Semester projects
Principles of Complex Systems
CSYS/MATH 300, Fall, 2011

Prof. Peter Dodds

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

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

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

Goals:
- Understand, critique, and communicate published work.
- Seed research papers or help papers along.
Narrative hierarchy

Presenting at many scales:

- 1 to 3 word encapsulation, a soundbite,
- a sentence/title,
- a few sentences,
- a paragraph,
- a short paper,
- a long paper,
- ...
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:

- Develop and elaborate an **online experiment** to study some aspect of **social phenomena**
- e.g., collective search, cooperation, cheating, influence, creation, decision-making, etc.
- Part of the PLAY project.
Rummage round in the papers (확) we’ve covered in our weekly Complex Systems Reading Group at UVM.
Study movement and interactions of people.


Barabasi’s group: tracking movement via cell phones [21].
matchings increases exponentially (Supplementary Information, section IV.C) and, as a result, the chance that a link does not participate in any control configuration decreases. For scale-free networks, we observe the same behaviour, with the caveat that previous work to a few dozen nodes at most (3)) to large networks is computationally prohibitive, limiting previous tools to address controllability for arbitrary network topologies (Fig. 5c, d). Our key finding, that and sizes. Our key finding, that the tools to address controllability for arbitrary network topologies is to identify and remove a set of critical links, the core, and therefore ensuring that there are no critical links in the system.

The fractions of critical (red, $l_c$), redundant (green, $l_r$) and ordinary (grey, $l_o$) links for the real networks named in Table 1. To make controllability robust to link failures, it is sufficient to double only the critical links, formally making each of these links redundant and therefore ensuring that there are no critical links in the system.

Sociotechnical phenomena—Foldit:

Figure 1 | Foldit screenshot illustrating tools and visualizations. The visualizations include a clash representing atoms that are too close (arrow 1); a hydrogen bond (arrow 2); a hydrophobic side chain with a yellow blob because it is exposed (arrow 3); a hydrophilic side chain (arrow 4); and a segment of the backbone that is red due to high residue energy (arrow 5). The players can make modifications including ‘rubber bands’ (arrow 6), which add constraints to guide automated tools, and freezing (arrow 7), which prevents degrees of freedom from changing. The user interface includes information about the player’s current status, including score (arrow 8); a leader board (arrow 9), which shows the scores of other players and groups; toolbars for accessing tools and options (arrow 10); chat for interacting with other players (arrow 11); and a ‘cookbook’ for making new automated tools or ‘recipes’ (arrow 12).

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

Also: zooniverse (('/:'), ESP game ('/'), captchhas ('/').

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[9]. 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).
Score-based voting versus rank-based voting:

- Balinski and Laraki\textsuperscript{[2]}
  “A theory of measuring, electing, and ranking”
The madness of modern geography:

- Explore distances between points on the Earth as travel times.
- See Jonathan Harris’s work [here](#) and [here](#).
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. [33]
- “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
Semester projects

The Plan

Suggestions for Projects

References

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

- Human Disease Network

![Diagram of Human Disease Network]
Explore and critique Fowler and Christakis et al. work on social contagion of:

- Obesity[10]
- Smoking cessation[11]
- Happiness[19]
- Loneliness[8]

One of many questions:

How does the (very) sparse sampling of a real social network affect their findings?

Figure 1. Loneliness clusters in the Framingham Social Network. This graph shows the largest component of friends, spouses, and siblings at Exam 7 (centered on the year 2000). There are 1,019 individuals shown. Each node represents a participant, and its shape denotes gender (circles are female, squares are male). Lines between nodes indicate relationship (red for siblings, black for friends and spouses). Node color denotes the mean number of days the focal participant and all directly connected (Distance 1) linked participants felt lonely in the past week, with yellow being 0–1 days, green being 2 days, and blue being greater than 3 days or more. The graph suggests clustering in loneliness and a relationship between being peripheral and feeling lonely, both of which are confirmed by statistical models discussed in the main text.
The problem of missing data in networks:

- Clauset et al. (2008) “Hierarchical structure and the prediction of missing links in networks” [12]
Explore “self-similarity of complex networks”[34, 35]
See accompanying comment by Strogatz[36]
See also “Coarse-graining and self-dissimilarity of complex networks” by Itzkovitz et al.[?]

The Plan
Suggestions for Projects
References
Related papers:

- “Origins of fractality in the growth of complex networks”
  Song et al. (2006a) [35]

- “Skeleton and Fractal Scaling in Complex Networks”
  Go et al. (2006a) [20]

- “Complex Networks Renormalization: Flows and Fixed Points”
  Radicchi et al. (2008a) [32]
topics:

- Explore patterns, designed and undesigned, of cities and suburbs.
"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 exurbia. 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/.../the-geometry-of-sprawl/ (䗖)
topics:

- 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]
Physics/Society—Wars: Study work that started with Lewis Richardson’s “Variation of the frequency of fatal quarrels with magnitude” in 1949.

Specifically explore Clauset et al. and Johnson et al.’s work [13, 24, 5] on terrorist attacks and civil wars

Richardson bonus: Britain’s coastline, turbulence, weather prediction, ...
Study Hidalgo et al.’s “The Product Space Conditions the Development of Nations” [22]

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?
Explore Dunbar’s number

See here and here for some food for thought regarding large-scale online games and Dunbar’s number. [http://www.lifewithalacrity.com]

Recent work: “Network scaling reveals consistent fractal pattern in hierarchical mammalian societies” Hill et al. (2008) [23].
Study scientific collaboration networks.
Mounds of data + good models.
See seminal work by De Solla Price\textsuperscript{[31]}. plus modern work by Redner, Newman, \textit{et al.}.
We will study some of this in class...
topics:

▶ Study Kearns et al.’s experimental studies of people solving classical graph theory problems\textsuperscript{[26]}
▶ “An Experimental Study of the Coloring Problem on Human Subject Networks”
▶ (Possibly) Run some of these experiments for our class.
topics:

➡️ Study **collective tagging** (or folksonomy)
➡️ e.g., del.icio.us, flickr
➡️ See work by Bernardo Huberman et al. at HP labs.
Study games (as in game theory) on networks.

For cooperation: Review Martin Nowak’s piece in Science, “Five rules for the evolution of cooperation.”\textsuperscript{[30]} and related works.

Much work to explore: voter models, contagion-type models, etc.
topics:

- **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.
topics:

- Explore proposed measures of system complexity.
- Study Stuart Kauffman’s \textit{nk boolean networks} which model regulatory gene networks\cite{25}
topics:

- Critically explore Bejan’s Constructal Theory.
- See Bejan’s book “Shape and Structure, from Engineering to Nature.” [3]
- Bejan asks why we see branching network flow structures so often in Nature—trees, rivers, etc.

- Read and critique “Historical Dynamics: Why States Rise and Fall” by Peter Turchin. [37]
- Can history Clyodynamics (/language), Psychohistory (/language), ... 
- “Big History” (/language)
- Also see “Secular Cycles” (/language).
Topics:

- Explore work by Doyle, Alderson, et al. as well as Pastor-Satorras et al. on the structure of the Internet(s).
topics:

- 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 [15, 16, 17]
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.

http://arxiv.org/abs/0809.0690 (◨)
Semester projects

The Plan

Suggestions for Projects

References

Vague/Large:

- Study amazon’s recommender networks.
  
  Customers Who Bought This Item Also Bought
  
  - Harry Potter Schoolbooks: Fantastic Beasts and... by J.K. Rowling
    
  - The Tales of Beedle the Bard, Collector's E... by J. K. Rowling
    
  - Harry, A History: The True Story of a Boy Wizar... by Melissa Anelli
    
  - Inkdeath (Inkheart) by Cornelia Funke

  See work by Sornette et al..

- Vague/Large:
  Study Netflix’s open data (movies and people form a bipartite graph).
topics:

Vague/Large:

- Study how the Wikipedia’s content is interconnected.
More Vague/Large:

- 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.
- [http://memetracker.org/](http://memetracker.org/) (⊞)
- Sport...
topics:

More Vague/Large:

- How does advertising work collectively?
- Does one car manufacturers’ ads indirectly help other car manufacturers?
- Ads for junk food versus fruits and vegetables.
- Ads for cars versus bikes versus walking.
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

The life-spans of empires.

A theory of measuring, electing, and ranking.

Shape and Structure, from Engineering to Nature.
References II


References III


References IV


References V


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 VIII


The Origins of Order. 

An experimental study of the coloring problem on 
human subject networks. 

[27] G. Kossinets. 
Effects of missing data in social networks. 

Empirical analysis of evolving social networks. 
References X


References XI


References XII

**Historical Dynamics: Why States Rise and Fall.**  