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 workplace 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 for 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: prosodic 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


Organic Networks: Building Computation around Human Networks

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

Personal Video Mining

The "I Sensed" Data Set
Clarkson, Pentland

Data Collection Wearable

The Similarity Measure

Sequence Similarity

Match one sequence to the other and accumulate frame-by-frame similarities.
The Viterbi Algorithm produces the best possible alignment.
Application: Smart Headphones

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

Mutual 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

Sawney, Clarkson, and Pentland, IEEE Wearables
Memory Glasses

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

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

Where Next?

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

Face-to-face interactions as primary medium
MeetingMiner
All-Around Voice Analysis
Nathan Eagle, Janet Boman, Alex Pentland
Digital, Cambridge
11 April 2000

...and (hopefully) improve the organization

Example: MeetingMiner
- Tracking Speaking Behaviors and Audience Interest

![Voicing Segments of Nine Speakers]

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

![Inference of Conversational Situation Diagram]
Some Preliminary Results

Chatting about what to order in the cafeteria:

Transcription:
Store going to stop and listen to type of its Cellular and files 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 complexity 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

<table>
<thead>
<tr>
<th>Rank</th>
<th>Location</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>eat in fast food restaurant</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>eat in restaurant</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>wait on table</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>you would go to restaurant because you</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>wait table</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>go to restaurant</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>know how much you owe restaurant</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>store food for people to purchase</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>sitting down while place order at bar</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>cook food</td>
<td></td>
</tr>
</tbody>
</table>

Conversations are More than Words

- 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:

<table>
<thead>
<tr>
<th></th>
<th>Father</th>
<th>Mother</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch (Hz)</td>
<td>444±90</td>
<td>472±98</td>
</tr>
<tr>
<td>Energy</td>
<td>.36±.30</td>
<td>.78±.67</td>
</tr>
<tr>
<td>Av. Gap (s)</td>
<td>1.03</td>
<td>0.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Father</th>
<th>Mother</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch (Hz)</td>
<td>372±51</td>
<td>555±99</td>
</tr>
<tr>
<td>Energy</td>
<td>.18±.10</td>
<td>.85±.78</td>
</tr>
<tr>
<td>Av. Gap (s)</td>
<td>3.13</td>
<td>1.85</td>
</tr>
</tbody>
</table>

Application: A self-management tool

Kumar, Basu, Pentland

High sleep
Lack of energy
Low concentration
Psychomotor dysfunction

Mania
Low sleep
Impulsivity
Grundiosity
Talkative

Depression
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

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

Social Network

Based on multi-dimensional scaling of geodesic distances

Brian Clarkson, Rich DiPietro, Vadim Goranov, Josh Weaver
**Effects of Distance**

![Graph showing the effects of distance on a variable.]

**Betweenness Centrality of the Participants**

![Bar chart showing 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

- Probability of person holding turn: 0.75, 0.35
- Probability of person giving up turn to conversation partner: 0.25, 0.65
- Probability of conversation partner holding turn: 0.65, 0.72
- Probability of conversation partner giving up turn: 0.25, 0.75

**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's "average-self"  A's "average partner"
B's "average-self"  B's "average partner"

**Influence Parameters:**

$$P(S_i | S_{i-1}^*, S_{i-2}^*, \ldots, S_{i-n}^*) = \sum_{j} \alpha_{ij} P(S_j | S_{j-1}^*)$$

- $\{\alpha_{ij}\}$: Amount of influence that person i has on person j
- $P(S_i | S_{i-1}^*)$: How person i is influenced by person j

**Correlation Influence Values with Centrality Scores**

- Aggregate Influence Values
  - Correlation: 0.90

- Betweenness Centrality Scores

**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