Organic Networks: Wearable Computers For Human Organization

Alex Pentland

Room E15-387, MIT Media Laboratory, 20 Ames St, Cambridge MA 02139

http://hd.media.mit.edu, pentland@media.mit.edu

Abstract

We propose that active analysis of face-to-face interactions within the workplace can radically improve the functioning of the organization. There are several different types of information inherent in such interactions: interaction features, participants, context, and content. By aggregating this information, high-potential collaborations and expertise within the organization can be identified, and information efficiently distributed. Examples of using wearable machine perception to characterize face-to-face interactions and using the results to initiate productive connections are shown, and privacy concerns are addressed.

1. Introduction

A series of studies on office interactions discovered that 35-80% of work time is spent in spoken conversation, 14-93% of work time is spent in opportunistic communication, and 7-82% of work time is spent in meetings [1]. Senior managers represent the high end of these scales. Clearly, face-to-face interactions within the worksplace is highly important, and critical pieces of information are often transmitted by word of mouth in a serendipitous fashion. The money and time spent on business travel and conferences further underscores the value of face-to-face and serendipitous interactions.

Given the importance of such communications, it is notable that the majority of working professionals already carry a microphone and speaker continually in the form of a cellular phone. Many also carry PDAs with computational horsepower similar to those found in desktop computers only a few years ago. This emerging foundation of mobile IP communications and wearable processing power within the workplace will enable an exciting suite of new business applications.

We therefore propose that if an organization were able to use wearable technology to characterize the face-to-face interactions of its employees, it would have an extraordinary resource for collaboration, team formation, knowledge management, and social network analysis (see Basu et al [2], Choudhury and Pentland [3], Eagle and Pentland [4], http://hd.media.mit.edu).

To explore this vision I will first discuss the information that can be obtained from many streams of sensor data using a wearable computation platform. I will then describe how this information can be combined with knowledge about human networks and common-sense knowledge about topics of conversation to characterize the semantics and the function of the interactions.

2. Applications

Synergistic collaborations, real-time expertise, and redundant work can be identified by clustering people based on profiles generated from an aggregate of conversation, email, location, and web data. Additionally, by leveraging recent advances in machine learning, robust computational models can be built to simulate the effects of organizational disruptions in the existing social networks, such as relocating a group to a different location or merging two departments. Indeed, such a data-driven model offers the potential to transcend the traditional org-chart, perhaps by drawing parallels to ad-hoc network optimization. Forming groups based on inherent communication behavior rather than rigid hierarchy or formal education may also yield significant improvements to the organizations performance. I will then describe a method of mathematically describing these interactions using probabilistic graphical models, potentially allowing far more effective management of the organization, and report on some of the applications our technology enables. Finally, I will discuss the privacy implications of such systems.

3. Conclusion

We have proposed that active analysis of interactions within the workplace can radically improve the functioning of the organization. There are several different types of information in face-to-face interactions that are measurable by inconspicuous wearable devices: prosidic and body language features, participants identity, context, and conversational content. By aggregating this information, interpreting it in terms of work tasks, and modeling the dynamics of the interactions, we hope to be better able to understand and manage complex organizations.

4. References

1. T. Allen, Architecture and Communication Among Product Development Engineers, 1997, MIT Press, Cambridge, MA, pp. 1-35
2. S. Basu, T. Choudhury, B. Clarkson, and A. Pentland. (2001) *Towards Measuring Human Interactions in Conversational Settings*. IEEE Int'l Workshop on Cues in Communication (CUES 2001) at CVPR 2001. Kauai, Hawaii.
3. T. Choudhury, and A. Pentland, (2003) *The Sociometer: Measuring Human Networks*, 23d Int'l Social Network Conf ., Cancun, MX Feb. 12-16, 2003
4. N. Eagle, and A. Pentland, (2003)., *Collaborative Conversations*, 23d Int'l Social Network Conf., Cancun, MX Feb. 12-16, 2003

Organic Networks: Building Computation around Human Networks

Prof. Alex (Sandy) Pentland
The Media Laboratory
Massachusetts Institute of Technology

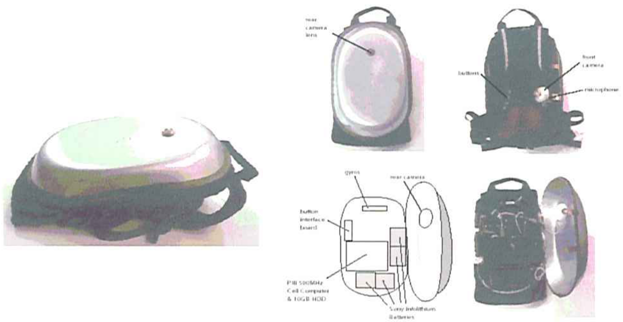
Personal, Context Aware Computation

WEARABLES TOKYO
Wearable Computers
Symposium & Fashion Show

Personal Video Mining

The "I Sensed" Data Set
Data Collection Wearable

Clarkson, Pentland

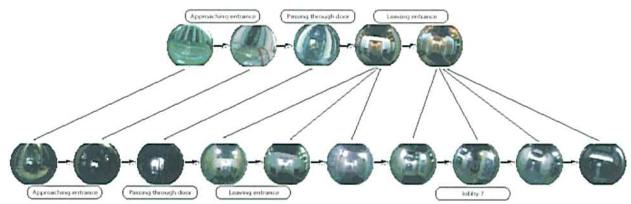


demo available

The Similarity Measure

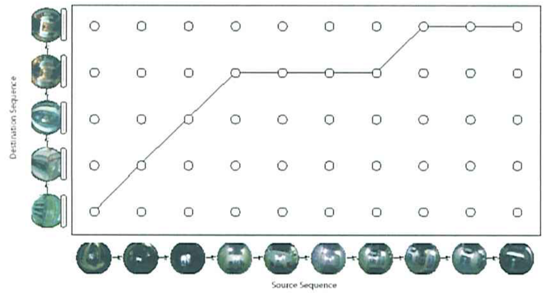
Sequence Similarity

Match one sequence to the other and accumulate frame-by-frame similarities.



The Similarity Measure
Alignment Path

The Viterbi Algorithm produces the best possible alignment.

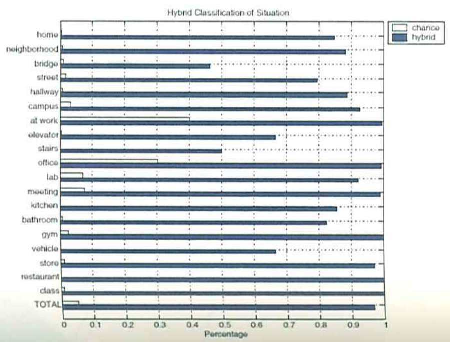


Situation Classification
19 Situations

home	apartment
neighborhood	Beacon St., Mass. Ave. (Boston-side)
bridge	Harvard Bridge, Longfellow Bridge
street	Kendall Sq., Boston Downtown, Main St., Memorial Dr., and more
hallway	Infinite Corridor and more
campus	inside & outside of bldg. 56, 66, 7, 10, and more
at work	anything in the Media Lab
elevator	any elevator
stairs	any stairs
office	my office at lab
lab	the area outside of my office
meeting	any meeting
kitchen	kitchen (at home and lab)
bathroom	any bathroom
gym	Dupont
vehicle	taxi, subway, bus
store	any store
restaurant	any restaurant
class	any class

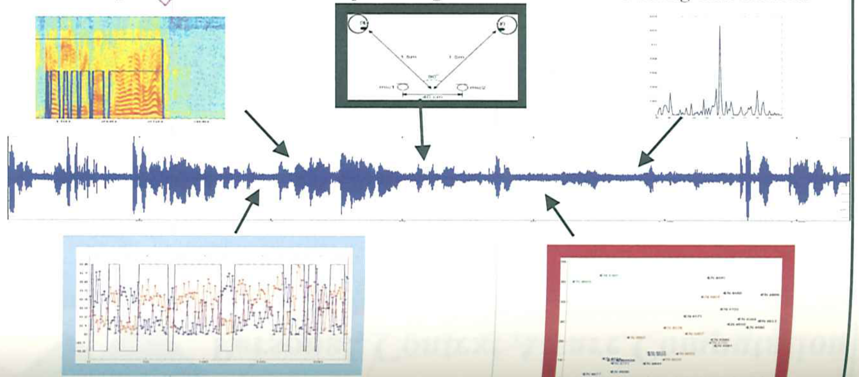
* Every 5 minute interval over 20 days was labeled with its situation(s).

Situation Classification
A Hybrid Classifier

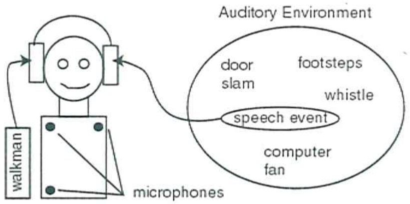


Total:
rank-1 = 97%

Audio Mining



Application: Smart Headphones



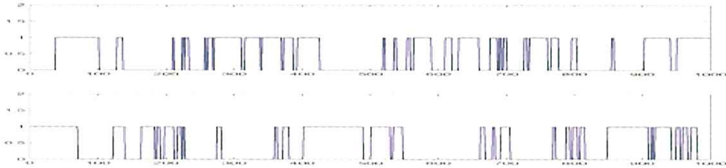
- **Interface options:**
  - Pass through speech, mix volume
  - Browsable record of speech events

Basu, Pentland. "Smart Headphones." CHI '2001 Short Paper

Mutal Information: Extremely reliable conversation detection

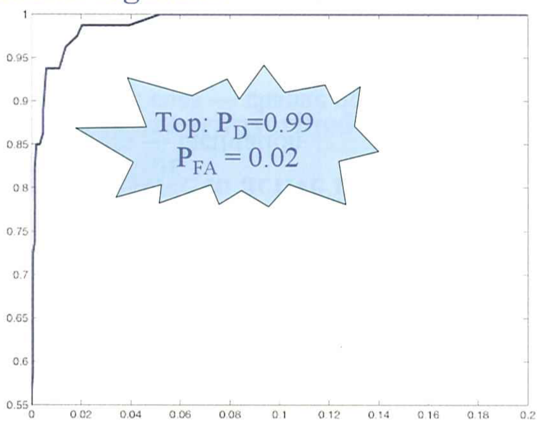
Voicing segments look like a pseudorandom bit sequence

- The conversational partner is a **noisy** complement
- use mutual information to detect conversations



Spectacular Accuracy at Detection

- **One-minute segments**



Socially Aware Telephony

Sawney, Clarkson, and Pentland, IEEE Wearables

Where Next?

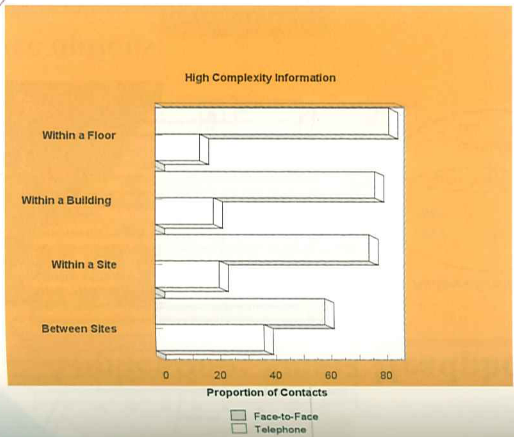
- Move from passive to active interfaces:
  - unequal access --- facilitation
  - proactive social cues --- change dominance
  - behavior modification --- shape interaction
  - unconscious learning --- subliminal, implicit

Memory Glasses

DeVaul, Pentland

- Subliminal cues work when you are primed!
- Cognitive aid without cognitive load
- Applications:
  - Reminders
  - Training

Face-to-face interactions as primary medium



Allen, T., *Architecture and Communication Among Product Development Engineers*. 1997, Sloan School of Magement, MIT: Cambridge.

Design computational network to support human networks

- Leveraging The New Corporate Uniform
  - Today's Personal Digital Assistant
  - Cellular Phone Headset Mics
- Computational Networking
  - Figuring Out Who You Are Talking To
  - Categorizing the Relationship
  - Inferring Aspects of Your Situation

(Bluetooth) Microphone / Headset

Sharp Zaurus

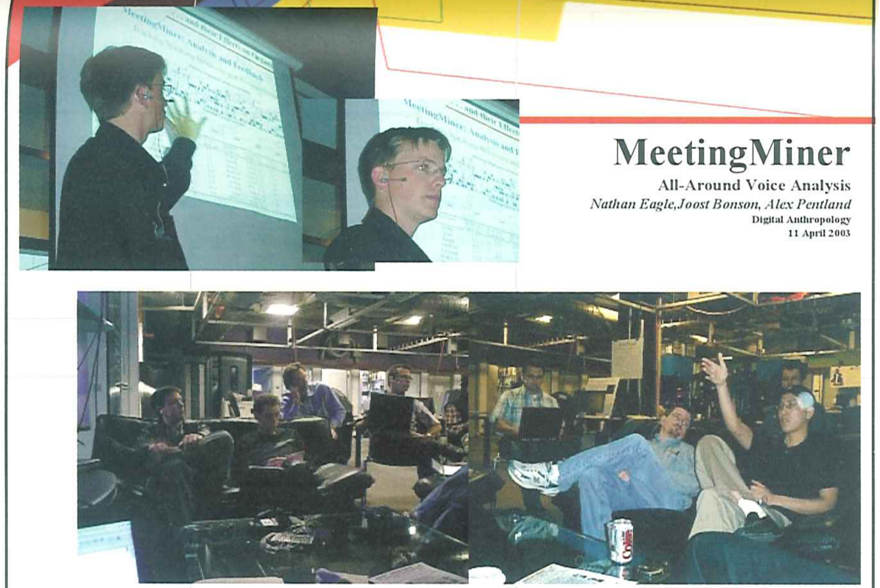
MeetingMiner

All-Around Voice Analysis

Nathan Eagle, Joost Bonson, Alex Pentland

Digital Anthropology

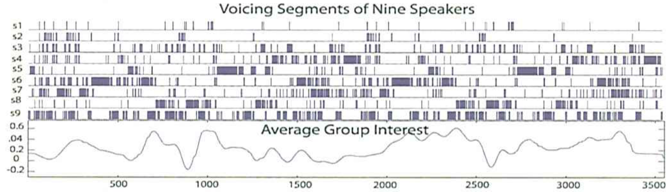
11 April 2003



JPB 2003

. . . and (hopefully) improve the organization

- **Example: MeetingMiner**
  - Tracking Speaking Behaviors and Audience Interest



Participant	Speaking time (%)	Avg (sec) Comment	Nearest Neighbor	Transition (Name, %)	Avg Interest	Group Interest
Ivan	1.5	4.1	Nathan	Nathan-27	.21	.44
Jon	2.2	2.2	Sandy	Sandy-47	.13	.36
Joost	9.9	3.5	Sandy	Jordan-22	.20	.22
Jordan	11.4	9.6	Mike S	Mike O-23	.05	.30
Leonelle	12.8	8.8	Mike S	Sandy-37	.18	.33
Mike O.	16.9	6.6	Jordan	Mike S-28	.09	.21
Mike S.	10.1	6.6	Jordan	Sandy-30	.19	.24
Nathan	10.8	10.9	Ivan	Sandy-26	.40	.32
Sandy	24.4	6.9	Mike S	Mike O-22	.17	.25

Listen to my Words

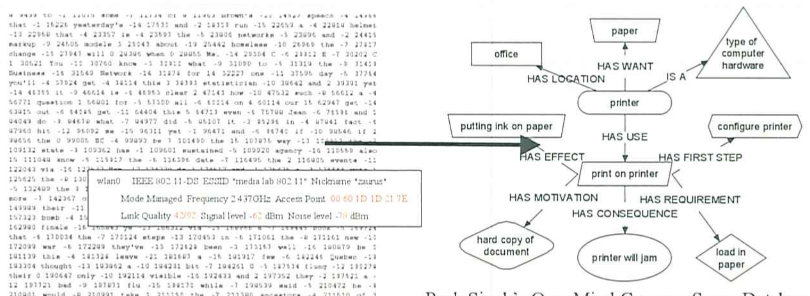
Eagle, Pentland

- **Omni-present "familiar" listens in a user's conversations**
- **Conversations are transcribed by a voice recognition engine**
- **Common sense engine to figure out a user's current context given new conversation data**

Inference of Conversational Situation

[Eagle, Singh, Pentland '03]

- **Discrete to Continuous Knowledge Representations**
  - Regularization of the noisy transcripts with semantic filtering
  - Build out the conditional probability distributions to reflect recent user behavior



Push Singh's OpenMind CommonSense Database

Some Preliminary Results

[Eagle, Singh, Pentland '03]

Chatting about what to order in the cafeteria:

Transcription:

Store going to stop and listen to type of its cellular and fries he backed a bill in the one everyone get a guess but that some of the past like a salad bar and some offense militias cambers the site fast food them and the styrofoam large chicken nuggets son is a pretty pleased even guess I as long as can't you don't have to wait too long its complicity sunrise against NAFTA pact if for lunch

Selected Keywords:

wait type store stop salad past lunch long long listen large fry food fast chicken cellular bill big bar back

Rank	Location Context
27	eat in fast food restaurant
21	eat in restaurant
18	wait on table
16	you would go to restaurant because you
16	wait table
16	go to restaurant
15	know how much you owe restaurant
12	store food for people to purchase
11	sitting down while place order at bar
11	cook food

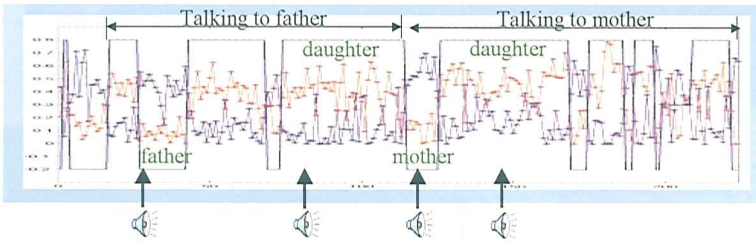
Conversations are More than Words

Basu, Pentland

- **Speech recognition is important but…**
- **Knowing "what's going on" is important:**
  - Is there a conversation?
  - **Who** is speaking **where** and **when**
  - **How** are they speaking (speaking style)
  - Recognizing conversational patterns
- **Conversations => scenes**
  - Who's holding the floor?
  - Characterizing conversations

Characterizing Audio Scenes

- **We speak differently to different people:**



	Father	Mother
Pitch (Hz)	444+90	472+98
Energy	.36+.30	.78+.67
Av. Gap (s)	1.03	0.66

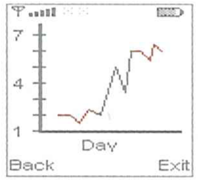
Daughter holding the floor

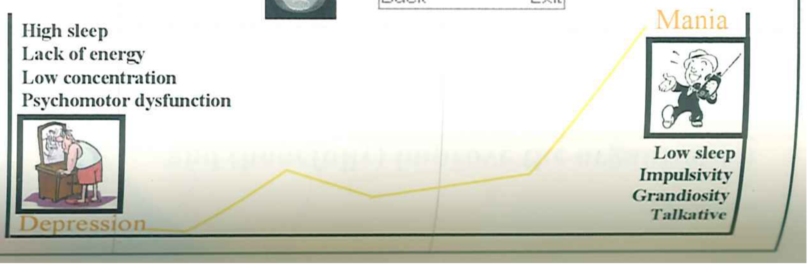
	Father	Mother
Pitch (Hz)	372+51	555+99
Energy	.18+.10	.85+.78
Av. Gap (s)	3.13	1.85

Daughter not holding the floor

Application: A self-management tool

Kumar, Basu, Pentland





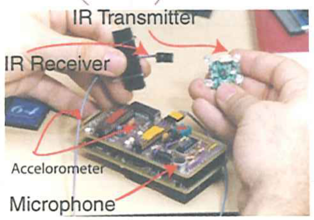
Where Next?

- **Active, purposeful management for group goals**
  - requires joint model of users and interaction
  - ability to estimate model parameters
  - ability to influence model evolution

How do we measure interactions ?

Sensor based approach

The Sociometer

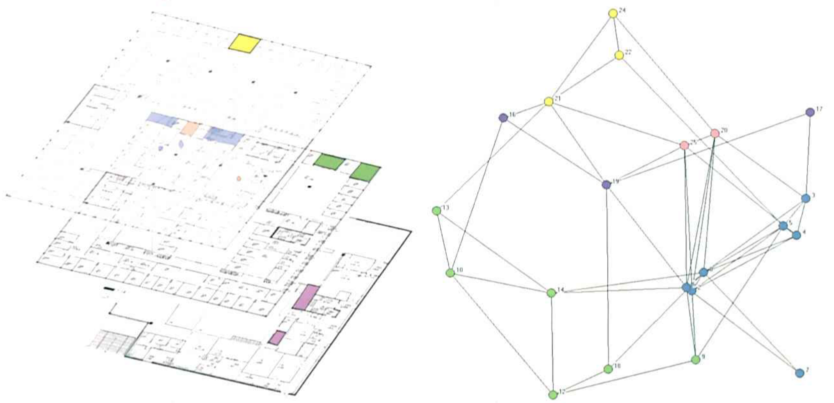


Not cumbersome
Effortless

wearability
Shape, size, attachment, weight, movement, aesthetics

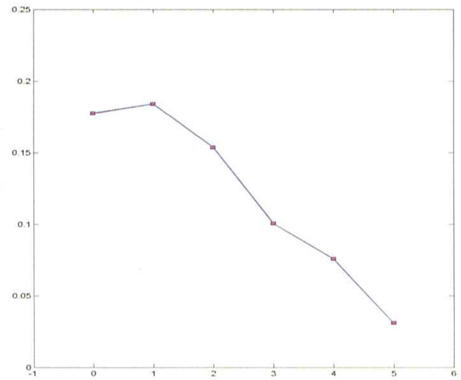
Brian Clarkson, Rich DeVaul, Vadim Gerasimov, Josh Weaver

Social Network



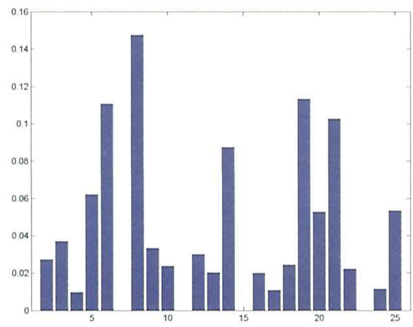
Based on multi-dimensional scaling of geodesic distances

Effects of Distance



X-axis label Distance

Betweenness Centrality of the Participants



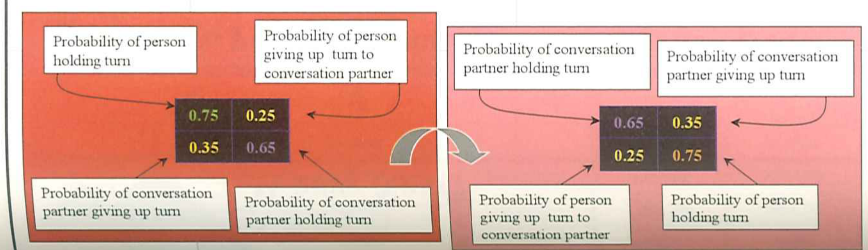
Betweenness centrality of individuals in the interaction network

Turn-taking Matrix

Person A converses with a given conversation partner.

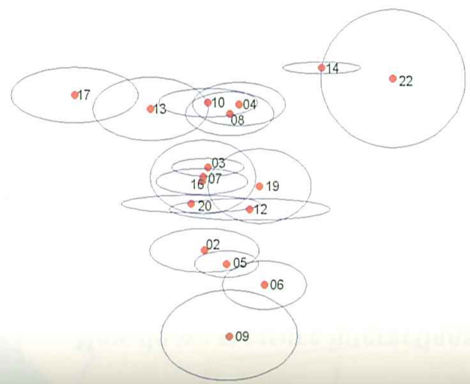
Turn-taking matrix for A

And turn-taking matrix for the partner

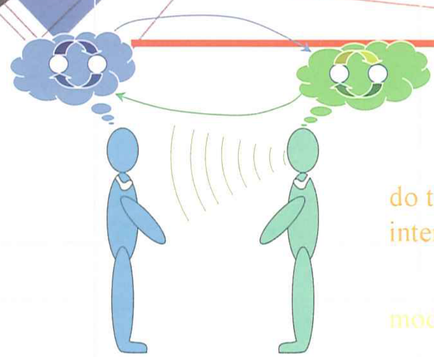


Average Turn-taking Style

For each individual we can estimate: "average self" and "average partner"



Do people affect each other's turn-taking?



do they affect each other's interaction

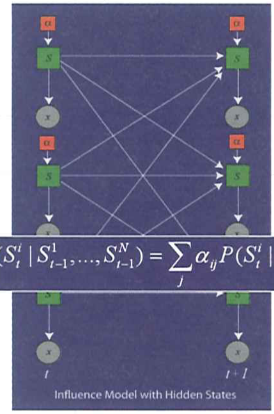
how do we model the effect?

a specific pair A and B

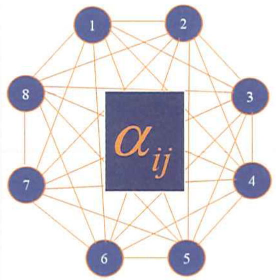
A's "average-self" | A's "average partner"

B's "average-self" | B's "average partner"

Influence Parameters:



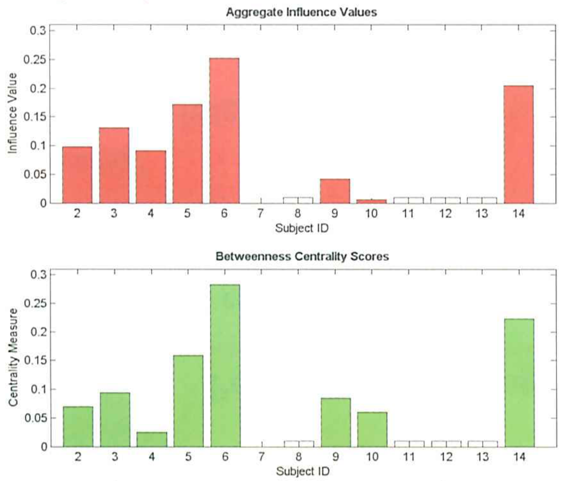
$$P(S_t^i \mid S_{t-1}^1, \ldots, S_{t-1}^N) = \sum_j \alpha_{ij} P(S_t^i \mid S_{t-1}^j)$$



$\{\alpha_{ij}\}$: Amount of influence that person i has on person j

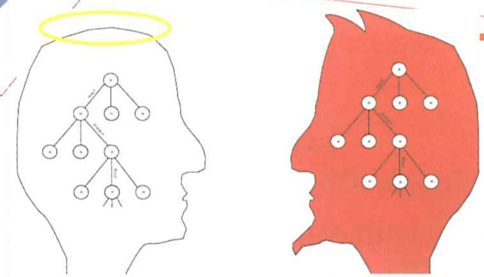
$P(S_t^i \mid S_{t-1}^j)$: How person i is influenced by person j

Correlation Influence Values with Centrality Scores



Correlation: 0.90

Game Theory



This is a formal framework to model interactions of strategy. Some examples:

- Chess
- Producer vs. Consumer pricing scenarios
- Negotiations

Generally, game theory allows us to model multi-agent interactions when each agent is:

- trying to achieve some goal (often different amongst participants)
- modeling the intentions of the other participants/opponents

Policy by Technology?

Can we design communications networks to achieve social goals?

- Transparency?
- Access?
- Self-knowledge?

Take Home Messages

- From computation for individual to computation for organization
- From passive to active interface
- From network of information channels to network of influences