Social Physics and the Data Driven Society

Prof. Alex Pentland, MIT

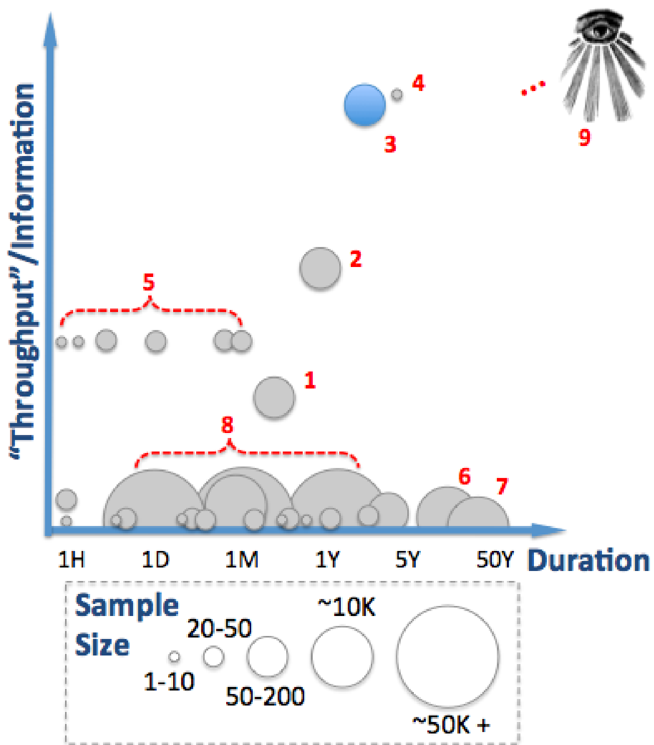
Connection Science and Engineering

WEF Big Data, Hyperconnected World

Outline

1. Big Data and computational social science
2. Distributed Intelligence
3. Network Intelligence
4. Big Data breaks science

1. Understanding Ourselves: The Big Data Revolution



Human Dynamics Observatories:

(1) MIT Reality Mining Study
(2) MIT Social Evolution,
(3) MIT Friends and Family (Current),
(4) MIT lifelog pioneers; MyLifeBits,
(5) Sociometric Badge studies,
(6) Midwest Field Station ,
(7) Framingham Heart Study,
(8) Large Call Record Datasets ,
(9) "Omniscient"/All-Seeing View

Ahrony, Pentland

Background:
Humans Have Two Types of Thought

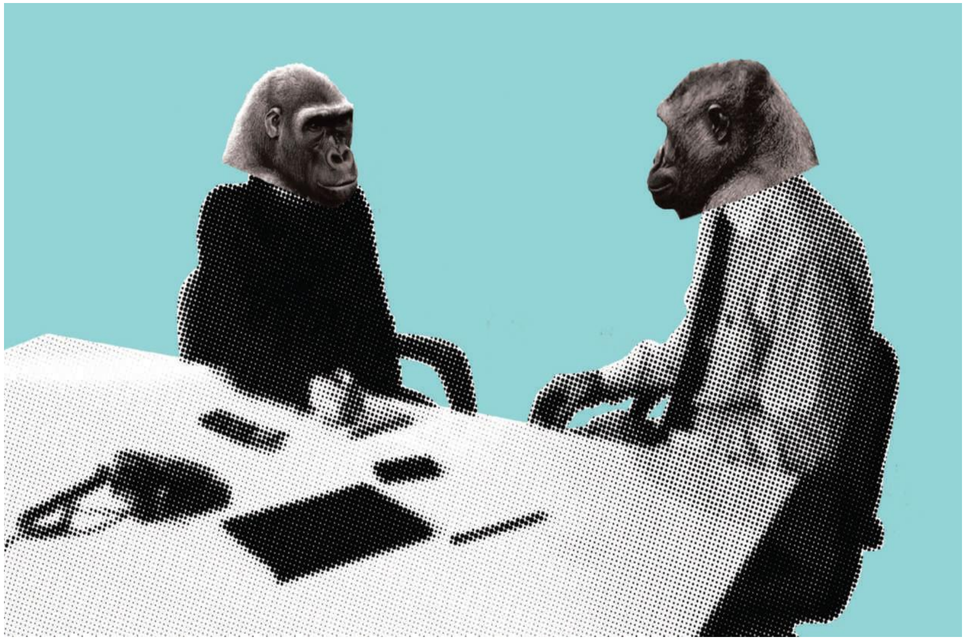
Nobel Prize winner Kahneman,

father of behavioral economics

Social Physics

<table>
<tr><th></th><th>Habitual
(System 1)</th><th>Attentive
(system 2)</th></tr>
<tr><td>Process</td><td>Fast
Parallel
Automatic
Associative</td><td>Slow
Serial
Controlled
Rule-based</td></tr>
<tr><td>Content</td><td colspan="2">Conceptual Representations
Past, Present, Future
Can be evoked by language</td></tr>
</table>

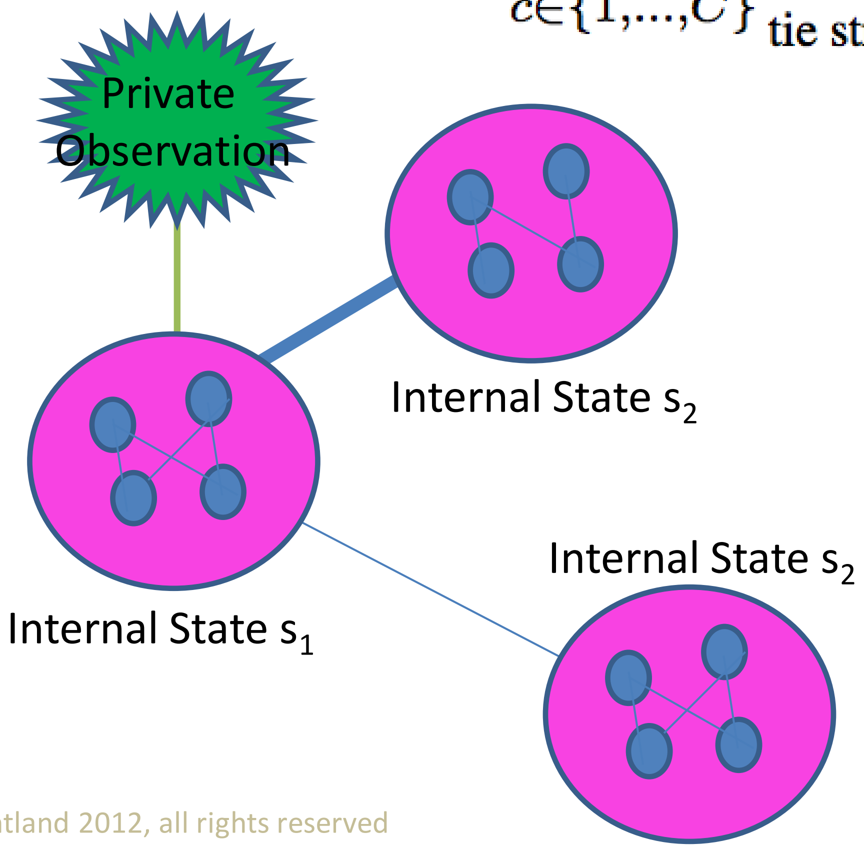
People Mostly Learn by Examples, not Arguments or Reasoning



90% - 10% balance
Rendell et al, Social Learning, Science 4/10

Influence Model & Idea Flow

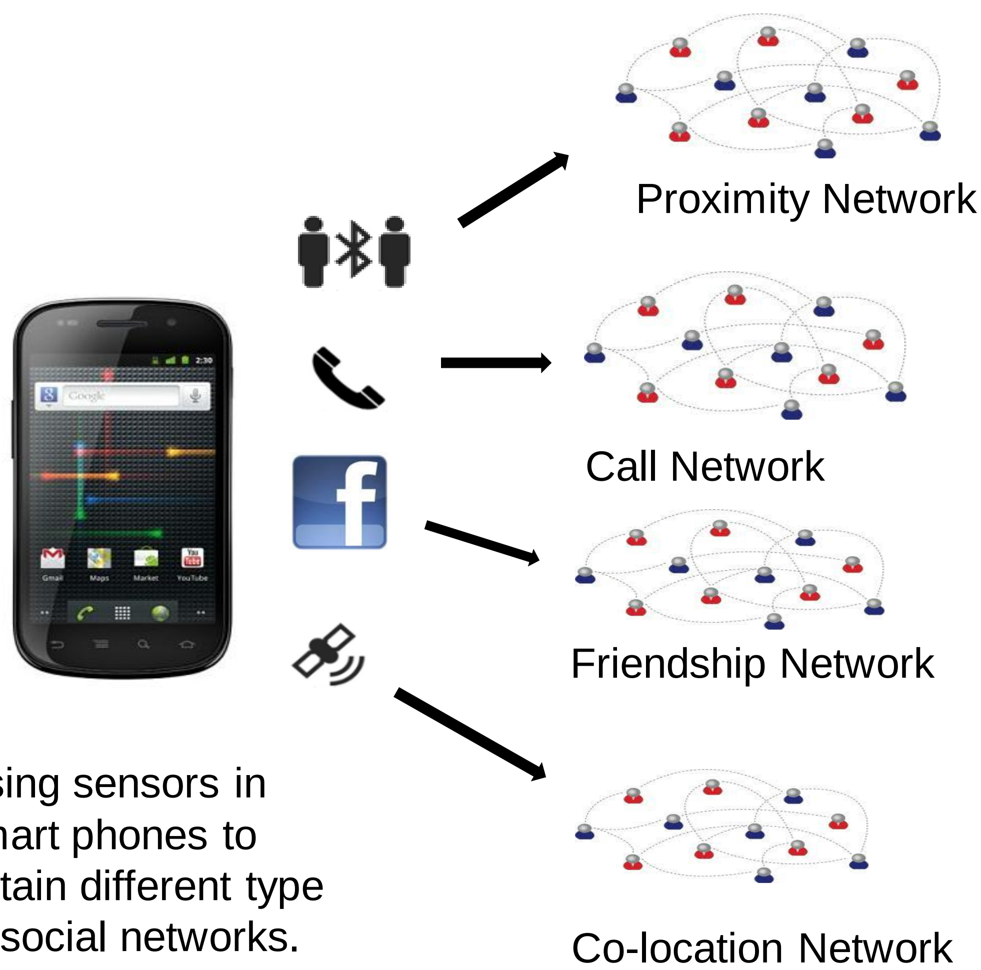
$$\mathrm{Prob}(h_t^{(c')}|h_{t-1}^{(1)},\ldots,h_{t-1}^{(C)}) = \sum_{c\in\{1,\ldots,C\}} \underbrace{\mathbf{R}_{c',c}}_{\text{tie strength}} \times \underbrace{\mathrm{Prob}(h_t^{(c')}|h_{t-1}^{(c)})}_{\text{cond. probability}}$$



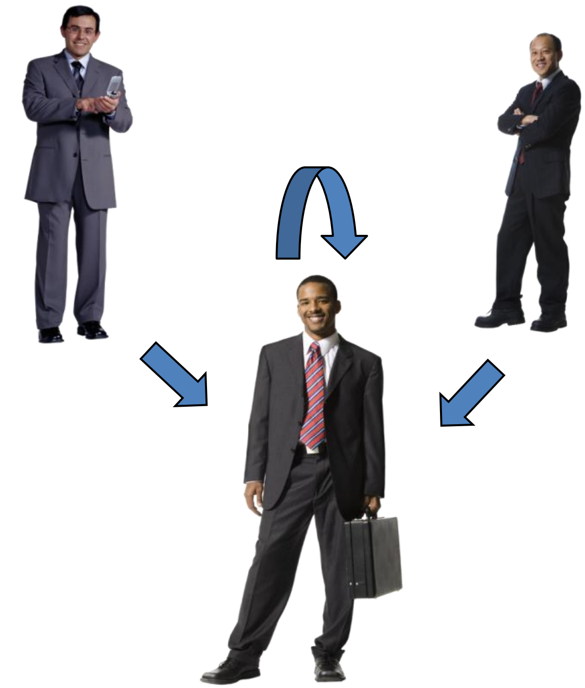
	s_1	s_2
s_1	0.8	0.2
s_2	0.2	0.8

Social Exposure Predicts Behavior

65 young families, 12 months data



Using sensors in smart phones to obtain different type of social networks.

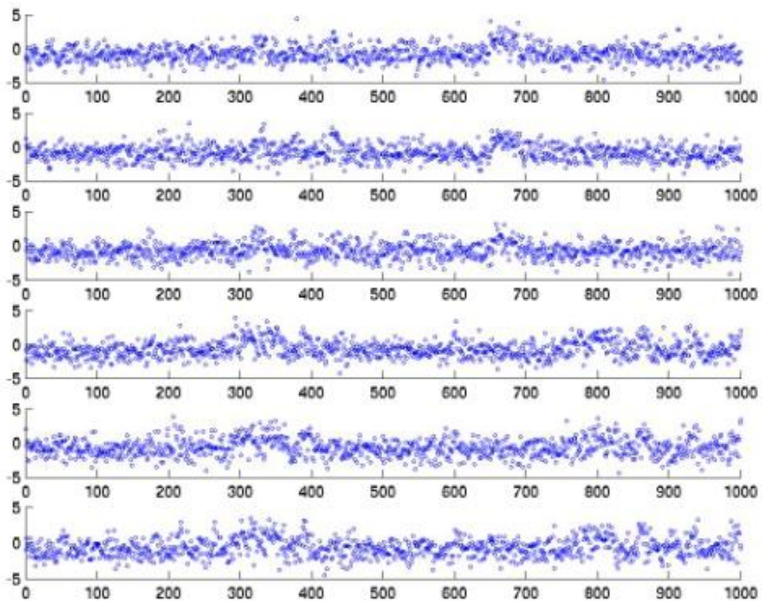


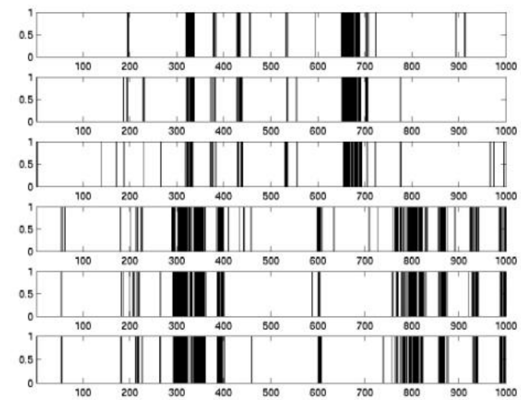
45% accuracy predicting app downloads. Gompertz function describes influence

Pan, Aharony, Pentland

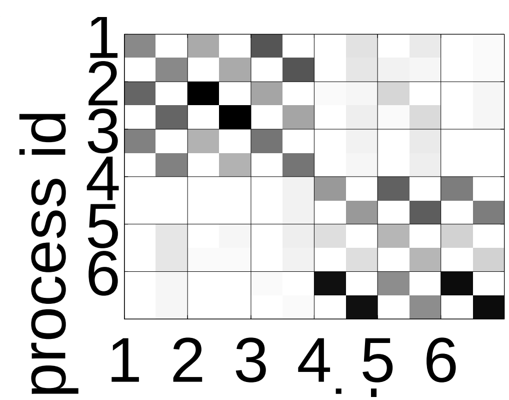
Inverse Problem: Discovery of Influence, Node State

Raw Observations From Nodes





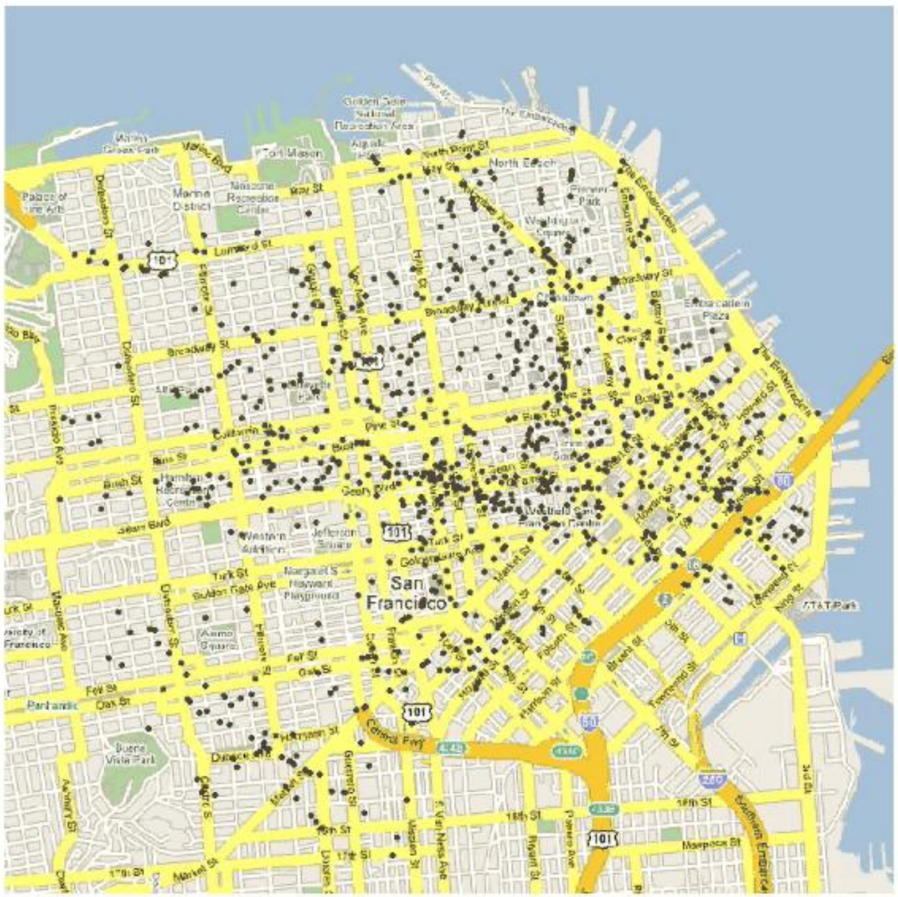
clean state estimate for nodes



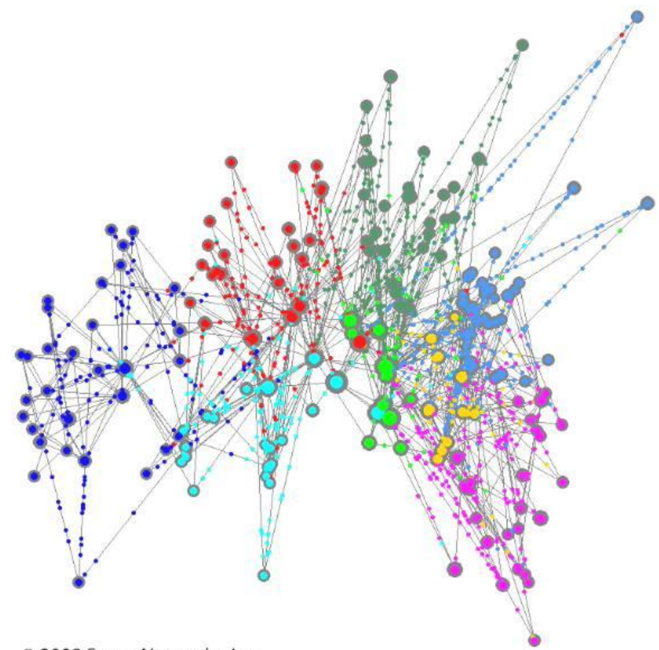
influence structure connecting nodes

$$\mathrm{Prob}(h_t^{(c')}|h_{t-1}^{(1)},\dots,h_{t-1}^{(C)}) = \sum_{c\in\{1,\dots,C\}} \underbrace{\mathbf{R}_{c',c}}_{\text{tie strength}} \times \underbrace{\mathrm{Prob}(h_t^{(c')}|h_{t-1}^{(c)})}_{\text{cond. probability}}$$

Understanding Ourselves: Behavioral Demographics



© 2008 Sense Networks, Inc.



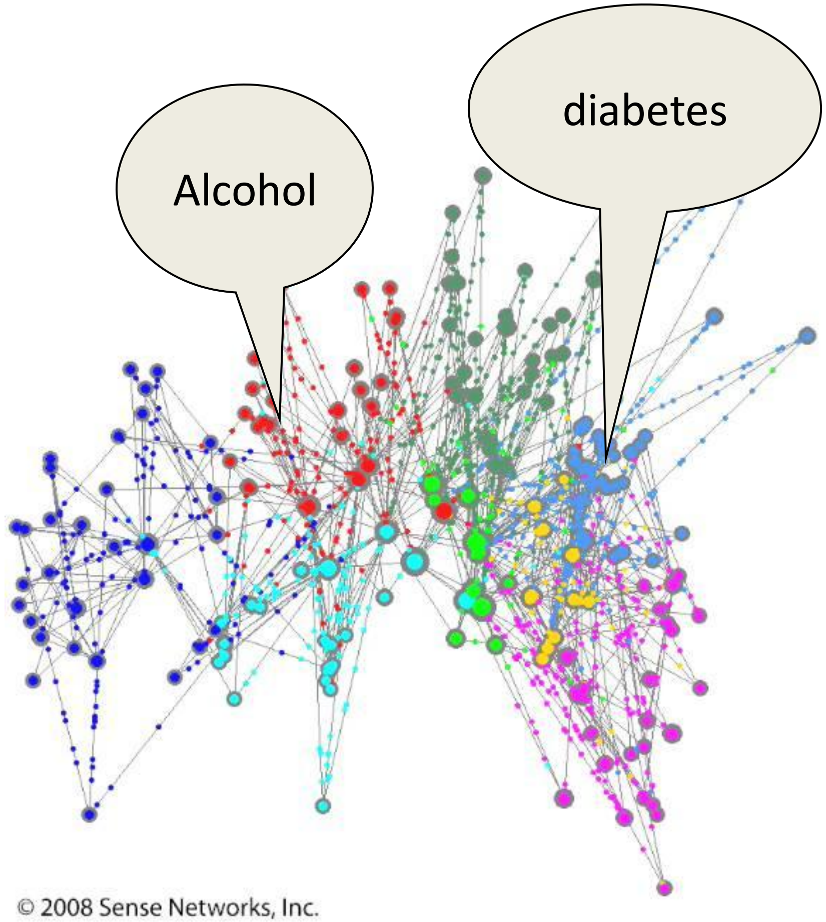
© 2008 Sense Networks, Inc.

Accuracy 4 times normal demographics

90 million people continuously

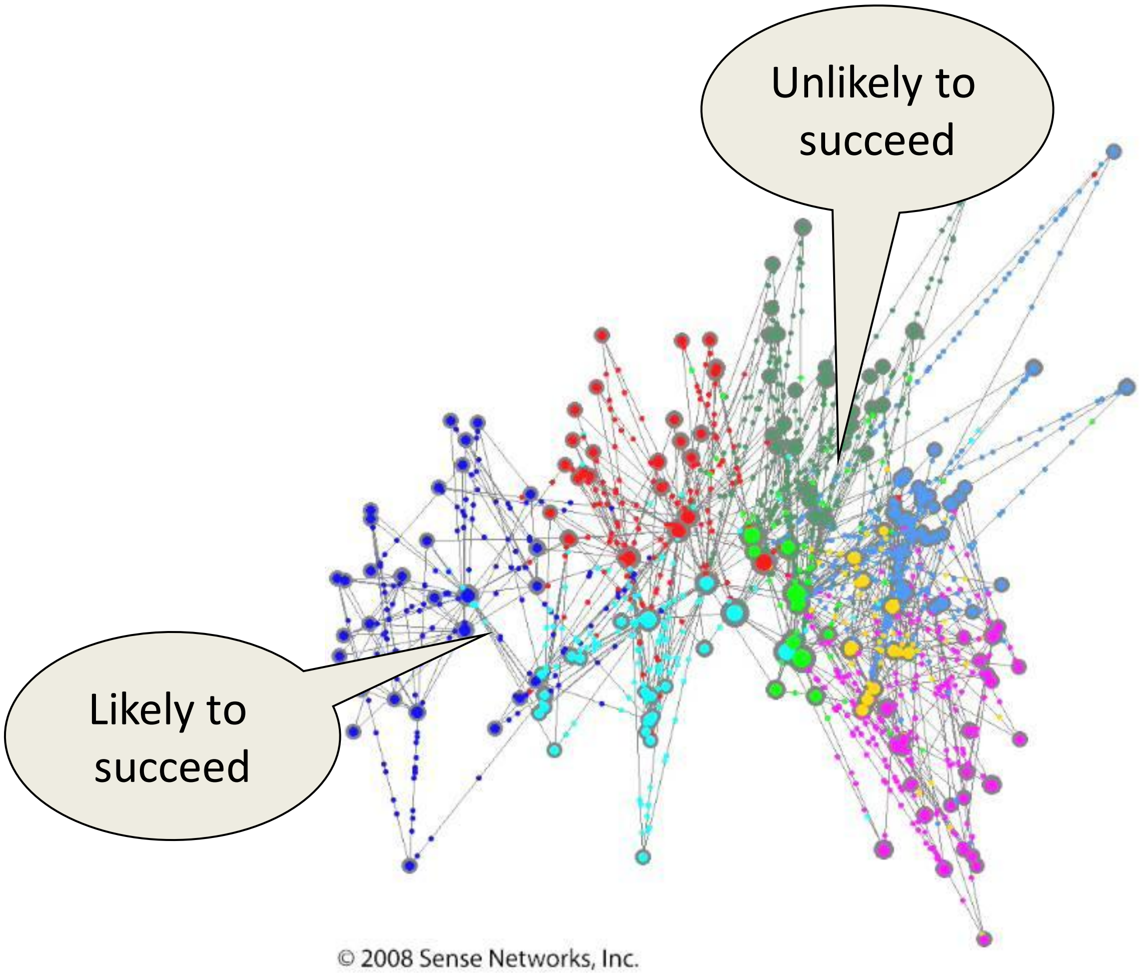
Patterns of Health

With MGH: Phenotypic + Genetic Characterization



Patterns of Finance

Success Scoring of Unbanked

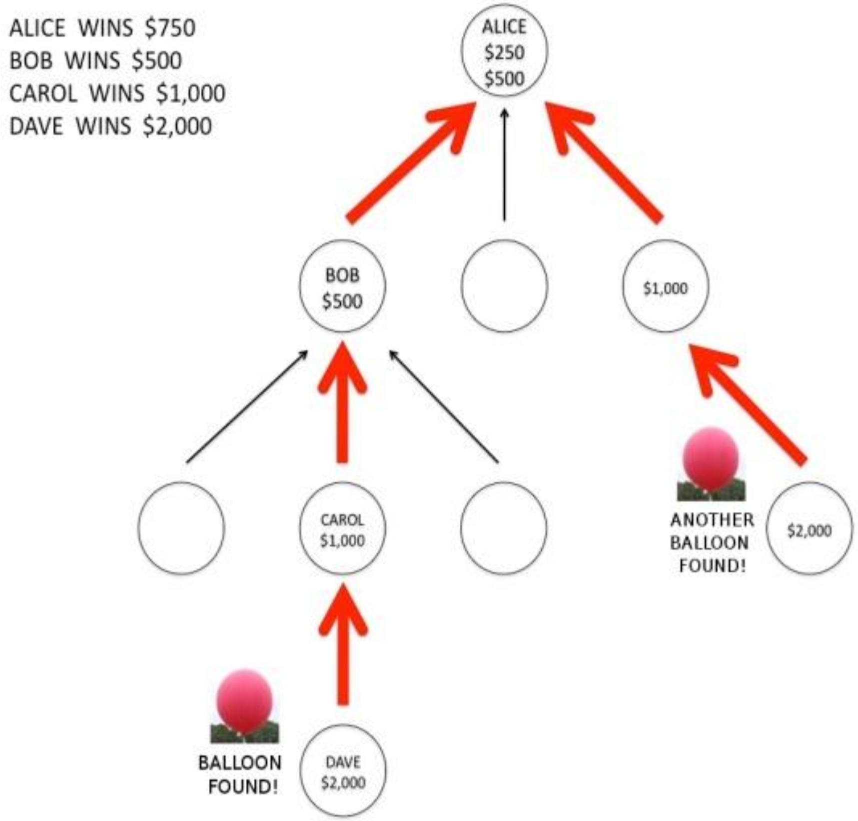


Including life coaching

2. Distributed Intelligence

Shaping By Social Incentives

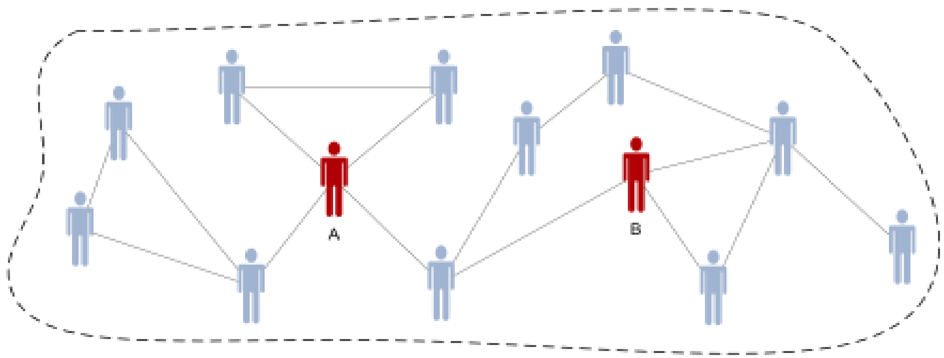
40th Anniversary of the Internet Grand Challenge



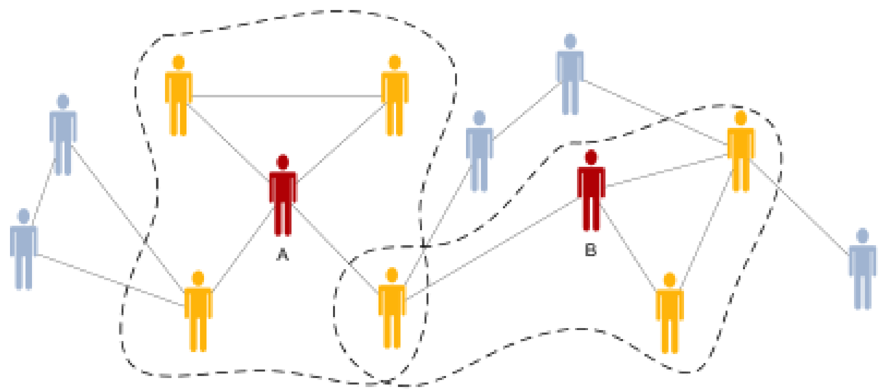
Pickard, Pan, Rahwan,Cebrian,Madan,Crane,Pentland

Shaping by Social Incentives

incentives that leverage social influence



Global externality:
tragedy of the commons



Localized externality:
The peers of individuals A and B receive rewards for behavior of A, B.

Mani, Rahwan, Pentland

Behavior Shaping By Social Influence

- Reward individuals for their peers' behavior

$$U_i(\mathbf{x}, \mathbf{p}) = u_i(x_i) - v_i \left(\sum_{j \neq i} x_j \right) - x_i \sum_{j \in Nbr(i)} p_{ji} - c \sum_{j \in Nbr(i)} p_{ij} + \sum_{j \in Nbr(i)} r_{ji}(x_j)$$

personal utility | externality cost | peer pressure | peer cost | Incentive

- The total reward distributed to the peers of actor *j* is less than the Pigouvian subsidies to *j* if

$$\sum_{i \neq j} \beta_i + \sum_{i \in Nbr(j)} \beta_i > -\alpha_j$$

$$\text{where } \alpha_j = c u_j''(x_j^\circ) \text{ and } \beta_i = v_i' \left(\sum_{k \neq i} x_k^\circ \right)$$

Mani, Rahwan, Pentland

Social Influence incentive mechanism is 3.5 times as efficient as standard incentive mechanism

Condition	Activity Change from Period 1 to Period 3	Reward in Period 3	Reward Efficiency (Δ/$)
Overall			
Control	0.037	$3.00	0.012
Exp 1	0.070	$2.77	0.0253
Exp 2	0.126	$3.04	0.0416
Close Buddies (both Buddies score 3 or higher)			
Exp 1	0.118	$2.68	0.0444
Exp 2	0.269	$3.00	0.0896
Stranger Buddies (both Buddies score 2 or lower)			
Exp 1	-0.007	$2.82	-0.0025
Exp 2	0.137	$2.95	0.0464
Mixed Buddies (one Close, one Stranger)			
Exp 1	0.154	$2.75	0.0560
Exp 2	0.053	$3.12	0.0171

Standard Incentive

Social media Incentive

Peer reward incentive

65 young families, 3 months data

Aharony, Pan, Pentland

3. Network Intelligence

Influence Model

Influence Model

- **Learn the underlying hidden influence network from historical data**
- **Use edge weights (of network) to derive adoption potential**

$$p_a(i) = \sum_{j \in \mathcal{N}(i)} w_{i,j} x_j^a$$

$w_{i,j}$: edge weight between *i* and *j* in the diffusion network

x_j^a = 1 if j has adopted strategy *a*, =0 if not.

Altshuler, Pentland

Behavior Propagation

Behavior Change Model

- Learn the underlying hidden influence network from historical data
- Use edge weights (of network) to derive adoption potential
- Calculate behavior predictions

Diffusion model:

$$P_{Local} = E_u \left[\text{Prob}\left(x_u^a = 1 \mid N(u) \right) \right] = 1 - e^{-s_u - p_a(u)}$$

$$\forall u, s_u \geq 0$$

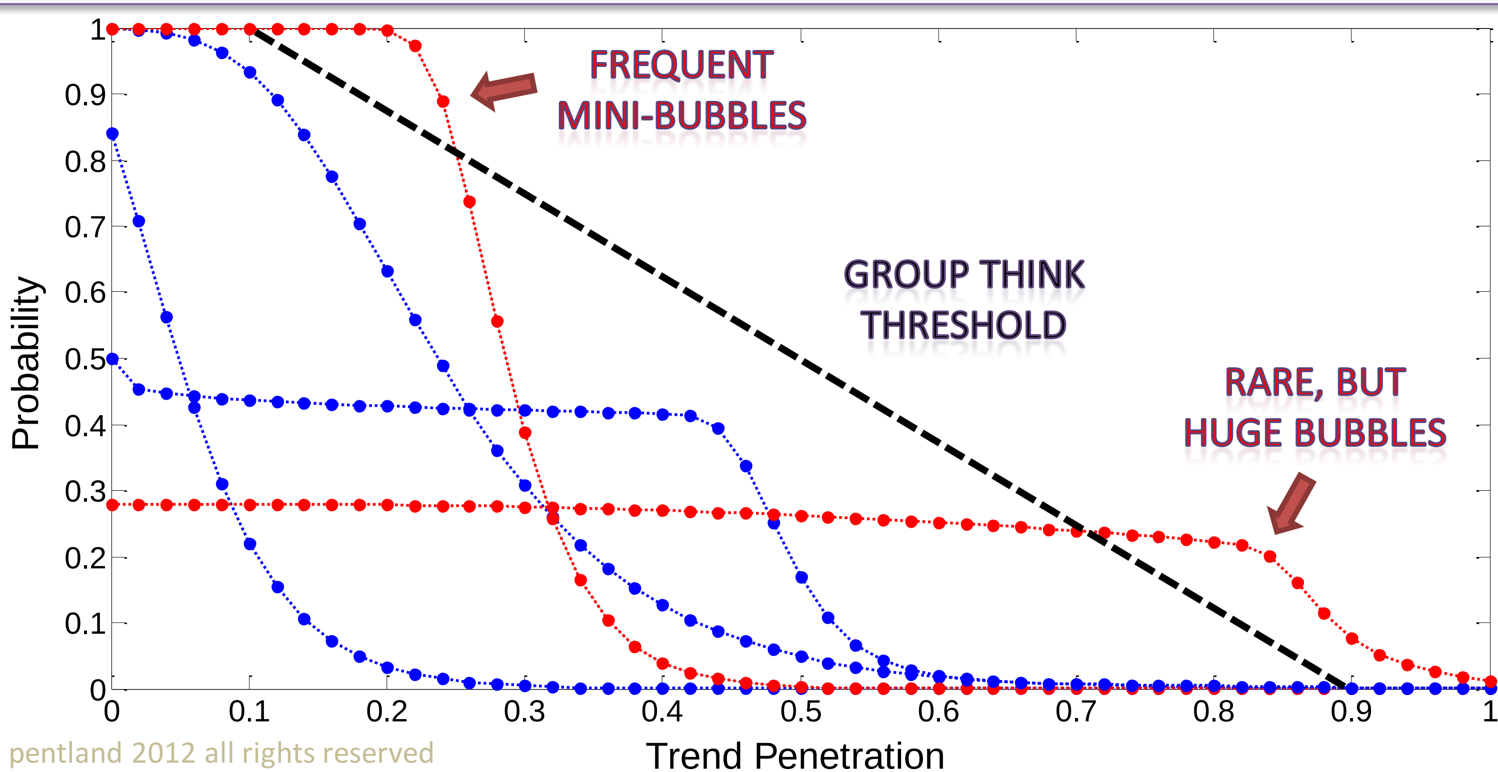
s_u is the individual susceptibility factor of user *u*

Altshuler, Pentland

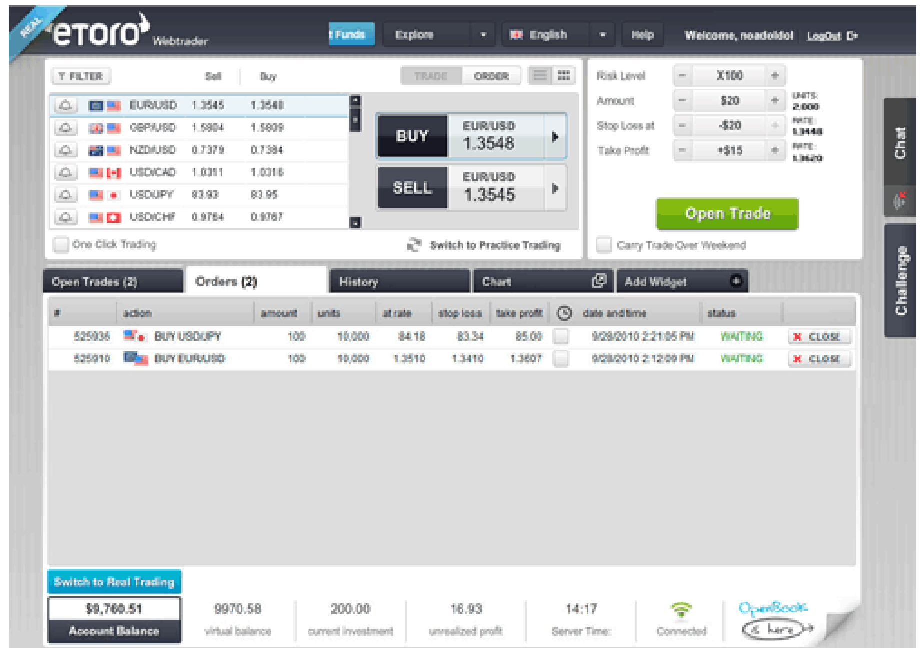
Probability of Idea-Behavior Flow, Φ(C)

Model

- Learn the underlying hidden influence network from historical data
- Use edge weights (of network) to derive adoption potential
- Calculate behavior predictions
- Predict cascade frequency and size, from local influence forces



Φ(C) and the Wisdom of the Crowd



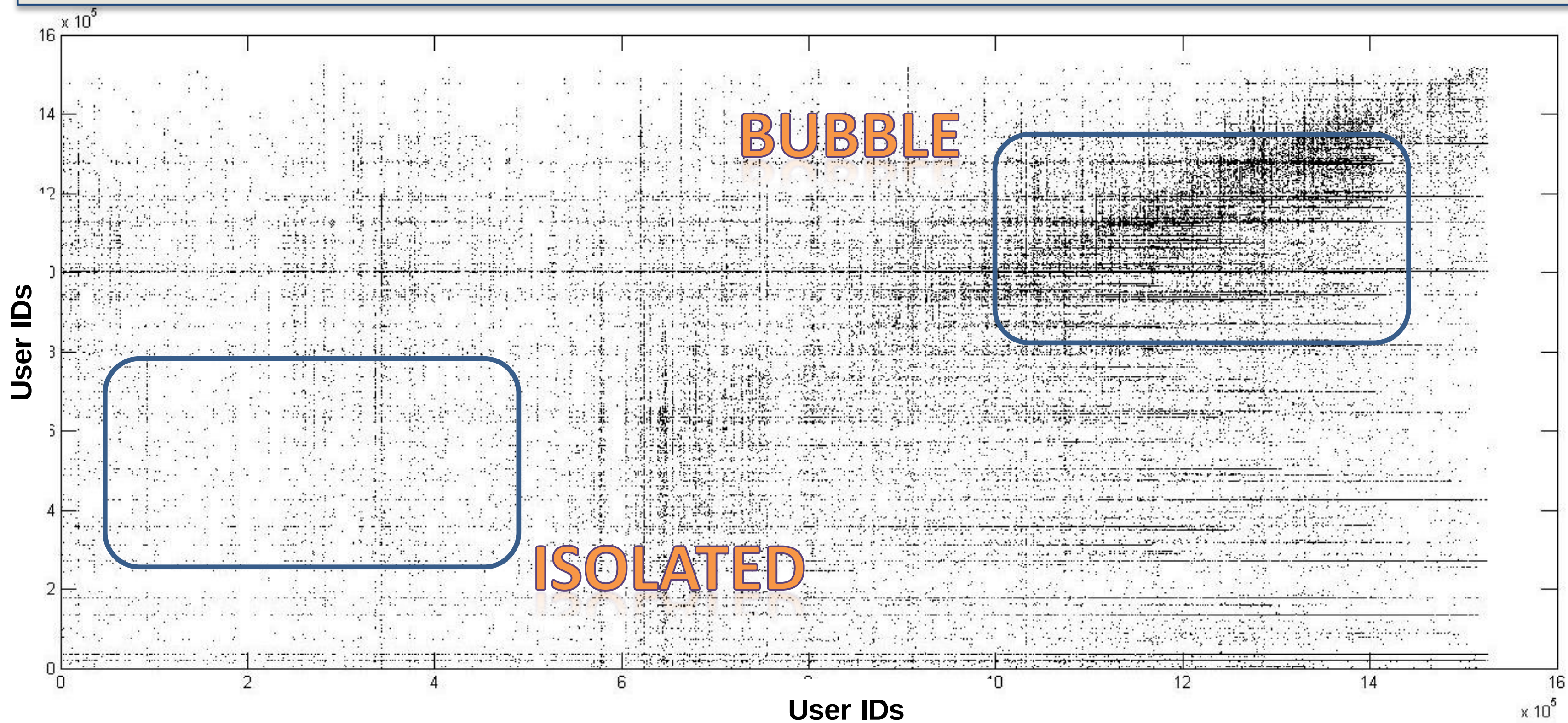
***Social trading*: users can see and copy trades of another user**

Altshuler, Pentland

Isolation and Herding

eToro – Social Trading Network

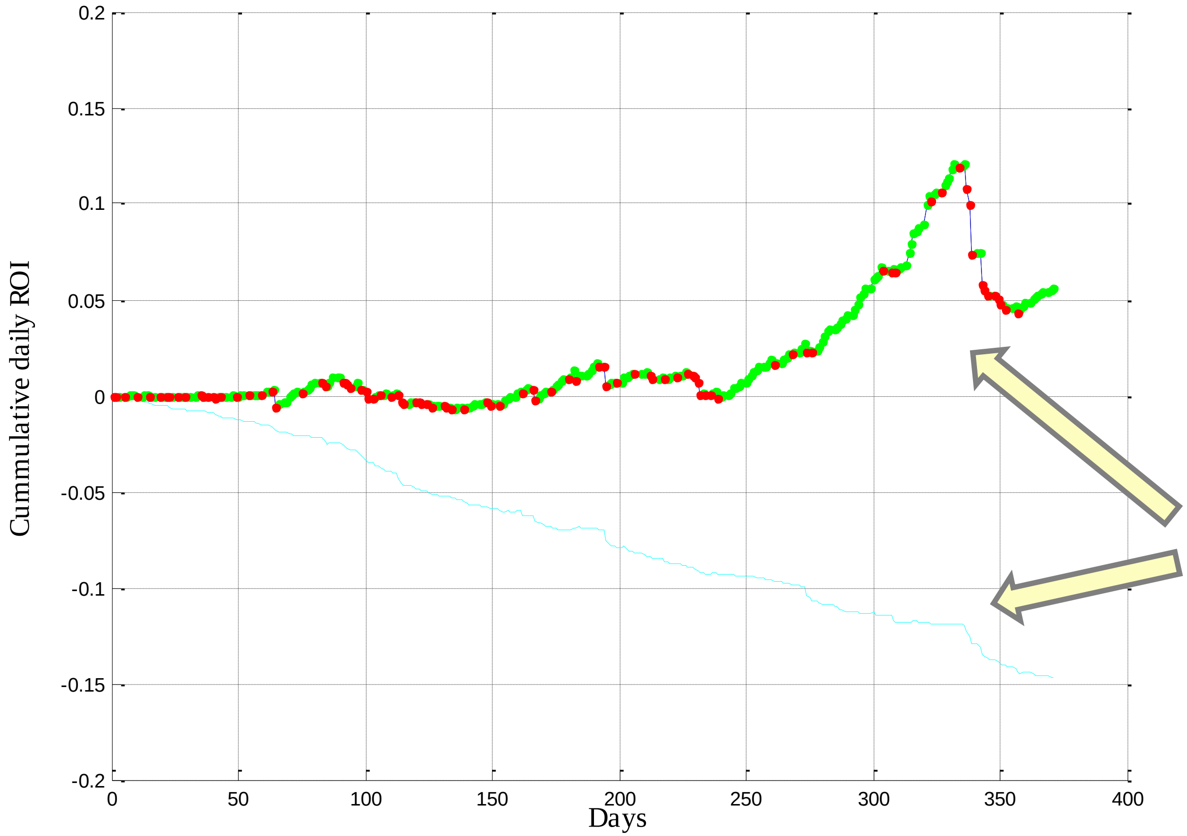
- 2.7 Million users
- "Twitter-like" social based financial trading
- Trading as collaborative problem solving



Altshuler,Pentland

Social Intelligence

Social Trading (Annual ROI)



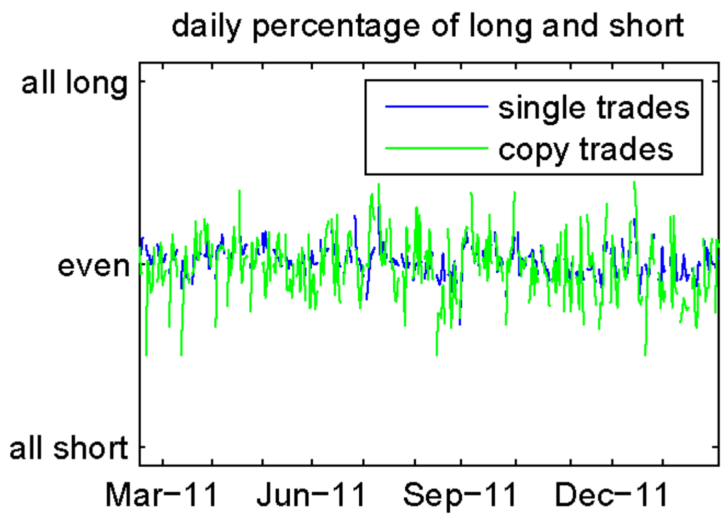
<u>Social Trading</u>

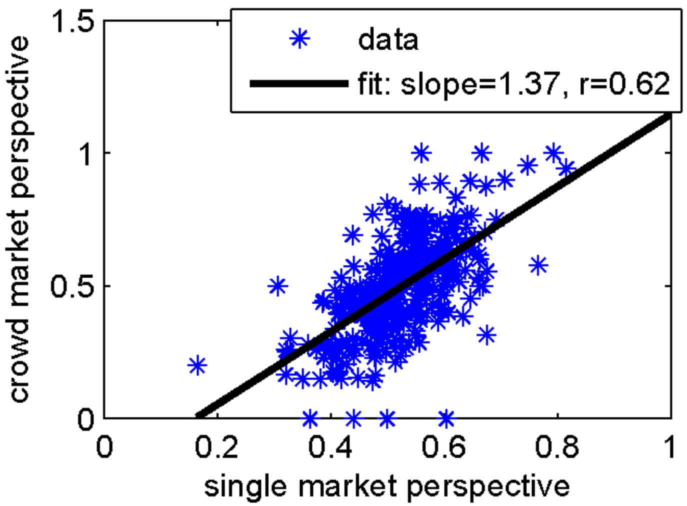
ROI : 5.58%
Max drawdown : -7.81%
Sharpe (yearly) : 1.03
Days win ratio : 0.7

Social Trading
Non-social trading

Altshuler,Pentland

Stupidity of the Crowd



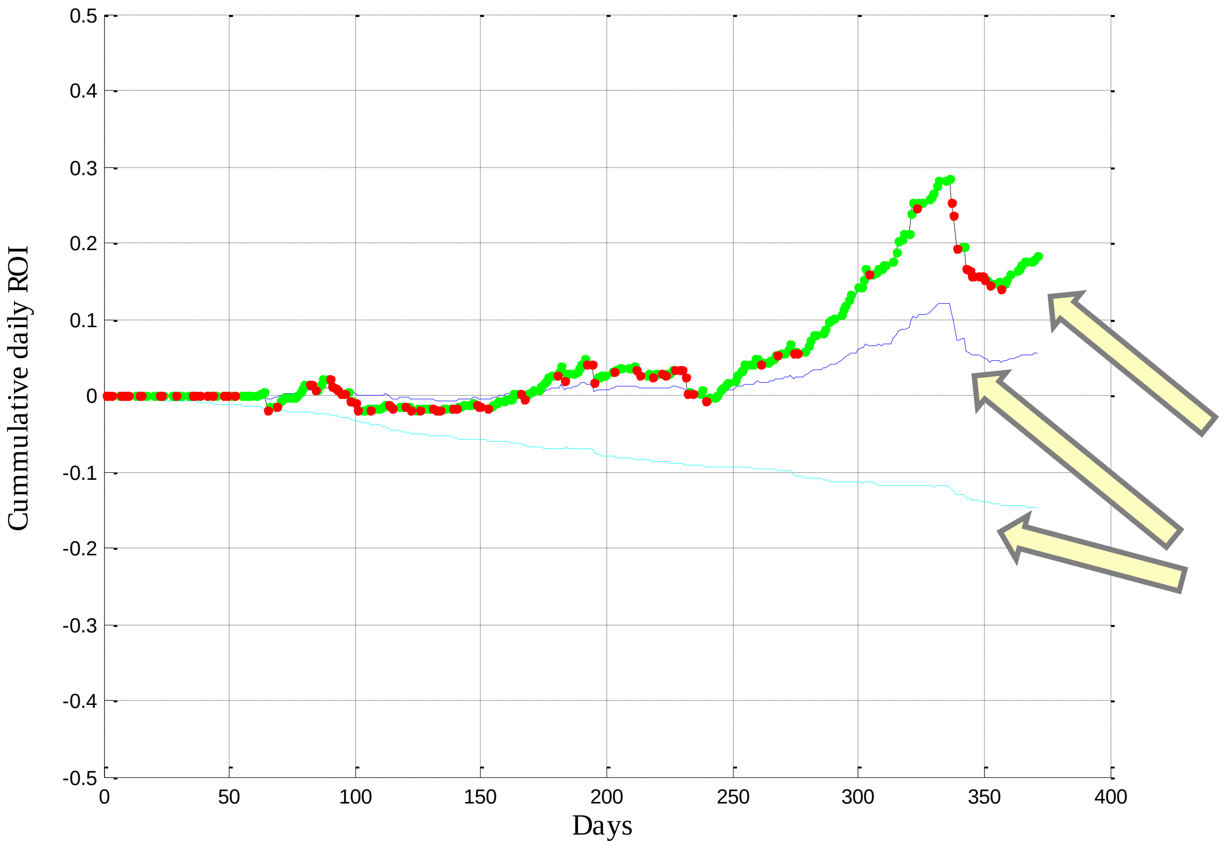


$$
\begin{aligned}
& E(X_i \text{ after consulting } j) \\
&\quad = \quad (p+a) \times \text{Prob}(X_j = 1) \\
&\quad \qquad + (p-a) \times \text{Prob}(X_j = 0) \\
&\quad = \quad (p+a)p + (p-a)(1-p) \\
&\quad = \quad (2a+1)p - a.
\end{aligned}
$$

Pan, Alshuler, Pentland

Experts

Guru Trading (Annual ROI)



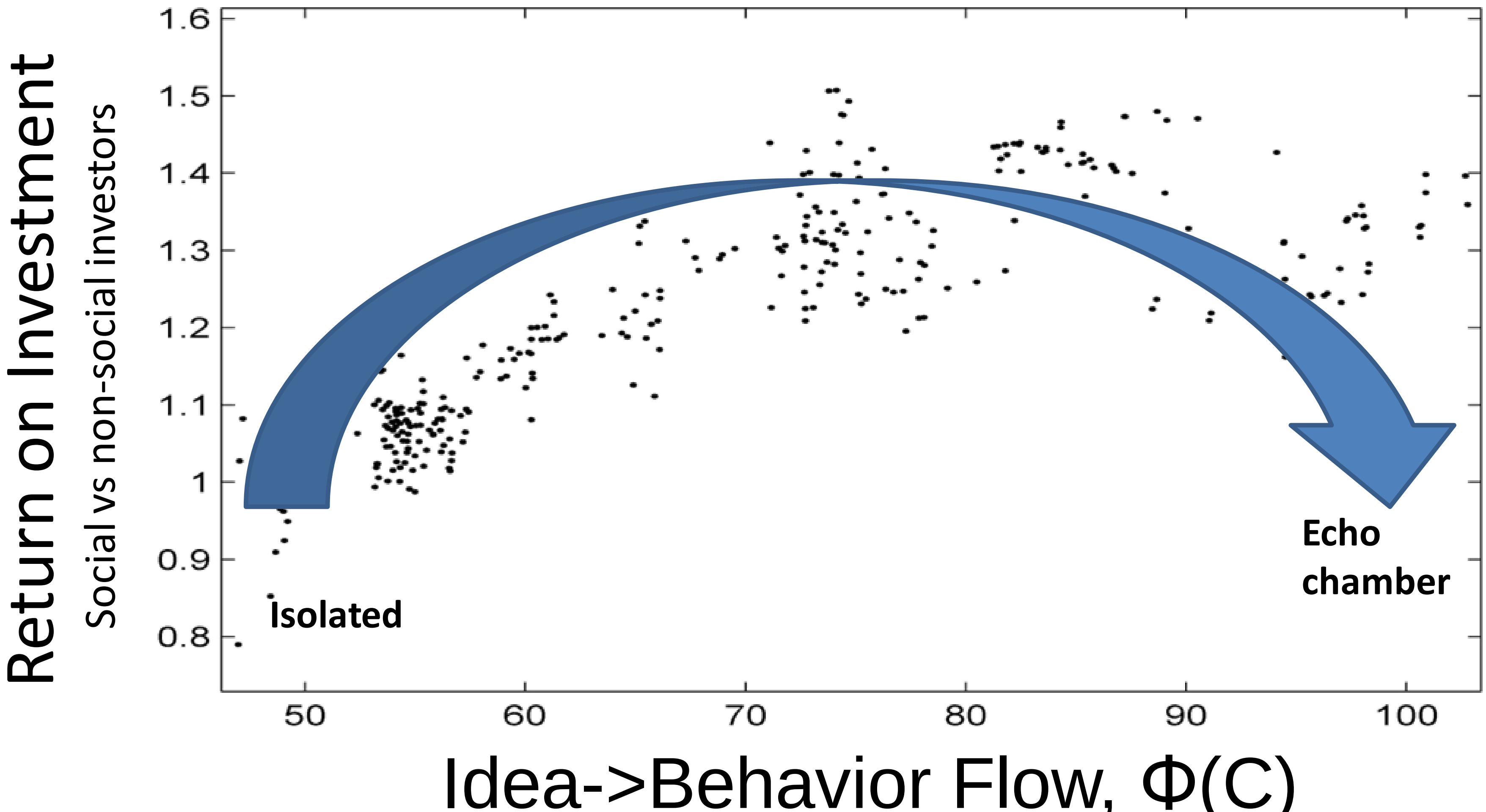
<u>“Guru Trading”</u>

ROI : 18.22%
Max drawdown : -14.49%
Sharpe (yearly) : 1.56
Days win ratio : 0.7

“Guru trading”
Social Trading
Non-social trading

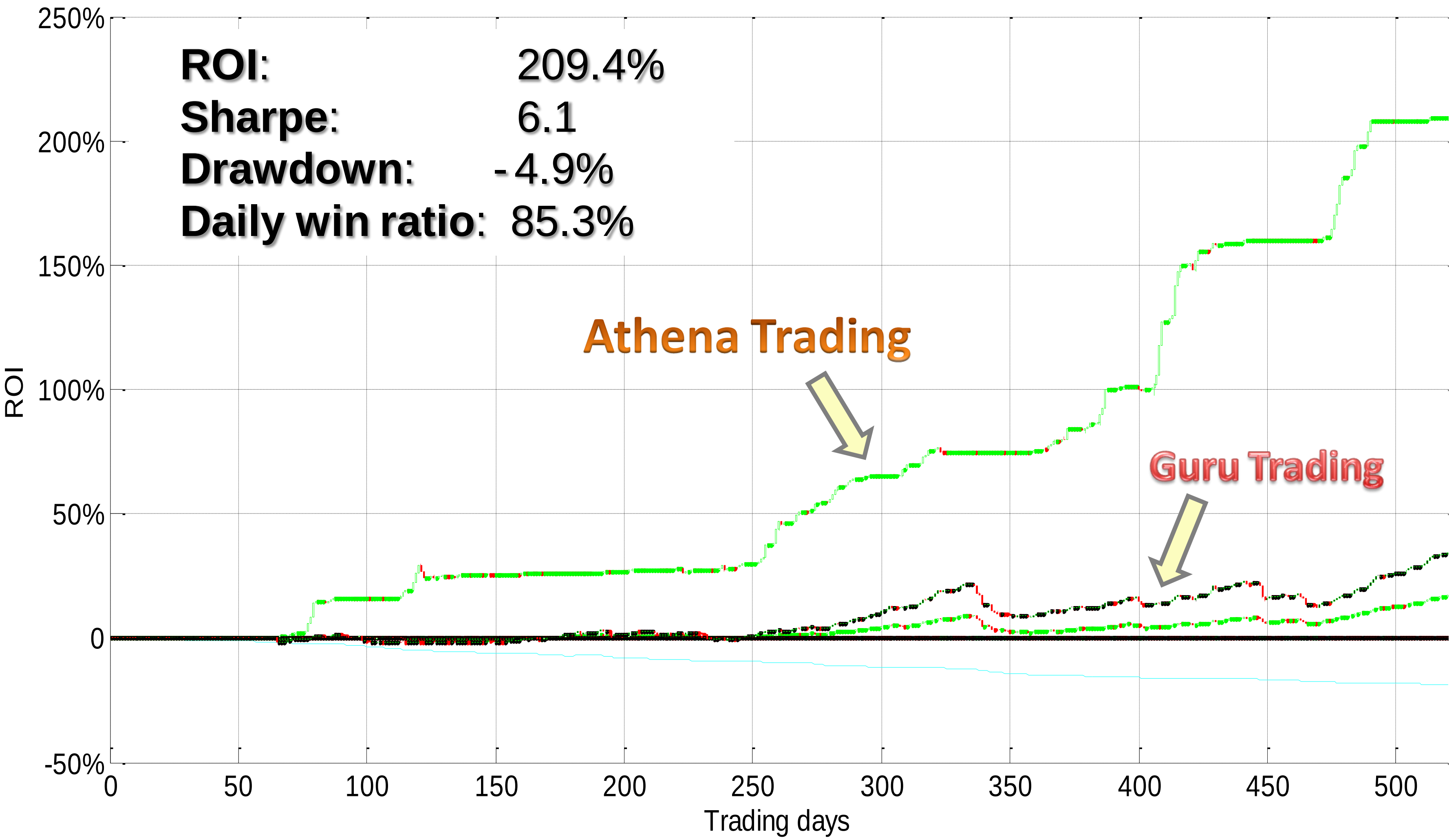
Altshuler,Pentland

Decision Accuracy Depends on Diversity of Information Sources



Altshuler,Pentland

Selecting Oracles

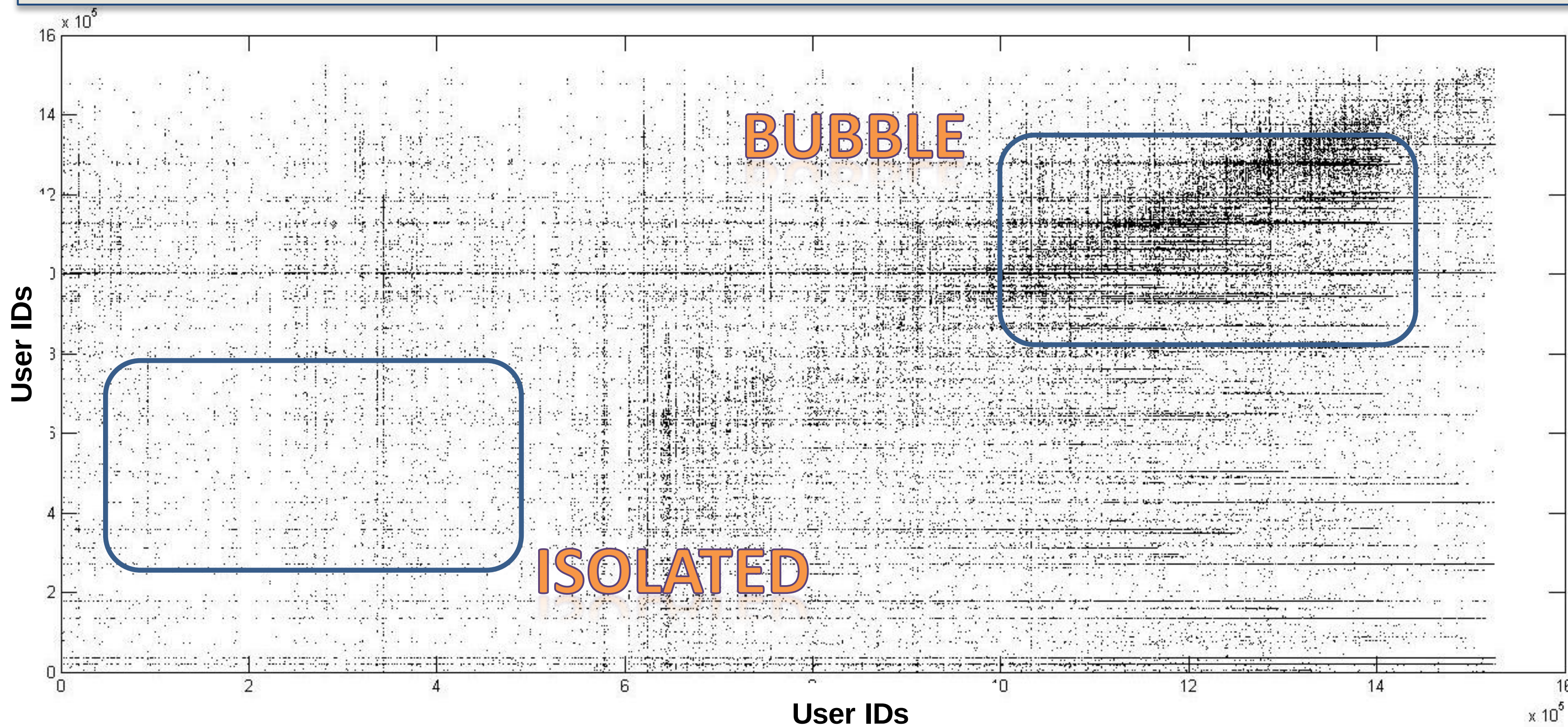


Altshuler,Pentland

Insuring Diversity of Information

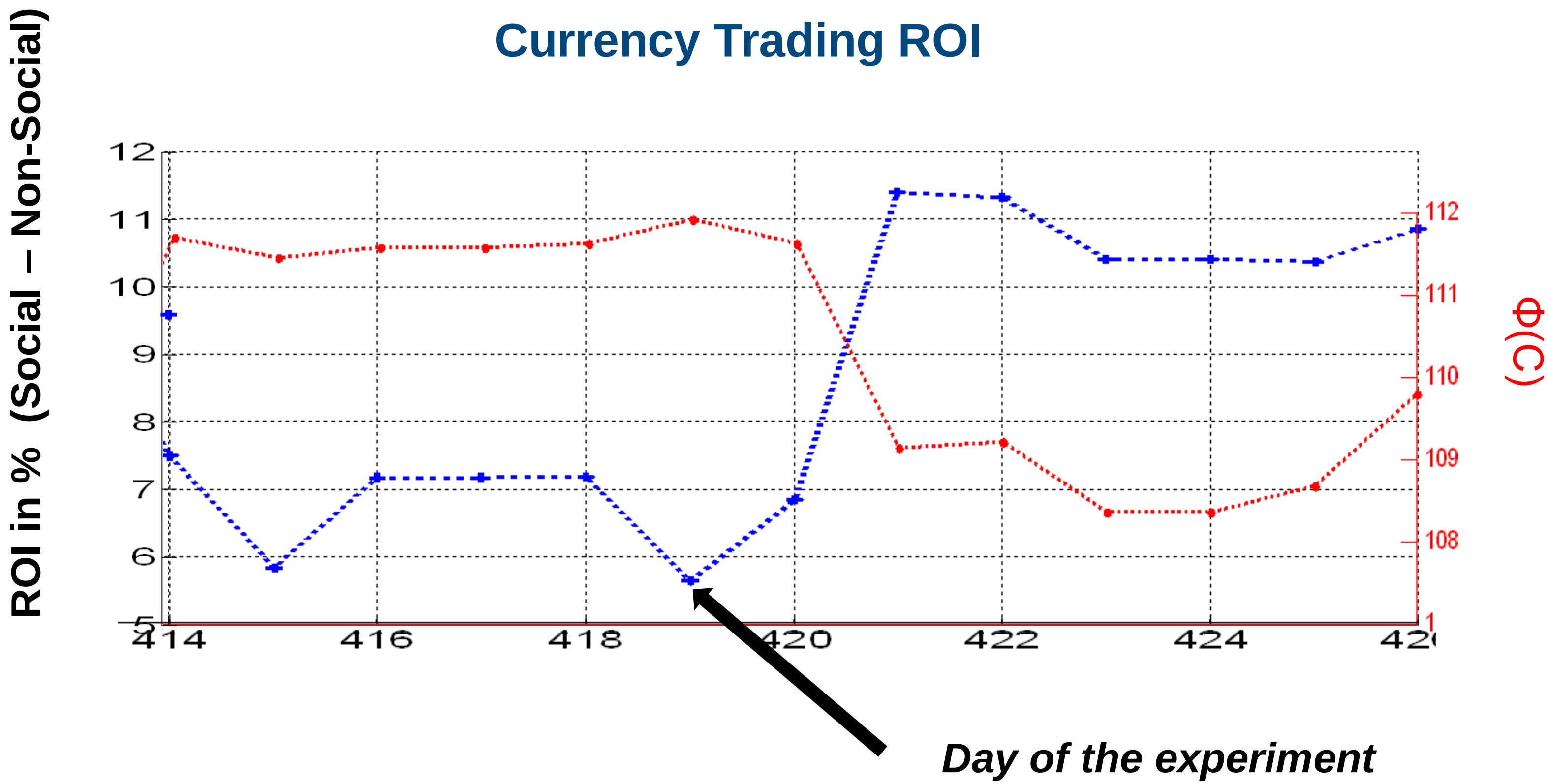
eToro – Social Trading Network

- 2.7 Million users
- "Twitter-like" social based financial trading
- Trading as collaborative problem solving



Altshuler,Pentland

Tune Network to Optimize Φ(C)

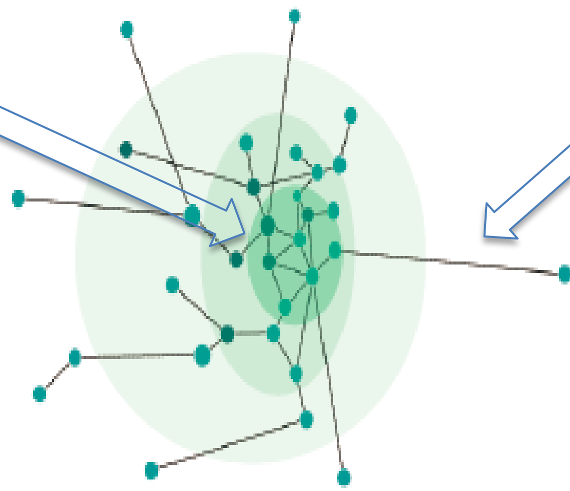


Altshuler,Pentland

Pattern of Social Ties and Φ(C)

Engagement:

Density of sharing
of information
within group



Exploration:

Harvesting New Ideas
outside of group;
`fat tails'

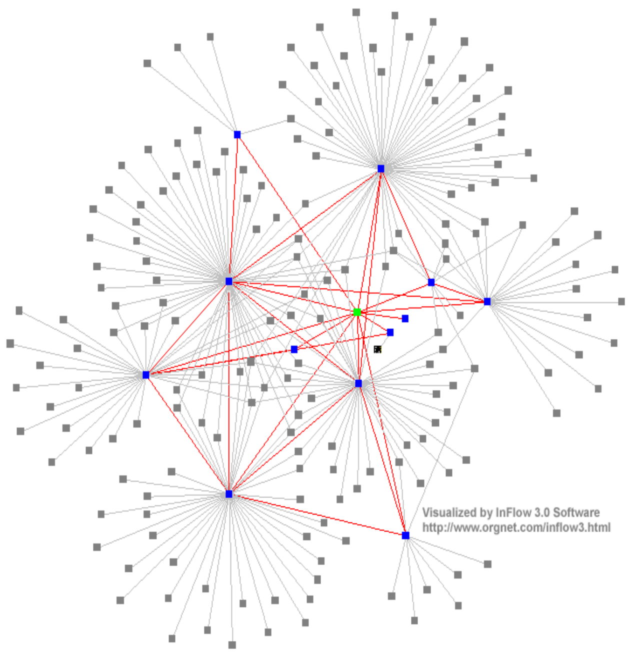
Exploration and Engagement: a study of white collar workers

Engagement in face-to-face accounts for

30% of between-group variation in productivity

Exploration in face-to-face accounts for

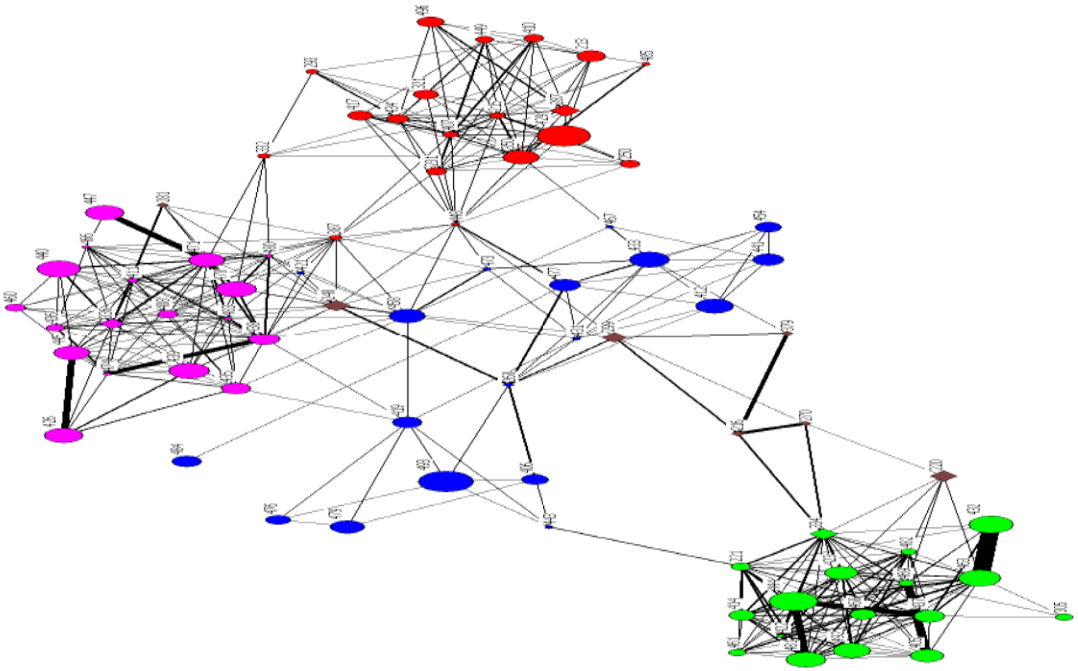
10% of between-group variation in productivity



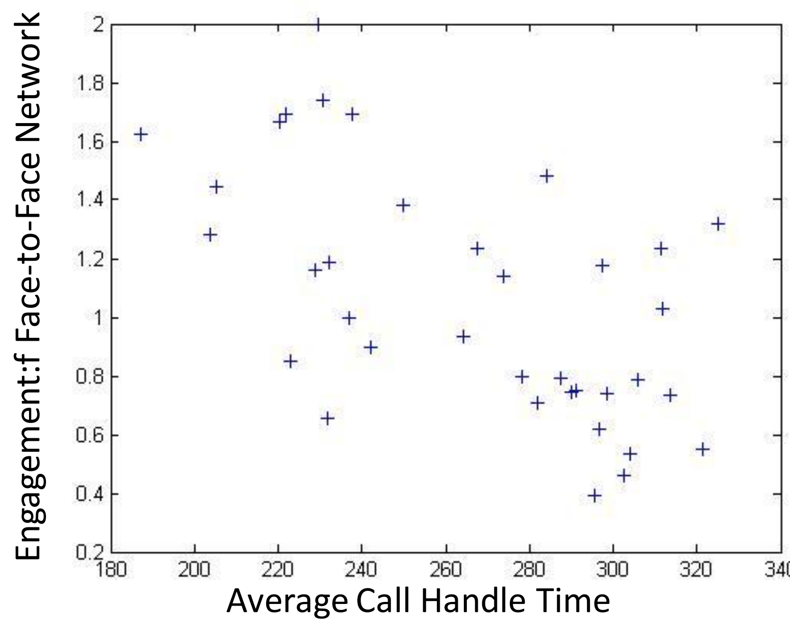
Best Research Paper, ICIS 2008

Wu,Waber,Aral,Brinjolfsson,Pentland

BAC Call Center Productivity Study



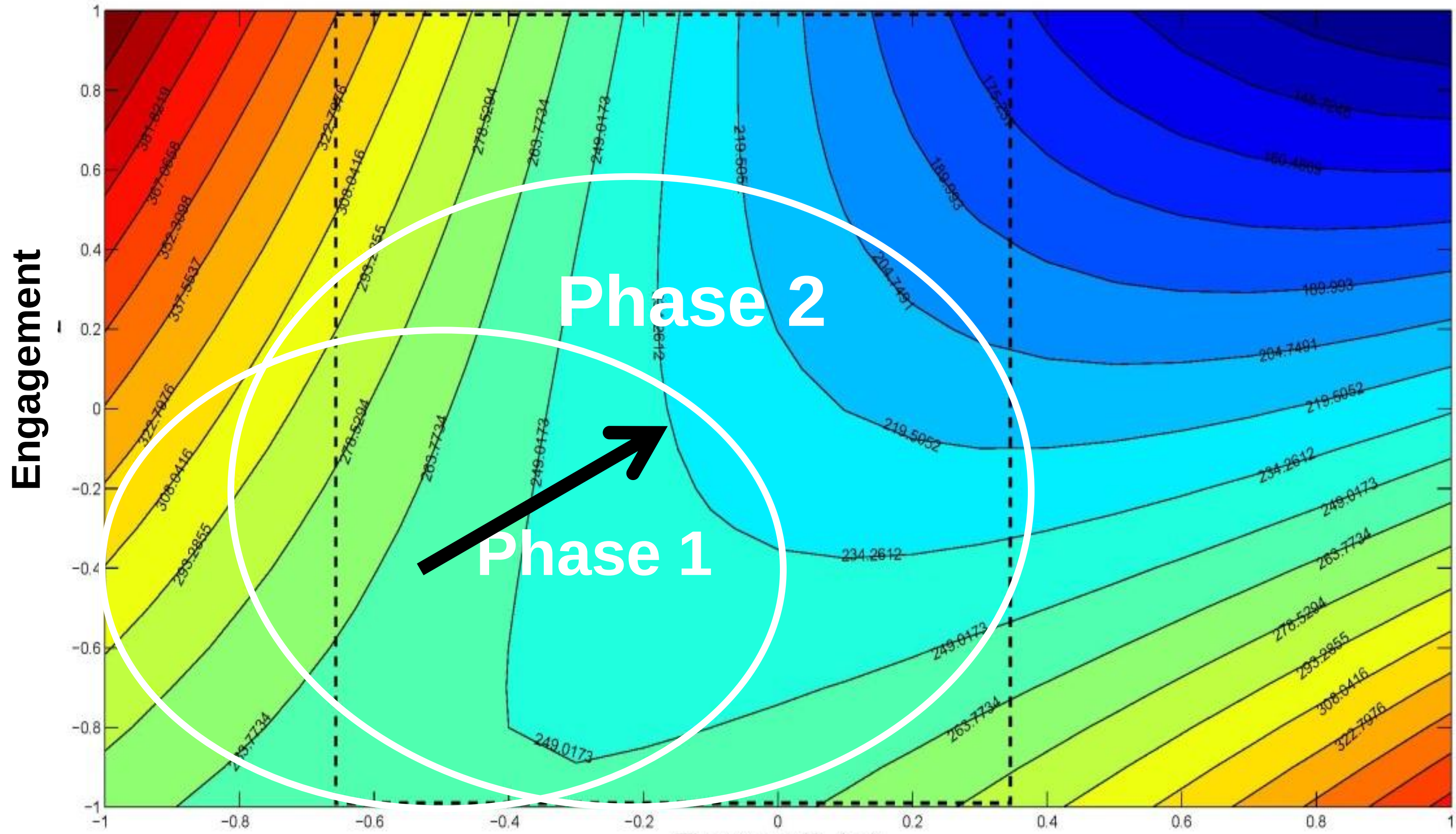
Productivity correlated with group engagement



sociometric solutions

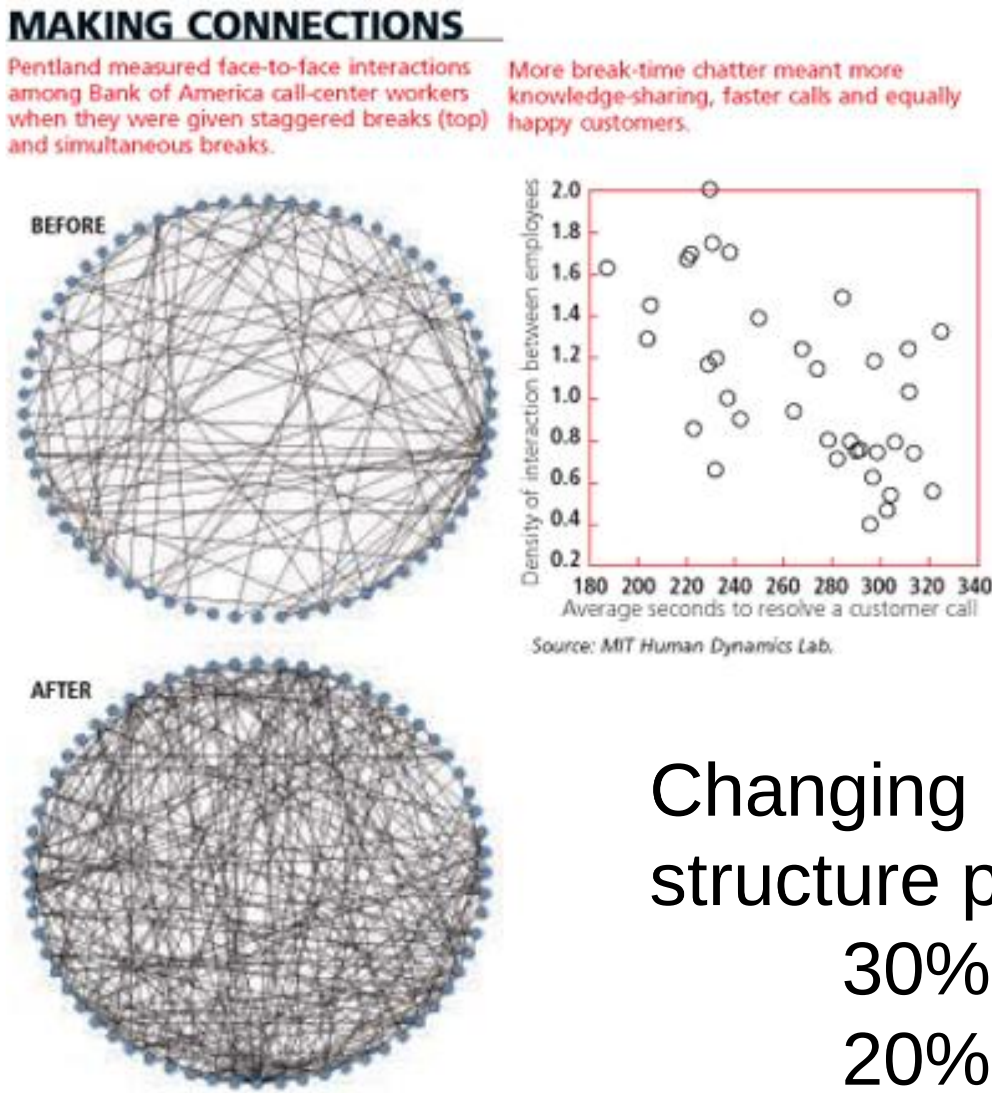
Optimize Idea Flow Φ(C)

Average Call Handling Time



Energy

Olguin, Waber, Kim, Pentland



sociometric solutions

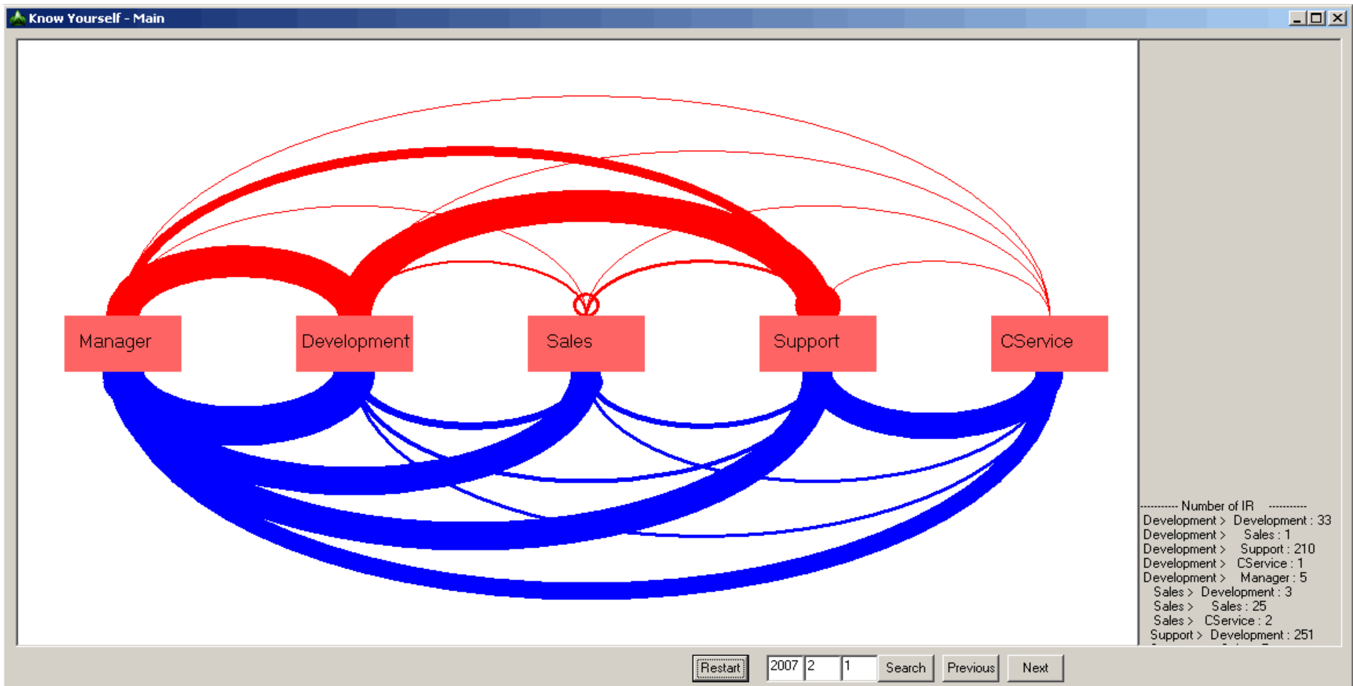
Changing coffee break structure produced:

30% increase engagement

20% decrease stress

$15M / year savings

Φ(C) Measures Are Typically 40% of Performance



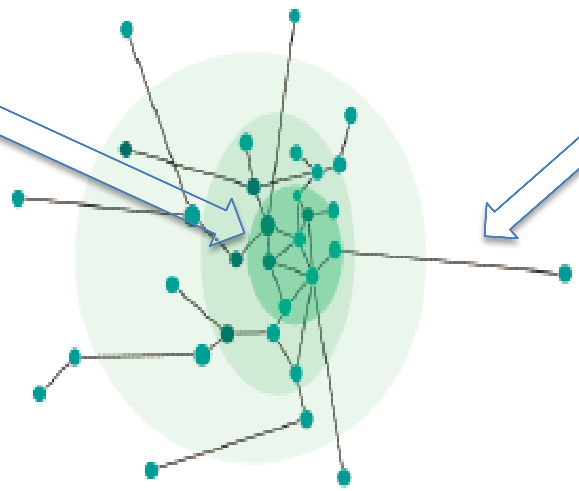
Harvard Business Review:
Breakthrough Idea of the Year

Olguin, Waber, Kim, Gloor, Pentland

Cities and Φ(C)

Engagement:
Density of sharing of information within group

Exploration:
Harvesting New Ideas outside of group

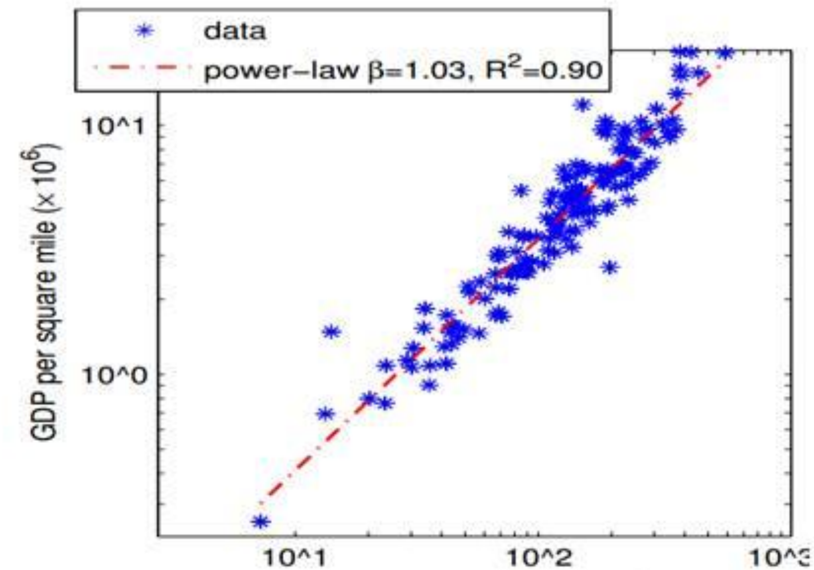


$$P_{ij} \propto \frac{1}{\mathrm{rank}_i(j)},$$

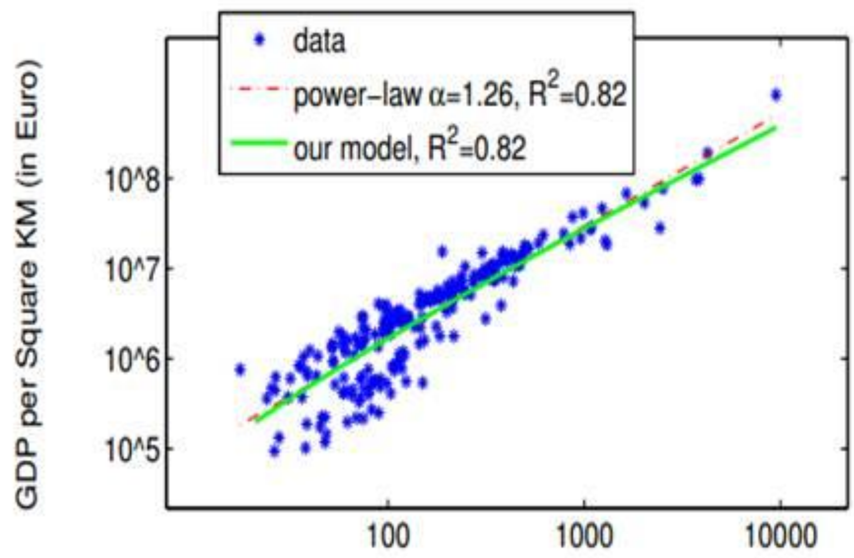
Lieben-Nowell; Krings et al;

Φ(C) and GDP

- EU and US Cities GDP vs Social Tie Pattern



Φ(G)



Φ(G)

Pan, Krumme, Ghoshal, Cebrian, Pentland

4. Big Data breaks science

Science as practiced assumes
strong theoretical understanding

Big Data is good for interpolation
but not for extrapolation

Big Data governance *requires* thousands of social *science* experiments

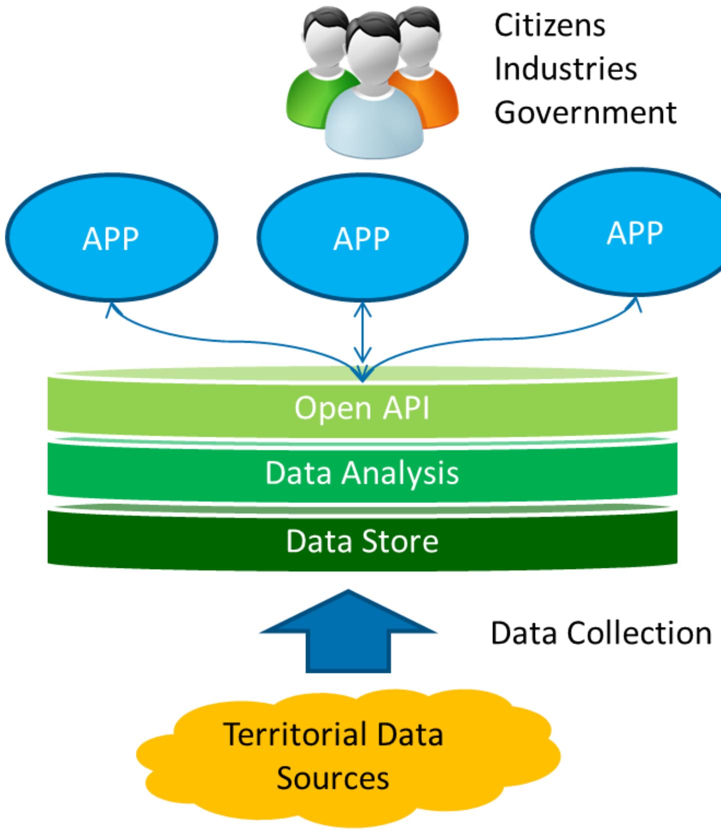
mtl
Mobile Territorial Lab

HOMEPAGE | JOIN IN | PROJECT | ABOUT US | CONTACT

TELECOM ITALIA
FONDAZIONE BRUNO KESSLER
MIT MEDIA LAB
Telefónica Digital
ID idcubed.org

Data from private companies and Provential Authority

Trentino Open Living Data Project (TOLD)



Application scenarios:

- **Mobility**:
  - Online efficient private traffic
  - Public transportation on the fly route balancing
- **Safety**:
  - Detection and support in dangerous situations (e.g. fires, avalanches, etc.)
- **Health**:
  - Recognition and prediction of epidemic spread
- **Urban & Local business planning**:
  - Understand economically depressed areas
  - Help companies to plan investment

- **A joint project between**

TELECOM ITALIA, trentorise, ENGINEERING, Massachusetts Institute of Technology

Data from individuals

Mobile Territorial Lab

A joint project between:

TELECOM ITALIA

Telefonica

Massachusetts Institute of Technology

trentorise
Open Innovation rooted in Research and Education

- Understand the needs and the behaviour of users.

- Provide individuals mobile phone equipped with a sensing middleware to collect the data generated to be analyzed (starting community: young families with newborns)

- Short term outcomes:

1. Developing and testing **a new model of DATA OWNERSHIP**

2. Understanding the **dynamics of people's needs**

1. Understanding **people's interactions** in the generated social networks

pentland@mit.edu

http://media.mit.edu/~pentland

Forbes, 8/10,
Mining Human Behavior