hierarchically
  - Number: $R_n = N_\omega / N_{\omega+1}$
  - Segment length: $R_l = \langle l_{\omega+1} \rangle / \langle l_\omega \rangle$
  - Area/Volume: $R_a = \langle a_{\omega+1} \rangle / \langle a_\omega \rangle$

Properties

9. network distances:

(a) shortest path length d_{ij}:

- Fewest number of steps between nodes i and j.
- (Also called the chemical distance between i and j.)

(b) average path length $\langle d_{ij} \rangle$:

- Average shortest path length in whole network.
- Good algorithms exist for calculation.
- Weighted links can be accommodated.

Properties

9. network distances:

- network diameter d_{max}:
  Maximum shortest path length between any two nodes.
- closeness $d_{cl} = [\sum_{ij} d_{ij}^{-1} / \binom{n}{2}]^{-1}$:
  Average 'distance' between any two nodes.
- Closeness handles disconnected networks ($d_{ij} = \infty$)
- $d_{cl} = \infty$ only when all nodes are isolated.
- Closeness perhaps compresses too much into one number

Properties

10. centrality:

- Many such measures of a node's 'importance.'
- ex 1: Degree centrality: k_i.
- ex 2: Node i's betweenness
  = fraction of shortest paths that pass through i.
- ex 3: Edge ℓ's betweenness
  = fraction of shortest paths that travel along ℓ.
- ex 4: Recursive centrality: Hubs and Authorities (Jon Kleinberg [12])

Properties

Interconnected networks and robustness (two for one deal):

"Catastrophic cascade of failures in interdependent networks" [6]. Buldyrev et al., Nature 2010.

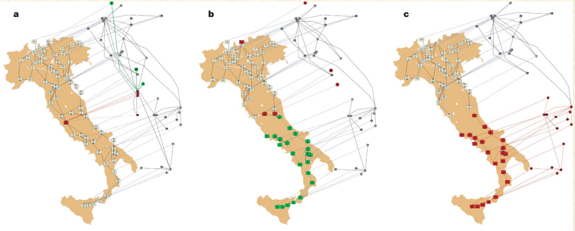


Figure 1 | Modelling a blackout in Italy. Illustration of an iterative process of a cascade of failures using real-world data from a power network (located on the map of Italy) and an Internet network (shifted above the map) that were implicated in an electrical blackout that occurred in Italy in September 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).

Nutshell:

Overview Key Points:

- The field of complex networks came into existence in the late 1990s.
- Explosion of papers and interest since 1998/99.
- Hardened up much thinking about complex systems.
- Specific focus on networks that are large-scale, sparse, natural or people-made, evolving and dynamic, and (crucially) measurable.
- Three main (blurred) categories:
  1. Physical (e.g., river networks),
  2. Interactional (e.g., social networks),
  3. Abstract (e.g., thesauri).
- To solve network problems: "Follow the edges."

scale-free-networks,

References I

[1] R. Albert and A.-L. Barabási.
Statistical mechanics of complex networks.
Rev. Mod. Phys., 74:47–97, 2002. pdf

[2] A.-L. Barabási and R. Albert.
Emergence of scaling in random networks.
Science, 286:509–511, 1999. pdf

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang.
Complex networks: Structure and dynamics.
Physics Reports, 424:175–308, 2006. pdf

[4] J. Bollen, H. Van de Sompel, A. Hagberg, L. Bettencourt, R. Chute, M. A. Rodriguez, and B. Lyudmila.
Clickstream data yields high-resolution maps of science.
PLoS ONE, 4:e4803, 2009. pdf

References II

[5] S. Bornholdt and H. G. Schuster, editors.
Handbook of Graphs and Networks.
Wiley-VCH, Berlin, 2003.

[6] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin.
Catastrophic cascade of failures in interdependent networks.
Nature, 464:1025–1028, 2010. pdf

[7] A. Clauset, C. Moore, and M. E. J. Newman.
Structural inference of hierarchies in networks, 2006. pdf

[8] S. N. Dorogovtsev and J. F. F. Mendes.
Evolution of Networks.
Oxford University Press, Oxford, UK, 2003.

References III

[9] M. Gladwell.
The Tipping Point.
Little, Brown and Company, New York, 2000.

[10] A. Halevy, P. Norvig, and F. Pereira.
The unreasonable effectiveness of data.
IEEE Intelligent Systems, 24:8–12, 2009. pdf

[11] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausman.
The product space conditions the development of nations.
Science, 317:482–487, 2007. pdf

[12] J. M. Kleinberg.
Authoritative sources in a hyperlinked environment.
Proc. 9th ACM-SIAM Symposium on Discrete Algorithms, 1998. pdf

References IV

[13] M. Kretzschmar and M. Morris.
Measures of concurrency in networks and the spread of infectious disease.
Math. Biosci., 133:165–95, 1996. pdf

[14] M. Newman.
Assortative mixing in networks.
Phys. Rev. Lett., 89:208701, 2002. pdf

[15] M. E. J. Newman.
The structure and function of complex networks.
SIAM Rev., 45(2):167–256, 2003. pdf

[16] I. Rodríguez-Iturbe and A. Rinaldo.
Fractal River Basins: Chance and Self-Organization.
Cambridge University Press, Cambrigde, UK, 1997.

References V

[17] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon.
Network motifs in the transcriptional regulation network of *Escherichia coli*.
Nature Genetics, 31:64–68, 2002. pdf

[18] F. Vega-Redondo.
Complex Social Networks.
Cambridge University Press, 2007.

[19] D. J. Watts.
Six Degrees.
Norton, New York, 2003.

[20] D. J. Watts and S. J. Strogatz.
Collective dynamics of 'small-world' networks.
Nature, 393:440–442, 1998. pdf

References VI

[21] E. Wigner.
The unreasonable effectivenss of mathematics in the natural sciences.
Communications on Pure and Applied Mathematics, 13:1–14, 1960. pdf