Semester projects
Principles of Complex Systems
CSYS/MATH 300, Spring, 2013

Prof. Peter Dodds
@peterdodds

Department of Mathematics & Statistics | Center for Complex Systems | Vermont Advanced Computing Center | University of Vermont

Narrative hierarchy

Presenting at many scales:
1. 1 to 3 word encapsulation, a soundbite,
2. a sentence/title,
3. a few sentences,
4. a paragraph,
5. a short paper,
6. a long paper,
7. ...

Twitter—living in the now:

Research opportunity: be involved in our socio-info-algorithmo-econo-geo-technico-physical systems research group studying Twitter and other wordful large data sets.

Topics:
1. Develop and elaborate an online experiment to study some aspect of social phenomena
2. e.g., collective search, cooperation, cheating, influence, creation, decision-making, etc.
3. Part of the PLAY project.

Outline

The Plan
Suggestions for Projects
References

Semester projects

Requirements:
1. 3 minute introduction to project (5th week).
2. 5-10 minute final presentation.
3. Report: ≥ 5 pages (single space), journal-style

Goals:
1. Understand, critique, and communicate published work.
2. Seed research papers or help papers along.

References

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.


Rummage round in the papers we’ve covered in our weekly Complex Systems Reading Group at UVM.

Semester projects

The Plan

Sociotechnical phenomena—Foldit:

Suggestions for Problems

References

Figure 1. Multi-resolution visualization tools and roadblocks: The visualization toolkit and the roadblocks. (a) The visualization tool. (b) The roadblocks. The visualization tool is a web-based tool that allows users to upload their own data and to visualize it in a variety of ways. The roadblocks are a set of visualizations that are created by the tool to help users understand their data. The roadblocks include a heat map, a scatter plot, a box plot, and a density plot.

> “Predicting protein structures with a multiplayer online game.” Cooper et al., Nature, 2010. [14]

Also: zooniverse [6], ESP game [7], captchas [8].

Semester projects


Figure 1. (a) Development of the cascading failures. The cascade of failures is triggered by a small number of nodes that are removed from the network. The cascade then propagates and affects a large number of nodes. The cascade is self-sustaining and continues until the network is broken. (b) Recovery of the network. The network is recovered by adding new nodes to the network. The network is then restored to its original state.

Semester projects

Voting

Score-based voting versus rank-based voting:

> Balinski and Laraki [2]

“A theory of measuring, electing, and ranking”


Figure 1. (a) Score-based voting. The score-based voting method is used to select the best candidate. The candidate with the highest score is selected. (b) Rank-based voting. The rank-based voting method is used to select the best candidate. The candidates are ranked based on their scores. The candidate with the highest rank is selected.
The madness of modern geography:

- Explore distances between points on the Earth as travel times.
- See Jonathan Harris’s work here (II) and here (III).

Topics:

- Explore general theories on system robustness.
- Are there universal signatures that presage system failure?
- See “Early-warning signals for critical transitions” Scheffer et al., Nature 2009. [35]
- “Although predicting such critical points before they are reached is extremely difficult, work in different scientific fields is now suggesting the existence of generic early-warning signals that may indicate for a wide class of systems if a critical threshold is approaching.”
- Later in class: Doyle et al., robust-yet-fragile systems

The Plan
Suggestions for Problems
References

Topics:

- Study the human disease and disease gene networks (Goh et al., 2007):

The problem of missing data in networks:

- Clauset et al. (2008) “Hierarchical structure and the prediction of missing links in networks” [12]
Self-similarity of complex networks

Chaoming Song

NATURE | VOL 433 | 27 JANUARY 2005 | www.nature.com/nature

Complex networks have been studied extensively owing to their relevance to many real systems such as the world-wide web, the Internet, energy landscapes and biological and social networks. A common metric in complex networks is the degree exponent, which determines how the number of nodes increases with the 'diameter' of networks. These networks are not self-similar, since self-similarity requires a power-law relation between the number of boxes needed to cover the network and the size of the box, where \( l \) is the 'size' of the box.

Page 1

Related papers:

- "Origins of fractality in the growth of complex networks" Song et al. (2006a)
- "Skeleton and Fractal Scaling in Complex Networks" Go et al. (2006a)
- "Complex Networks Renormalization: Flows and Fixed Points" Radicchi et al. (2008a)

Page 2

Explore "self-similarity of complex networks" [36, 37]
See accompanying comment by Strogatz [38]
See also "Coarse-graining and self-dissimilarity of complex networks" by Itzkovitz et al. [39]

Page 3

"Looking at Gielen's work, it's tempting to propose a new branch of the human sciences: geometric sociology, a study of nothing but the shapes our inhabited spaces make. Its research agenda would ask why these forms, angles and geometries emerge so consistently, from prehistoric settlements to the fringes of exuberia. Are sites like these an aesthetic pursuit, a mathematical accident, a calculated bending of property lines based on glitches in the local planning code or an emergent combination of all these factors? Or are they the expression of something buried deep in human culture and the unconscious, something only visible from high above?"

http://opinionator.blogs.nytimes.com/the-geometry-of-sprawl/

Page 4

Study collective creativity arising out of social interactions
Productivity, wealth, creativity, disease, etc. appear to increase superlinearly with population
Start with Bettencourt et al.'s "Growth, innovation, scaling, and the pace of life in cities" [4]
Lastly, agents that remain inactive for longer than a certain number of steps are removed from the network. This rule is motivated by the observation that agents can become permanent non-participants in the network. For concreteness, let us assume that an agent with a certain probability $p$ will join the network as a new member, and another probability $q$ will choose an incumbent to be added to the team. Let us further assume that the network reaches a steady state after a transient period. The model predicts that the network evolves according to the following rules:

1. For each of the five fields for which we consider the network, we used our model to produce the network of collaborations, one at a time. The fractions $p$ and $q$ were considered for each field, and the model enabled us to simulate the evolution of the network.

2. We found that the network of collaborations, for each of the five fields, is similar for three of the four fields considered: economics, ecology, and astronomy. The fraction of new members is $p = 0.32$, and for the other two fields, we found $p = 0.32$ and $p = 0.35$.

3. We also observed that the fraction of new members is $q = 0.4$ for one field and $q = 0.5$ for the other two fields.

4. For the second and subsequent agents selected from the incumbents' pool: (i) with probability $m = 0.3$, we were drawn from the pool of incumbents and a probability $1 - m$ from the pool of newcomers. We used this parameter to investigate how the network evolves under different conditions.

5. We found that the system undergoes a percolation transition at a critical line $\theta = 0.55$. That is, the system experiences a transition from a network with a large connected component to a network with a large isolated component. For values of $\theta$ less than 0.55, the network contains a large connected cluster.

6. For each field, we found that the network evolves according to the following rules: (a) the network changes very slowly. In contrast, low turnover and very fast dynamics. (b) The network of collaborations among the practitioners, comprising a substantial fraction of the agents or is the network of collaborations among the practitioners, comprising most of the agents.

7. The study of collaboration networks is triguingly, the relative sizes of the giant components of the network depend on each other for water, energy, people (immigration), investments, and other factors.

8. The study of collaboration networks is also related to situations where multiple people collaborate on a task. In these situations, the network of collaborations is often referred to as the "invisible college." The invisible college is a network of collaborations among the practitioners, comprising a substantial fraction of the agents.

9. We will study some of this in class...
Study games (as in game theory) on networks.
Much work to explore: voter models, contagion-type models, etc.

Semantic networks: explore word-word connection networks generated by linking semantically related words.
Also: Networks based on morphological or phonetic similarity.
More general: Explore language evolution
One paper to start with: “The small world of human language” by Ferrer i Cancho and Solé [18]
Study spreading of neologisms.
Examine new words relative to existing words—is there a pattern? Phonetic and morphological similarities.
Crazy: Can new words be predicted?
Use Google Books n-grams as a data source.

Explore proposed measures of system complexity.
Study Stuart Kauffman’s nk boolean networks which model regulatory gene networks [26]

Review: Study Castronova’s and others’ work on massive multiplayer online games. How do social networks form in these games? [9]
See work by Johnson et al. on gang formation in the real world and in World of Warcraft (really).
Study phyllotaxis, how plants grow new buds and branches.

Some delightful mathematics appears involving the Fibonacci series.

Excellent work to start with:

“Phyllotaxis as a Dynamical Self Organizing Process: Parts I, II, and III” by Douady and Couder

▶

Social networks:

Study social networks as revealed by email patterns, Facebook connections, tweets, etc.


“Community Structure in Online Collegiate Social Networks” Traud et al., 2008.

Study how the Wikipedia’s content is interconnected.

▶

Vague/Large:

Study amazon’s recommender networks.

See work by Sornette et al.

Vague/Large:

Study Netflix’s open data (movies and people form a bipartite graph).

How do countries depend on each other for water, energy, people (immigration), investments?

How is the media connected? Who copies whom?

(Problem: Need to be able to measure interactions.)

Investigate memetics, the ‘science’ of memes.

▶

Sport...
More Vague/Large:

- Study spreading of anything where influence can be measured (very hard).
- Study any interesting micro-macro story to do with evolution, biology, ethics, religion, history, food, international relations, . . .
- Data is key.

References I

The life-spans of empires. 

A theory of measuring, electing, and ranking. 

Shape and Structure, from Engineering to Nature. 

References II

Growth, innovation, scaling, and the pace of life in cities. 

Common ecology quantifies human insurgency. 

The scaling laws of human travel. 

References III

Catastrophic cascade of failures in interdependent networks. 

Alone in the crowd: The structure and spread of loneliness in a large social network. 

Synthetic Worlds: The Business and Culture of Online Games. 

References IV

The spread of obesity in a large social network over 32 years. 

The collective dynamics of smoking in a large social network. 

Hierarchical structure and the prediction of missing links in networks. 

References V

On the Frequency of Severe Terrorist Events. 

Predicting protein structures with a multiplayer online game. 

Phyllotaxis as a dynamical self organizing process Part I: The spiral modes resulting from time-periodic iterations. 
References VI

Phyllotaxis as a dynamical self organizing process
Part II: The spontaneous formation of a periodicity
and the coexistence of spiral and whorled patterns.

Phyllotaxis as a dynamical self organizing process
Part III: The simulation of the transient regimes of
ontogeny.

The small world of human language.

References VII

Dynamic spread of happiness in a large social
network: longitudinal analysis over 20 years in the
Framingham Heart Study.
BMJ, 337:article #2338, 2008. pdf

Skeleton and fractal scaling in complex networks.

Understanding individual human mobility patterns.

References VIII

Team assembly mechanisms determine
collaboration network structure and team
performance.

[23] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and
R. Hausman.
The product space conditions the development of
nations.

Network scaling reveals consistent fractal pattern in
hierarchical mammalian societies.

References IX

O. Becerra, J. C. Bohorquez, N. Suarez, E. M.
Restrepo, and R. Zarama.
Universal patterns underlying ongoing wars and
terrorism, 2006. pdf

The Origins of Order.

[27] M. Kearns, S. Suri, and N. Montfort.
An experimental study of the coloring problem on
human subject networks.

Effects of missing data in social networks.

References X

[29] G. Kossinets and D. J. Watts.
Empirical analysis of evolving social networks.

Controllability of complex networks.

Gray, The Google Books Team, J. P. Pickett,
D. Hoiberg, D. Clancy, P. Norvig, J. Orwant,
S. Pinker, M. A. Nowak, and E. A. Lieberman.
Quantitative analysis of culture using millions of
digitized books.

References XI

Five rules for the evolution of cooperation.

Networks of scientific papers.

[34] F. Radicchi, J. J. Ramasco, A. Barrat, and
S. Fortunato.
Complex networks renormalization: Flows and fixed
points.
References XII


References XIII