Engineering Systems
Doctoral Seminar

ESD.83  –  Fall 2011

Session 10

Faculty: Chris Magee and Joe Sussman
TA: Rebecca Kaarina Saari
Guest: Professor Abhijit Banerjee
Session 10: Agenda

- Welcome and Overview of class 10 (5 min.)
- Dialogue with Professor Banerjee (55 min)
- Break (10 min.)
- Discussion of other papers (30-40 min)
- Theme and topic integration (Magee)
  - Cumulative Progress in understanding socio-technical systems
  - Further discussion of readings;
  - Observations vs. experiments in complex socio-technical systems
  - Complexity and abstractions in observation and experiment
  - Global Poverty from an ES perspective
  - More on Six degrees of separation

- Next Steps - preparation for week 11: (5 min.)
Two Titles for today’s session

1. **In weekly note**: “Experiments in Complex systems and Global Poverty”

2. **In syllabus** “Poverty as a problem in complex systems (& empirical study)”
A Research Process

1. Development of conceptual understanding (qualitative framework)
2. Development of quantitative model
3. Observe (system)
4. Analyze observations
5. Generalize or simplify/complicate model

There are other “research processes”—what is fundamental about them for accumulation?
Idealized Research Process

Generalization

Formal Modeling

Conceptualization

Measurement

(Dis)Proving/Testing
Pathology # 1: “Guruism”

(Dis)Proving/Testing

Observation

Persuasion

Generalization

Measurement

Conceptualization

Formal Modeling

© 2009 Chris Magee, Engineering Systems Division, Massachusetts Institute of Technology
Pathology #2: “Air castling”

Generalization

(Dis)Proving/Testing

Conceptualization

Formal Modeling
A Research Process 2

1. Development of conceptual understanding (qualitative framework)

2. Development of quantitative model

3. Observe (system)
   - Design a specific version of a known procedure
   - Develop a new observational procedure
   - Find, and/or extract and combine data

4. Analyze observations
   - Use existing models to “reduce” data to model-relevant
   - Develop new models to “reduce” data
   - “Consilience” among observations of various kinds

5. Generalize or simplify/complicate model
## Session 10: Agenda

- Welcome and Overview of class 10 (5 min.)
- Dialogue with Professor Banerjee (55 min)
- Break (10 min.)
- Discussion of other papers (lead Stephen Zoepf, 30 - 40 min)
- Theme and topic integration (Magee)
  - Cumulative Progress in understanding socio-technical systems
  - Further discussion of readings;
  - Observations vs. experiments in complex socio-technical systems
  - Complexity and abstractions in observation and experiment
  - Global Poverty from an ES perspective
  - More on Six Degrees of Separation

- Next Steps - preparation for week 11: (5 min.)
Questions

1. How do Banerjee, Hall, Milgram and Watts relate to ES?
   - Are they studying Complex, Socio-technical Systems? Why or why not?
   - Are any of the challenges or pitfalls that they face applicable to ES?
   - What do they teach us about empirical observations and analysis?
Readings

- Banerjee –essential approach to observations and measurement? To overall system understanding?
- (massive) controlled experiments, “nibbling away” (with careful use of theory)
- Hall –essential approaches?:
- Massive surprisingly indirectly measurement and extensive theoretical analysis and visualization
- Milgram- essential approaches?:
- Simple creative direct test
- but was the experimental side very rigorous?
Questions

1. **How do Banerjee, Hall, Milgram and Watts relate to ES?**
   - Are they studying Complex, Socio-technical Systems? Why or why not?
   - What do they teach us about empirical observations and analysis?
   - Are any of the challenges or pitfalls that they face applicable to ES?

2. **How is empirical work done in ES?**
ES Observational Techniques

- Need for extensive data analysis and experiment vs. observational study are key differentiating factors among observational techniques

- Case studies (N = 1)
  - Implications of a singular fact
  - In-situ: Ethnographic study, surveys, interviews, document study, email studies, minutes, calendar analysis, quantitative and qualitative, etc.
  - Historical analysis: primary and secondary documents, interviews, quantitative and qualitative, etc.
  - Deep analysis and theory combined

- Medium N- as above but time limited

- High N (possibility of experiment)-
  - Randomly assigned field, natural experiment
  - Instrumental variable, laboratory, others

Look for “regularity” of various kinds
Examples from ESD 83

- Gonzalez; Dodds, Watts and Sabel; Gastner and Newman
- Zebras
- Kauffman
- Webster
- Rhodes
- Kaiser
ES Observational Techniques II

- Need for extensive data analysis and experiment vs. observational study are key differentiating factors among observational techniques
<table>
<thead>
<tr>
<th>Experiment (control of experimenter is necessary)</th>
<th>Observational Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly sophisticated quantitative analysis –use reliable theories to examine new theory</td>
<td>Many examples in natural science and in social science</td>
</tr>
<tr>
<td>Little quantitative analysis before use of data</td>
<td>Examples are not common in most natural and engineering sciences</td>
</tr>
<tr>
<td></td>
<td>Common in social science but significant qualitative analysis is often done</td>
</tr>
<tr>
<td>Experiment (control of experimenter is necessary)</td>
<td>Observational Study</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Highly sophisticated quantitative analysis</td>
<td>Many examples in natural science and in social science</td>
</tr>
<tr>
<td>Little quantitative analysis before use of data</td>
<td>Examples are not common in most natural sciences</td>
</tr>
</tbody>
</table>

ES Opportunity
ES Data Analysis Discussion

- Do mathematically elaborate (like Hall) Observational Methods exist in complex Socio-technical systems?
- Can we invent new ones?..involving unexpected “instrumental” variables
- David Broniatowski thesis completed in 2010
Agenda

- Motivation
- Literature Review
- Case Study: US FDA Medical Device Approval Panel Meetings – Circulatory Systems Devices Panel
- Methodology development – social networks from meeting transcripts
- Empirical Findings and Implications
- Preliminary Modeling Findings and Implications
- Contributions
Literature Review: Research on Group Decision-Making

Real-World

Semi-Automated:
- Network Text Analysis
  - (Roberts 1997), (Carley 1992), (Corman et al. 2002)

Automated:
- Opportunity for Contribution

Manual:
- Latent Coding & Cognitive Maps
  - (Axelrod 1976)

Honest Signals: (Pentland 2008)

Appreciative (e.g., Dogulas 1986, Cobb and Elder, 1983)
- Language expresses Group identity and influences perception of data
- Language drives decision outcomes via framing

- Voting outcome dictated by preferences, voting rules, and strategic interaction
- Treats all experts as homogeneous or assumes perfect communication

- Empirical but relies on undergraduates
- Group decisions as information processing mechanisms driven by information sharing
- Decisions strongly affected by status
How to construct a network?

- Each author pair links a certain number of times over all AT samples
  - Number of links is link frequency
- Histogram of link frequencies typically shows bimodality.
  - We would like to construct a network out of the frequent linkers
    - More likely than not to be linked
    - Establish Bonferroni threshold at 125 (assumes ~15 voting members).

Histogram of Link Frequencies
Contributions

1. Quantitative **automated method** for the extraction of social networks from meeting transcripts

2. Empirical Results showing that learning does occur on FDA panels:
   - The potential **technical** impacts of **ambiguity** on voting behavior and meeting social structure
   - The **cognitive** impacts of **medical specialty** on voting and linguistic behaviors
   - The **social** impacts of **speaking order** on voting and linguistic behaviors
   - The **social** role of the **committee chair** in mediation among panel members

3. Model Results showing that learning does occur on FDA panels:
   - The **technical** impacts of device **ambiguity, complexity** and **quality** on voting and linguistic behaviors, panel consensus, and panel correctness
   - The **cognitive** impacts of **learning/breadth** and **depth** of **expertise** on voting and linguistic behaviors, panel consensus, and panel correctness
   - The **social** impacts of **speaking order** on voting and linguistic behaviors
Questions

1. How do Banerjee, Hall, Milgram and Watts relate to ES?
   - Are they studying Complex, Socio-technical Systems? Why or why not?
   - What do they teach us about empirical observations and analysis?
   - Are any of the challenges or pitfalls that they face applicable to ES?

2. How is empirical work done in ES?
   - What empirical methods did guests and authors use?
   - What other methods of observation, experiment and data analysis have we seen applied to systems?
   - What methods might you use in your own research?

3. Cumulative Knowledge and the Cycle of Observation/Model/Theory/Practice
   - Could you write a PhD thesis that relies on just one of these elements? Or two?
Global Poverty from an ES Perspective

- Not an ES but perhaps a side effect or unintended consequence of an ES
- **Scope and Scale**
- ES = Global economic development system
- GEDS is huge (GOVTs, NGOs, Corp, indiv.)
- Boundary depends on problem - what is it?
- Geographic and political boundaries
- What are sub-issues to consider in these bounded areas? How well did Banerjee and Duflo do in selecting the sub-issues?
Global Poverty from an ES Perspective II

- Function of GEDS
- To increase global GDP?
- To increase national GDP?
- To increase personal GDP
- Would it help B/D to consider function
- Structure of GEDS (or National EDS)
- System of Systems?, layered, hierarchical, education, finance, health, Helpful to B/D?
- Temporality
- My view is that a rich temporality perspective would add the most to B/D analysis
Network Navigation: Milgram’s experiment

- The unpublished (but widely circulated) paper of Kochen and Pool using simple random graph models indicated the possibility of short paths through social networks. This instigated the famous social scientist Stanley Milgram to try an experiment.
- “route” letter to person XXX who is a stockbroker living in Sharon, MA who works in Boston.
- The letters can only be sent to someone who the recipient knows on a first name basis but in a way to get “closer” to person XXX. Participants also were asked to record and send along the routing information.
Social networks: Milgram’s experiment

Milgram’s experiment

- “route” letter to person XXX who is a stockbroker living in Sharon, MA who works in Boston.
- The letters can only be sent to someone who the recipient knows on a first name basis but in a way to get closer” to person XXX.

Some guesses that it would take hundreds of steps were refuted by results that “showed” it took (actually can take) much less
Six-degrees of separation

In the Nebraska Study the chains varied from two to 10 intermediate acquaintances with the median at five.

Milgram, *Psych Today* 2, 60 (1967)
Milgram’s experiment II

- “route” letter to person XXX who is a stockbroker living in Sharon, MA who works in Boston.
- The letters can only be sent to someone who the recipient knows on a first name basis but in a way to get closer” to person XXX.

- Some guesses that it would take hundreds of steps were refuted by results that “showed” it took (actually can take) much less.

→ A play was written and coined the phrase “six degrees of separation” as its Title and Milgram’s result became something “everyone knows”.

 Everyone is separated by only six removes from everyone else on the planet!

- But What did Milgram really show?
What did Milgram really show?

- Of 300 letters in original experiment, only 96 (random Nebraska) sampled tested the “everyone” part of what “everyone” knows.
- Only 18 of these were ever returned (the preceding graph contained very “non-random” Nebraska letters).
- Other trials that were random and tested the everyone basis had even smaller return rates than Milgram’s initial experiment.

Issues
- Is everyone really 6 or less steps from everyone?
Lecture 15 Outline

- Introductory Remarks
- Search and Navigation, search (briefly)
- Navigation
  - Milgram’s experiment and critiques
  - Small World and predecessor models
  - Kleinberg’s first model
  - The influence of structure and Kleinberg’s second model (and the Watts, Dodds and Newman model)
- Modeling Overview
  - Search/navigation as case study of model evolution
  - Modeling limitations and benefits
Poisson Random Graph

- Rapaport and later Erdos and Renyi and others such as Bollobas have studied a very simple model in some depth. This is the one where each node in a network is connected with probability $p$ to other nodes. Ensembles with variable numbers of links $<k>$ are studied and the degree distribution is
  \[ p_k \approx \frac{<k>^k e^{-<k>}}{k!} \]

- The path length can be formally shown to be
  \[ l \approx \frac{\ln n}{\ln <k>} \]
  and is thus consistent with a “Small World”

- Clustering is simply equal to the random probability of a link between 2 nodes and is
  \[ C = \frac{<k>}{n} \]
It is generally stated that this model is nice for intuition but describes no real networks. It also provides a benchmark.

If we look at a wide variety of “real world” graphs

What do we see?

Path Length, $l$, is generally small (small worlds) and often approximately equal to path length for a Poisson random network

Clustering is usually orders of magnitude higher than predicted by random networks for the large networks and is \( \sim \)constant with $n$
Small World Problem as seen by Watts

**Lattice**

$L(N) = N^{1/d}$

$C(N) \approx \text{const.}$

**Random graph**

$L(N) = \log N$

$C(N) \approx N^{-1}$
Small World Network Model (1D)

K is the number of nearest neighbors originally with links
(=3 below)
Small-world networks

- Large clustering coeff.
- Short typical path

Watts & Strogatz,

Small World Clustering Estimation

- Watts and Strogatz got results from simulation

- Later Work by Barrett and Weigt on their model derived a clustering coefficient of

\[ C = \frac{3(K - 1)}{2(2K - 1)} (1 - p)^3 \]

- An improved model by Newman and Watts and independently by Monasson gives for the clustering coefficient

\[ C = \frac{3(K - 1)}{2(2K - 1) + 4Kp(p + 2)} \]

- These estimates are sufficiently high for real networks
Simulation based by Watts and Strogatz showed that path lengths were small and scaled with \( \ln n \).

No exact solution (yet) but Barthelemy and Amaral proposed a scaling relation that was later derived by Newman and Watts. It shows that the transition to “Small World Path Length Dependence” occurs at smaller \( p \) as \( n \) increases. Indeed, **the number of shortcuts needed to give small world behavior is constant** (for given \( K \)) as \( n \) increases.
Ubiquity of small-world networks

\[ L(p, N) \sim N_* \frac{N}{N_*} \]


Image by MIT OpenCourseWare.
Small World Models

- Small world models thus
  - Show that it is relatively easy to have higher clustering and yet short paths. In large networks a few long paths is all that is needed - brain now understood this way as are some other large scale complex systems.

- However, the specific models have only marginal connection to any real systems as they are stylistic and notional.

- Small World Models have been relatively widely used as a “substrate” for studies of such as iterated games, epidemics. The rewiring approach has also proven useful even if the specific models are not real.
Potential short paths

- There are almost surely relatively short paths between any two individuals.

- The path length is apparently about that calculated for random networks:

\[ l \approx \frac{\ln n}{\ln <k>} \]

- For \( n \) representative of the whole world, this would give path lengths as large as 10-20. Even though 10 degrees of separation does not sound as impressive, it is still small.

- As a model, the Small World Model is obviously primitive as a “Systems Formation” Model. For this phenomena/purpose (explaining Milgram’s experiment), is this its most serious shortfall?
Search and Navigation

- **Search**
  - “To look over carefully in order to find something, to explore”, “to make an effort to find something” seek, hunt, quest.
  - Network literature: “to find the node containing information that is desired”

- **Navigate**
  - “To plan, record and control the position of..” “to follow a planned Course” or “to make one’s way”
  - Network literature: “to get from one to another specific node by a(he) short(est) path using only local information”
Kleinberg’s initial model

- Most important insight
  - Milgram’s experiment did not only show that short paths exist but more importantly **that people can** (at least sometimes and in some circumstances) **find and access them**.

- Model assumptions
  - Small World (with shortcuts added onto a **lattice** of connections) – **not randomly** but with a probability that depends on distance,
“Navigation in Small Worlds: It is easier to find short chains in some networks than others”
Kleinberg’s initial model

- Most important insight
  - Milgram’s experiment did not only show that short paths exist but more importantly **that people can** (at least sometimes and in some circumstances) **find them**.

- Model assumptions
  - Small World (with shortcuts added onto a **lattice** of connections) – **not randomly** but with a probability that depends on distance,
    \[ p_s \propto r^{-\alpha} \]
  - \( S_{mean} \geq cn^\beta \) steps to find
    \[ \beta = \frac{(2 - \alpha)}{3} \text{ for } \alpha < 2 \text{ and } \]
    \[ \beta = \frac{(\alpha - 2)}{(\alpha - 1)} \text{ for } \alpha > 2 \]
Exponent $\beta$ in lower bound on $T$

\[ \frac{2-\alpha}{3}, \quad \frac{\alpha-2}{\alpha-1} \]

Image by MIT OpenCourseWare.
The existence of short paths does not guarantee that they can be found with local information. It takes network structure of a certain kind ($\alpha = 2$) to be able to do this and to get Milgram’s result. The structure Kleinberg showed worked seems quite artificial but it was a start because it showed that networks can be designed that allow for rapid search with “greedy” algorithms based on local information (“gossip” algorithms). Based on this work, even if everyone was connected to everyone, it is surprising that anyone could find the short path.

Thus Milgram’s famous result is not explained by this model.
Next Generation structural models for navigation

- Kleinberg and independently Watts, Dodds and Newman later proposed a structure that allows such search and appears consistent with the structure of social networks.
- How would you try to route a letter to a stockbroker in Omaha?
- This structure is derived starting from clues from the “Reverse Small World Experiments” which indicate how people actually navigate social networks by looking for common “features” between their targets and their acquaintances.
- This structure introduces hierarchy into the social network and defines a “social distance”.
Basic model structure

Source: *Six Degrees: The Science of a Connected Age*, Duncan J. Watts, Fig. 4.6, 2003
Assumptions in 2ndG Navigation/Search Models I

1. Individuals have *links and identities*

2. Individuals partition the world (identities of others) into a layered hierarchy and *distance*, \( x_{ij} \) is assumed to be the height of the lowest common parent. The branching ratio, \( b \), and levels, \( l \) define this abstraction.
Assumptions in 2ndG Navigation/Search Models II

1. Individuals have *links and identities*

2. Individuals partition the world into a layered hierarchy and *distance*, $d_{ij}$ is assumed to be the height of the lowest common parent. The branching ratio, $b$, and levels, $l$ define this abstraction.

3. Group membership signifies not only identity but also is a primary basis for determining social interaction:

4. Individuals hierarchically partition the world in more than one way and the model first assumes these distinctions are independent (Kleinberg shows this assumption can be relaxed with qualitatively similar results).

5. Individuals construct a measure of “social distance” which is the minimum over all dimensions between the nodes:

$$p_x = c \exp[-\alpha x]$$
H = 1

Dimension 1
(e.g., Geography)

H = 2

Dimension 2
(e.g., Occupation)

Source: Six Degrees: The Science of a Connected Age, Duncan J. Watts, Fig. 5.7, 2003
Assumptions in 2ndG Navigation/Search Models III

1. Individuals have **links and identities**

2. Individuals partition the world into a layered hierarchy and **distance**, $x_{ij}$, is assumed to be the height of the lowest common parent. The branching ratio, $b$, and levels, $l$ define this abstraction.

3. Group membership signifies not only identity but also is a primary basis for determining social interaction.

4. Individuals hierarchically partition the world in more than one way and the model first assumes these distinctions are independent.

5. Individuals construct a measure of “social distance” which is the minimum over all dimensions between the nodes:

$$p_x = c \exp[-\alpha x]$$

6. Individuals forward messages based only on knowledge of their nearest neighbors and **their identities**. Forwarding the message to someone closer to the target is the “greedy” or “gossip” algorithm used.
Results I

- Successful search assumes a decent probability (.05) of finishing the chain even though the probability of terminating the search at each step is fairly high (0.25 or higher).

- Key result is that searchable networks occupy a broad range of parameter space ($\alpha$, $H$) with almost all searchable networks having $\alpha > 0$ and $H > 1$. 
Search success for different size networks with alpha and H

Image by MIT OpenCourseWare.
Results II

- Successful search assumes a decent probability (0.05) of finishing the chain even though the probability of terminating the search at each step is fairly high (0.25 or higher).

- Key result is that searchable networks occupy a broad range of parameter space ($\alpha$, $H$) with almost all searchable networks having $\alpha > 0$ and $H > 1$.

- Increasing group dimension beyond $H = 1$ yields a dramatic increase in search success (reduction in delivery time) but “the improvement is lost as $H$ increases further.”
Probability of successful search
Results III

- Successful search assumes a decent probability (.05) of finishing the chain even though the probability of terminating the search at each step is fairly high (0.25 or higher).

- Key result is that searchable networks occupy a broad range of parameter space (\( \alpha, H \)) with almost all searchable networks having \( \alpha > 0 \) and \( H > 1 \).

- Increasing group dimension beyond \( H = 1 \) yields a dramatic increase in search success (= reduction in delivery time) but the improvement is lost as \( H \) increases further.

- For plausible values of all parameters, agreement with Milgram results are found.
Distribution predicted vs. Milgram distributions

Image by MIT OpenCourseWare.
The Iterative Learning Process

Objectively obtained quantitative data (facts, phenomena)

deduction \rightarrow \text{induction}

\rightarrow \text{hypothesis (model, theory that can be disproved)}

Models are “hardened” only by intensive simultaneous observational studies of relevant \textit{reality}. The result can be

The rapid facilitation of a transition to engineering (vs. craft approaches) for the design of complex social/technological systems

The emergence of a cumulative science in this area.
The Iterative Learning Process

Objectively obtained quantitative data (facts, phenomena)

\[ \text{dEDUCTION} \rightarrow \text{INDUCTION} \rightarrow \text{dEDUCTION} \rightarrow \text{INDUCTION} \]

hypothesis (model, theory that can be disproved)

Models are “hardened” only by intensive simultaneous observational studies of relevant reality.

What social distance (communication) exists in real social networks?
Random network models indicate relatively short paths might exist.
Milgram does an experiment and short paths (small worlds) exist.
Random networks do not describe clustering and short paths
Small world model is consistent and ubiquitous- Milgram experiment is revisited
Kleinberg points out navigation issue and introduces a model which treats it but does not agree with Milgram. A 2nd generation navigation model introduces structure into the social network and agrees with Milgram result.
Data Analysis Discussion

- Mapping the Millennium- Experiment or Observational Study?
- Report from the Front- experiment or ?
- What is learned from these cases (and some others like the earth’s core, galaxy structure etc.)
- The models and theories deployed in data analysis are as elaborate (sophisticated) as the models for understanding the system
Discussion

- Assignment 1- examples of observational methods
- How would you differentiate between an experiment and an observational study?
  - Experiment = system (individuals treated, nature of treatment, measures of outcomes, etc.) under control of the investigator
- Experiment or observational study?
  - Duflo et al
  - Hall
  - Travers and Milgram
  - Report from the Front
A Research Process 3

1. Development of conceptual understanding (qualitative framework)
2. Development of quantitative predictive model
3. Observe (system)
4. Analyze observations
5. Generalize or simplify/complicate model

Research styles (1, 2, 3, 4, 5 repeat; 1, 3, 5 repeat; 1/3, 2/4, 5/1; 3, 4, 1, 2; etc.)
ESD.83 Doctoral Seminar in Engineering Systems
Fall 2011

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.